



ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Mécanique

Abdoul Magid Amadou Sanoko, Simon Essongue, Joseph Paux and Léo Morin

Reformulation and influence of discretization in Fourier-series-based FFT methods for heterogeneous materials

Volume 354 (2026), p. 633-662

Online since: 26 June 2026

<https://doi.org/10.5802/crmeca.369>



This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*The Comptes Rendus. Mécanique are a member of the
Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1873-7234



Research article

Reformulation and influence of discretization in Fourier-series-based FFT methods for heterogeneous materials

Abdoul Magid Amadou Sanoko ^{a,b}, Simon Essongue ^{a,b}, Joseph Paux ^c and Léo Morin ^{*,a,b}

^a Univ. Bordeaux, CNRS, Bordeaux INP, I2M, UMR 5295, F-33400 Talence, France

^b Arts et Metiers Institute of Technology, CNRS, Bordeaux INP, I2M, UMR 5295, F-33400 Talence, France

^c Université Sorbonne Paris Nord, Laboratoire des Sciences des Procédés et des Matériaux, LSPM, CNRS, UPR 3407, F-93430, Villetaneuse, France

E-mails: abdoul-magid.amadou-sanoko@u-bordeaux.fr (A.M. Amadou Sanoko),
simon.essongue-boussougou@u-bordeaux.fr (S. Essongue), joseph.paux@univ-paris13.fr (J. Paux),
leo.morin@u-bordeaux.fr (L. Morin)

Abstract. The aim of this work is to revisit the formulation of FFT-based methods for heterogeneous materials in the periodic setting. These numerical methods are based on the iterative resolution of an auxiliary problem involving a reference homogeneous material and a polarization tensor. Assuming a description of the fields by Fourier series, we show the equivalence between three discretization approaches based either on the strong and the weak formulation of the problem. A special emphasis is put on the representation of the local fields described by Fourier series including the possibility of using non-uniform grid. Numerical experiments are performed on a model problem of conductivity with a checkerboard microstructure for which an analytical solution allows to assess the effect of Fourier modes together with the grid discretization. The occurrence of oscillations is finally addressed by studying (generalized) discrete Green operators (still in the context of Fourier series) and interface spreading approaches based on smoothing techniques.

Keywords. FFT-based solvers, Strong and weak formulations, Fourier series, Green operators.

Manuscript received 27 February 2026, revised 11 May 2026, accepted 12 May 2026, online since 26 June 2026.

1. Introduction

FFT-based methods for heterogeneous materials, introduced in the seminal article of Moulinec and Suquet (1998) [1], are a class of efficient methods that have been developed in the context of periodic homogenization of mechanical properties in heterogeneous microstructures, mainly for *micromechanics* (see the review papers of [2] and [3]). This method is based on an iterative solution of the so-called Lippmann-Schwinger equation using fast Fourier transforms (FFT). The main characteristics of these methods are that (i) calculations are based on the computational complexity of fast Fourier transforms in $O(N \log N)$, (ii) it is image-based (obtained from imaging techniques such as scanning electron microscopy (SEM) or tomography) so it

* Corresponding author

does not require meshing operations and (iii) it can be suited for parallel implementation (see e.g. [4]). It is generally used to calculate the local and overall fields, and has been applied to various problems in heterogeneous microstructures including notably [1], J2-plasticity [1], crystal (elasto)viscoplasticity [5], dislocation-mediated plasticity [6,7], conductivity [8], piezoelectricity [9], porous ductile solids [10,11].

The original method developed by Moulinec and Suquet (1998) [1] is implicitly based on the resolution of the strong formulation (of the equations of elastostatics) using infinite Fourier series. It was then shown by [12] the formulation of [1] is equivalent to a Galerkin method using trigonometric polynomials as interpolation basis functions. The use of a Galerkin method based on a weak formulation instead of a strong formulation is particularly important in extensions of FFT-based methods to non-periodic boundary conditions using sine and cosine series [13–17].

The scheme of [1] relies on the introduction of a Green operator which can be expressed easily in Fourier space using Fourier transforms and standard spectral derivative rule, defining the so-called *continuous* Green operator; the terminology “continuous” has been notably introduced by [18], to highlight its differences with other Green operators constructed using different derivative rules and/or different formulations. In the work of [18], a finite difference scheme is used to approximate the partial derivatives defining the local elliptic problem (see also [19]) and leads to the so-called *discrete* Green operator that was shown to improve both the convergence of the iterative procedure and the local fields. It was also shown in several works that the use of uniform finite elements on periodic Cartesian grids leads to similar discrete Green operators depending on the type of elements considered [20,21]: the FFT solver can then be seen as a preconditioner of the linear system arising in the FEM discretization.

The aim of this work is to revisit several aspects of the method of [1]. First, we compare several techniques of the numerical resolution of the auxiliary problem, either based on strong or weak formulations, by describing explicitly the unknown fields using (infinite or partial) Fourier series following initial ideas of [12]; this allows to clearly show the contribution of the number of modes, aliasing and underintegration in the procedure of resolution which are classical topics in the spectral methods literature [22,23]. As the fields are described by Fourier series, a particular emphasis is put on the representation of solution fields between grid nodes with the possibility of using non-uniform sampling: the occurrence of spurious oscillations is carefully assessed especially between grid nodes. To address the possible remedies for oscillations, we derive, still in the context of Fourier series, generalized Green operators that are based on modified derivative laws which allows to retrieve the so-called continuous and discrete operators [18]. The importance of “spreading” (i.e. smoothing material properties) [24] is also assessed in order to remove these spurious oscillations.

The paper is organized as follows. In Section 2, the framework of Fourier series and transforms is recalled as well as the physical problem of conductivity considered throughout the article. The numerical discretization of the auxiliary problem is then done in Section 3, based either on strong and weak formulations together with numerical approximations. The representation of solution fields on an arbitrary grid is then addressed in Section 4. Numerical experiments are performed in Section 5 on a problem of conductivity with a checkerboard microstructure. Finally, remedies for oscillations, i.e. discrete Green operators and smoothing techniques, are discussed in Section 6.

2. Preliminaries

2.1. Equations of conductivity

We consider a problem of conductivity (as a model problem for FFT-based solvers) which consists in the computation of the electric field $\mathbf{E}(\varphi(\mathbf{x}))$, current $\mathbf{J}(\mathbf{x})$ and electric potential $\varphi(\mathbf{x})$, at each

point $\mathbf{x}$ in Ω , for a given conductivity field $\mathbf{c}(\mathbf{x})$. The unit-cell is a square or cubic domain $\Omega = [-L/2, L/2]^d$ ($d = 2$ or 3) with L the period of the microstructure. Tensorial components refer to a system of Cartesian coordinates $(\mathbf{e}_1; \mathbf{e}_2)$ in 2D and $(\mathbf{e}_1; \mathbf{e}_2; \mathbf{e}_3)$ in 3D. Periodic boundary conditions are considered and the local electric field $\mathbf{E}$ is split into its average $\bar{\mathbf{E}} = \langle \mathbf{E} \rangle$ (where $\langle \cdot \rangle$ denotes the spatial average over Ω) and a fluctuation term $\mathbf{E}(\varphi^*(\mathbf{x}))$:

$$\mathbf{E}(\varphi(\mathbf{x})) = \mathbf{E}(\varphi^*(\mathbf{x})) + \bar{\mathbf{E}} \quad \text{or equivalently} \quad \varphi(\mathbf{x}) = \varphi^*(\mathbf{x}) - \bar{\mathbf{E}} \cdot \mathbf{x}. \quad (1)$$

Periodic boundary conditions imply that the fluctuating part φ^* is periodic (notation: $\varphi^* \#$) and that the term $\mathbf{J} \cdot \mathbf{n}$ is anti-periodic on the boundary between two neighboring cells with $\mathbf{n}$ the outer normal along the boundary $\partial\Omega$ of Ω (notation: $\mathbf{J} \cdot \mathbf{n} - \#$).

The equations of conductivity are

$$\begin{cases} \operatorname{div} \mathbf{J}(\mathbf{x}) &= 0 \\ \mathbf{J}(\mathbf{x}) &= \mathbf{c}(\mathbf{x}) : (\mathbf{E}(\varphi^*(\mathbf{x})) + \bar{\mathbf{E}}) \\ \mathbf{E}(\varphi^*(\mathbf{x})) &= -\nabla \varphi^*(\mathbf{x}), \end{cases} \quad (2)$$

where $\mathbf{c}(\mathbf{x})$ is the local second-order conductivity tensor. In the particular case of isotropic conductivity, it reads

$$\mathbf{c}(\mathbf{x}) = c(\mathbf{x}) \mathbf{I}_d \quad (3)$$

with $c(\mathbf{x})$ the local (scalar) conductivity field and $\mathbf{I}_d$ the second-order identity tensor. Hereafter we will consider isotropic constituents for simplicity (although the method is designed for arbitrary anisotropic behaviors).

2.2. Fourier series and the discrete Fourier transform

We recall important results related to Fourier series and the associated discrete Fourier transforms.

Fourier series: the 1-d case. We consider a periodic function f with period L . Such function can be described by the Fourier series

$$f(x) = \sum_{i=-\infty}^{+\infty} F_i \exp(\iota \xi_i x), \quad \forall x \in [0, L], \quad (4)$$

where ι is the imaginary unit and ξ_i is some “frequency” parameter that is given by

$$\xi_i = 2\pi \frac{i}{L}, \quad i = -\infty, \dots, +\infty, \quad (5)$$

and F_i are the Fourier coefficients given by

$$F_i = \frac{1}{L} \int_0^L f(x) \exp(-\iota \xi_i x) dx. \quad (6)$$

It is interesting to note that the (first) derivative of f is the Fourier series

$$f'(x) = \sum_{i=-\infty}^{+\infty} \iota \xi_i F_i \exp(\iota \xi_i x), \quad \forall x \in [0, L]. \quad (7)$$

In addition, the associated partial (i.e. truncated) series of order $M+1$ (for every positive even integer M) is defined by

$$f(x) = \sum_{i=-M/2}^{M/2} F_i \exp(\iota \xi_i x), \quad \forall x \in [0, L]. \quad (8)$$

Discrete Fourier transform: the 1d-case. The computation of the Fourier coefficients F_i can be done approximately using the discrete Fourier transform (DFT). An efficient calculation of the DFT is classically performed using fast Fourier transforms (FFT) [25].

The 1-d domain of size $[0, L]$ is discretized uniformly with $N + 1$ segments so the spatial scale associated with the uniform grid is $\Delta x = L/(N + 1)$ and we denote by $x_a = a\Delta x$ (for $a = 0, \dots, N$) the grid points. The grid points values of function f are denoted by $f_a = f(x_a)$ (for $a = 0, \dots, N$) and may thus be written as a vector $\mathbf{f}$ of size $N + 1$ as

$$\mathbf{f} = (f_0 \ f_1 \ \dots \ f_{N-1} \ f_N) \quad (9)$$

The discrete Fourier transform of $\mathbf{f}$, denoted by $\text{DFT}(\mathbf{f})$, is the vector $\hat{\mathbf{F}}$

$$\hat{\mathbf{F}} = \text{DFT}(\mathbf{f}) = \left(\hat{F}_{-\frac{N}{2}} \ \hat{F}_{-\frac{N}{2}+1} \ \dots \ \hat{F}_{\frac{N}{2}-1} \ \hat{F}_{\frac{N}{2}} \right), \quad (10)$$

where the $\hat{F}_i$ are defined by

$$\hat{F}_i = \sum_{j=0}^N f_j \exp\left(-\frac{2\pi i}{N} i j\right) = \sum_{j=0}^N f_j \exp(-i \xi_i x_j), \quad i = -N/2, \dots, N/2. \quad (11)$$

The inverse discrete Fourier transform, denoted by $\mathbf{f} = \text{IDFT}(\hat{\mathbf{F}})$, is given by

$$f_i = \frac{1}{N} \sum_{j=-N/2}^{N/2} \hat{F}_j \exp\left(\frac{2\pi i}{N} i j\right). \quad (12)$$

The coefficients $\hat{F}_i$ provide a numerical approximation of the coefficients F_i of the Fourier series given by equation (6), using a standard trapezoidal rule [26]

$$F_i \simeq \frac{1}{N} \hat{F}_i, \quad i = -N/2, \dots, N/2. \quad (13)$$

Extension to the 3-d case. The extension of Fourier series and transforms to the 3-d case is straightforward. We consider a periodic function f in the prismatic domain $\Omega = [0, L_x] \times [0, L_y] \times [0, L_z]$; its Fourier series simply reads

$$f(x, y, z) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} F_{ijk} \exp(i \xi_i^x x) \exp(i \xi_j^y y) \exp(i \xi_k^z z), \quad \forall (x, y, z) \in \Omega, \quad (14)$$

where the frequencies ξ_i^x , ξ_j^y and ξ_k^z are given by

$$\xi_i^x = 2\pi \frac{i}{L_x}, \quad \xi_j^y = 2\pi \frac{j}{L_y}, \quad \xi_k^z = 2\pi \frac{k}{L_z}, \quad (15)$$

with $i = -\infty, \dots, +\infty$, $j = -\infty, \dots, +\infty$ and $k = -\infty, \dots, +\infty$. The Fourier coefficients F_{ijk} are given by

$$F_{ijk} = \frac{1}{L_x L_y L_z} \int_{\Omega} f(x, y, z) \exp(-i \xi_i^x x) \exp(-i \xi_j^y y) \exp(-i \xi_k^z z) dx dy dz. \quad (16)$$

The coefficients F_{ijk} can be approximated using the 3-d discrete Fourier transform. The 3-d domain Ω is discretized uniformly with $(N_x + 1) \times (N_y + 1) \times (N_z + 1)$ voxels so that the spatial scales are $\Delta x = L_x/(N_x + 1)$, $\Delta y = L_y/(N_y + 1)$ and $\Delta z = L_z/(N_z + 1)$. The coordinates of the grid points are denoted by $x_a = a\Delta x$ (for $a = 0, \dots, N_x$), $y_b = b\Delta y$ (for $b = 0, \dots, N_y$) and $z_c = c\Delta z$ (for $c = 0, \dots, N_z$). The values of the function f at the grid points are denoted by $f_{abc} = f(x_a, y_b, z_c)$ (for $a = 0, \dots, N_x$, $b = 0, \dots, N_y$ and $c = 0, \dots, N_z$) and we denote by $\mathbf{f}$ the array of size $(N_x + 1) \times (N_y + 1) \times (N_z + 1)$ containing the values of function f at the grid points. In 3-d, the discrete Fourier transform of $\mathbf{f}$, denoted by $\text{DFT3}(\mathbf{f})$, is the $(N_x + 1) \times (N_y + 1) \times (N_z + 1)$ array $\hat{\mathbf{F}}$

$$\hat{\mathbf{F}} = \text{DFT3}(\mathbf{f}), \quad (17)$$

whose components are defined by

$$\hat{F}_{ijk} = \sum_{l=0}^{N_x} \sum_{m=0}^{N_y} \sum_{n=0}^{N_z} f_{lmn} \exp\left(-\frac{2\pi i}{N_x} il\right) \exp\left(-\frac{2\pi i}{N_y} jm\right) \exp\left(-\frac{2\pi i}{N_z} kn\right). \quad (18)$$

As in the 1-d case, the coefficients $\hat{F}_{ijk}$ provide a numerical approximation of the coefficients F_{ijk} using the trapezoidal rule:

$$F_{ijk} \simeq \frac{1}{N_x N_y N_z} \hat{F}_{ijk}, \quad i = -N_x/2, \dots, N_x/2, \quad j = -N_y/2, \dots, N_y/2, \quad k = -N_z/2, \dots, N_z/2. \quad (19)$$

3. Discretization approaches

In the seminal paper of [1], the type of formulation (strong or weak) and the associated numerical discretization leading to the numerical algorithm was not explicitly detailed, although it is generally considered that [1]'s scheme was obtained using a strong formulation (see e.g. [27]). In this section, we discuss the numerical resolution of the auxiliary problem from different approaches (strong and weak formulations) and discretization methods (truncation and DFT approximation), following ideas of [12].

3.1. The auxiliary problem

First, the problem of heterogeneous conductivity (2) is written as the auxiliary problem,

$$\begin{cases} \operatorname{div} \mathbf{J}(\mathbf{x}) &= 0 \\ \mathbf{J}(\mathbf{x}) &= c_0 \mathbf{E}(\varphi^*(\mathbf{x})) + \boldsymbol{\tau}(\mathbf{x}) \\ \mathbf{E}(\varphi^*(\mathbf{x})) &= -\nabla \varphi^*(\mathbf{x}), \end{cases} \quad (20)$$

where c_0 is the conductivity of some homogeneous reference medium and $\boldsymbol{\tau}$ is the polarization tensor defined by,

$$\boldsymbol{\tau}(\mathbf{x}) = (\mathbf{c}(\mathbf{x}) - c_0 \mathbf{I}_d) : (\mathbf{E}(\varphi^*(\mathbf{x})) + \bar{\mathbf{E}}) + c_0 \bar{\mathbf{E}}. \quad (21)$$

The principle of resolution is to determine the fluctuation φ^* solution of the problem (20), assuming that $\boldsymbol{\tau}$ is known, and then iterate to find $\boldsymbol{\tau}$ corresponding to the solution of the problem (2). From the knowledge of $\boldsymbol{\tau}$, the numerical determination of φ^* solution of the problem (20) can then be done using several discretization strategies.

3.2. Discretization of the strong formulation by DFT

The strong formulation was implicitly used in the seminal article of [1] by assuming that all fields are described by (full) Fourier series; the conductivity field and the fluctuation field then read

$$\varphi^*(x, y, z) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \Phi_{ijk}^* \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \quad (22)$$

$$c(x, y, z) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} C_{ijk} \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \quad (23)$$

As a consequence, the polarization field is also described by (full) Fourier series and can be written as

$$\tau_x(x, y, z) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} T_{ijk}^x \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \quad (24)$$

$$\tau_y(x, y, z) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} T_{ijk}^y \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \quad (25)$$

$$\tau_z(x, y, z) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} T_{ijk}^z \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \quad (26)$$

With this description of functions, the auxiliary problem reduces to the equation

$$\begin{aligned} c_0 \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right) \Phi_{ijk}^* \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z) \\ = - \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} i \left(\xi_i^x T_{ijk}^x + \xi_j^y T_{ijk}^y + \xi_k^z T_{ijk}^z \right) \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \end{aligned} \quad (27)$$

which leads to the relation between the Fourier series coefficients of the fluctuation and the components of the polarization

$$\Phi_{ijk}^* = -i \frac{\xi_i^x T_{ijk}^x + \xi_j^y T_{ijk}^y + \xi_k^z T_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)}. \quad (28)$$

It is important to note that this relation holds for all frequencies, i.e. for $i = -\infty, \dots, +\infty$, $j = -\infty, \dots, +\infty$ and $k = -\infty, \dots, +\infty$ (except in the particular case $i = j = k = 0$ where one has $\Phi_{000}^* = 0$), and is expressed in terms of the *Fourier series coefficients of τ* . (28) provides the analytical expression of the Fourier coefficients of φ^* , *which need to be evaluated numerically*, leading to a first numerical approximation. Moreover, the numerical computation of the Fourier coefficients Φ_{ijk}^* from (28) is only performed for a finite number of indexes ijk , leading to a second numerical approximation in the estimation of φ^* due to the *truncation* of its Fourier series into a partial one.

In practical applications, the Fourier series coefficients are approximated by the coefficients of the discrete Fourier transform as shown by equation (19), which need to be calculated using some discretization of the domain. The prismatic Ω is therefore discretized uniformly with $(N_x + 1) \times (N_y + 1) \times (N_z + 1)$ voxels so that the spatial scales are $\Delta x = L_x / (N_x + 1)$, $\Delta y = L_y / (N_y + 1)$ and $\Delta z = L_z / (N_z + 1)$. Following equation (19), the discrete Fourier transforms of the fields φ^* , τ_x , τ_y and τ_z provide a numerical approximation of the Fourier series coefficients:

$$\Phi_{ijk}^* \simeq \frac{1}{N_x N_y N_z} \widehat{\Phi}_{ijk}^*, \quad T_{ijk}^x \simeq \frac{1}{N_x N_y N_z} \widehat{T}_{ijk}^x, \quad T_{ijk}^y \simeq \frac{1}{N_x N_y N_z} \widehat{T}_{ijk}^y, \quad T_{ijk}^z \simeq \frac{1}{N_x N_y N_z} \widehat{T}_{ijk}^z, \quad (29)$$

for $i = -N_x/2, \dots, N_x/2$, $j = -N_y/2, \dots, N_y/2$ and $k = -N_z/2, \dots, N_z/2$. This finally leads to

$$\widehat{\Phi}_{ijk}^* = -i \frac{\xi_i^x \widehat{T}_{ijk}^x + \xi_j^y \widehat{T}_{ijk}^y + \xi_k^z \widehat{T}_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)}, \quad (30)$$

again for $i = -N_x/2, \dots, N_x/2$, $j = -N_y/2, \dots, N_y/2$ and $k = -N_z/2, \dots, N_z/2$, except in the particular case $i = j = k = 0$ where one has $\widehat{\Phi}_{000}^* = 0$.

It is worth noting that, if the FFT algorithm suggests the computations of the Fourier coefficients for these indexes, it is possible to take fewer Fourier modes, i.e. compute (30) for $i = -M_x/2, \dots, M_x/2$, $j = -M_y/2, \dots, M_y/2$ and $k = -M_z/2, \dots, M_z/2$ with $M_x \leq N_x$, $M_y \leq N_y$ and $M_z \leq N_z$. While this choice seems to discard easy-to-compute coefficients and to reduce the order of the Fourier truncation of φ^* , therefore downgrades its numerical estimation, one must

keep in mind that the Fourier coefficients of τ in (29) are numerically estimated by DFT, therefore they have a numerical error which increases with i , j and k (see, for instance, [28]). Considering $M_x \leq N_x$, $M_y \leq N_y$ and $M_z \leq N_z$, therefore suppressing *high frequency* coefficients, discards the coefficients with highest numerical error coming from the DFT of τ . The influence of this approximation will be evaluated in the next section.

Equation (30) provides the coefficients of the discrete Fourier transform of φ^* as a function of the coefficients of the discrete Fourier transform of τ_x , τ_y and τ_z and thus it implicitly defines the so-called Green operator on the electric potential field. Owing to the relation between the electric potential field and the electric field, this leads to the relations

$$\begin{cases} \hat{E}_{ijk}^x = -\xi_i^x \frac{\xi_i^x \hat{T}_{ijk}^x + \xi_j^y \hat{T}_{ijk}^y + \xi_k^z \hat{T}_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)} \\ \hat{E}_{ijk}^y = -\xi_j^y \frac{\xi_i^x \hat{T}_{ijk}^x + \xi_j^y \hat{T}_{ijk}^y + \xi_k^z \hat{T}_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)} \\ \hat{E}_{ijk}^z = -\xi_k^z \frac{\xi_i^x \hat{T}_{ijk}^x + \xi_j^y \hat{T}_{ijk}^y + \xi_k^z \hat{T}_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)}, \end{cases} \quad (31)$$

for $i = -N_x/2, \dots, N_x/2$, $j = -N_y/2, \dots, N_y/2$ and $k = -N_z/2, \dots, N_z/2$ (except in the case $i = j = k = 0$). In the case $i = j = k = 0$, one has

$$\hat{E}_{000}^x = \bar{E}_x, \quad \hat{E}_{000}^y = \bar{E}_y, \quad \hat{E}_{000}^z = \bar{E}_z, \quad (32)$$

where $\bar{E}_x$, $\bar{E}_y$ and $\bar{E}_z$ are the components of the average of the electric field $\bar{\mathbf{E}}$. In equation (31), $\hat{E}_{ijk}^x$, $\hat{E}_{ijk}^y$ and $\hat{E}_{ijk}^z$ are the coefficients of the discrete Fourier transforms of the components of the local electric field $\mathbf{E}$:

$$\hat{\mathbf{E}}^x = \text{DFT3}(\mathbf{E}_x), \quad \hat{\mathbf{E}}^y = \text{DFT3}(\mathbf{E}_y), \quad \hat{\mathbf{E}}^z = \text{DFT3}(\mathbf{E}_z). \quad (33)$$

Equation (31) can be alternatively written

$$\hat{\mathbf{E}}(\boldsymbol{\xi}) = -\hat{\Gamma}^0(\boldsymbol{\xi}) : \hat{\boldsymbol{\tau}}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \neq \mathbf{0}, \quad \hat{\mathbf{E}}(\mathbf{0}) = \bar{\mathbf{E}}, \quad (34)$$

where $\boldsymbol{\xi}$ denotes the frequency in Fourier space. This leads to the definition of the so-called Green operator $\hat{\Gamma}^0$ which has an explicit form in Fourier space:

$$\hat{\Gamma}^0(\boldsymbol{\xi}) = \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{c_0 |\boldsymbol{\xi}|^2}. \quad (35)$$

3.3. Discretization of the strong formulation by truncation

In the previous section, the fields have been described initially using Fourier series, i.e. with an infinite number of Fourier coefficients. In practice, this description is not compatible with a standard discretization of functions which requires a finite number of terms. The problem can be alternatively written by considering partial (truncated) Fourier series (of order $(M_x + 1) \times (M_y + 1) \times (M_z + 1)$) for the conductivity field and the fluctuation field:

$$\varphi^*(x, y, z) = \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} \Phi_{ijk}^* \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z) \quad (36)$$

$$c(x, y, z) = \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} C_{ijk} \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \quad (37)$$

With this description, the local electric field $\mathbf{E}$ is also described by a partial (truncated) Fourier series (of order $(M_x + 1) \times (M_y + 1) \times (M_z + 1)$).

Since the polarization tensor is given by equation (21), as the product of the conductivity field and the local electric field, it is necessarily described by partial (truncated) Fourier series of order $(2M_x + 1) \times (2M_y + 1) \times (2M_z + 1)$, i.e.,

$$\tau_x(x, y, z) = \sum_{i=-M_x}^{M_x} \sum_{j=-M_y}^{M_y} \sum_{k=-M_z}^{M_z} T_{ijk}^x \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \quad (38)$$

$$\tau_y(x, y, z) = \sum_{i=-M_x}^{M_x} \sum_{j=-M_y}^{M_y} \sum_{k=-M_z}^{M_z} T_{ijk}^y \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \quad (39)$$

$$\tau_z(x, y, z) = \sum_{i=-M_x}^{M_x} \sum_{j=-M_y}^{M_y} \sum_{k=-M_z}^{M_z} T_{ijk}^z \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \quad (40)$$

Consequently, the auxiliary problem reduces to the equation

$$\begin{aligned} c_0 \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right) \Phi_{ijk}^* \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z) \\ = - \sum_{i=-M_x}^{M_x} \sum_{j=-M_y}^{M_y} \sum_{k=-M_z}^{M_z} i \left(\xi_i^x T_{ijk}^x + \xi_j^y T_{ijk}^y + \xi_k^z T_{ijk}^z \right) \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \end{aligned} \quad (41)$$

In contrast with the previous case where functions were described by infinite Fourier series, equation (41) does not possess a direct solution as it involves Fourier series of orders $(M_x + 1) \times (M_y + 1) \times (M_z + 1)$ and $(2M_x + 1) \times (2M_y + 1) \times (2M_z + 1)$; the polarization tensor contains contributions at frequencies exceeding the maximum frequencies of the unknown field φ^* , which can be considered as *aliasing*.

To overcome this issue, we will simply remove the high-frequencies of the polarization tensor, i.e. we will assume that it is expressed as

$$\begin{aligned} \tau_x(x, y, z) &= \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} T_{ijk}^x \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z) \\ \tau_y(x, y, z) &= \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} T_{ijk}^y \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z) \\ \tau_z(x, y, z) &= \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} T_{ijk}^z \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \end{aligned} \quad (42)$$

With this assumption, equation (41) now reads

$$\begin{aligned} c_0 \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right) \Phi_{ijk}^* \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z) \\ = - \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} i \left(\xi_i^x T_{ijk}^x + \xi_j^y T_{ijk}^y + \xi_k^z T_{ijk}^z \right) \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z), \end{aligned} \quad (43)$$

which leads to the relation between the Fourier series coefficients of the fluctuation and the components of the polarization

$$\Phi_{ijk}^* = -i \frac{\xi_i^x T_{ijk}^x + \xi_j^y T_{ijk}^y + \xi_k^z T_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)}, \quad (44)$$

for $i = -M_x/2, \dots, M_x/2$, $j = -M_y/2, \dots, M_y/2$ and $k = -M_z/2, \dots, M_z/2$ (except in the particular case $i = j = k = 0$ where one has $\Phi_{000}^* = 0$).

Then as previously, these coefficients are approximated by the coefficients of the discrete Fourier transform which need to be calculated using a discretization of the domain. The prismatic Ω is discretized uniformly with $(N_x + 1) \times (N_y + 1) \times (N_z + 1)$ voxels so that the spatial scales are $\Delta x = L_x/(N_x + 1)$, $\Delta y = L_y/(N_y + 1)$ and $\Delta z = L_z/(N_z + 1)$. The discrete Fourier transforms of the fields φ^* , τ_x , τ_y and τ_z provide a numerical approximation of the Fourier series coefficients as shown previously. The only precaution that is required is to have a *minimum number of "integration" points for the trapezoidal rule* (i.e. N_x , N_y and N_z) with respect to the *number of modes* chosen in the description of the partial Fourier series (i.e. M_x , M_y and M_z), that is, again,

$$N_x \geq M_x, \quad N_y \geq M_y, \quad N_z \geq M_z. \quad (45)$$

This finally leads to

$$\widehat{\Phi^*}_{ijk} = -i \frac{\xi_i^x \widehat{T}_{ijk}^x + \xi_j^y \widehat{T}_{ijk}^y + \xi_k^z \widehat{T}_{ijk}^z}{c_0 \left((\xi_i^x)^2 + (\xi_j^y)^2 + (\xi_k^z)^2 \right)}, \quad (46)$$

for $i = -M_x/2, \dots, M_x/2$, $j = -M_y/2, \dots, M_y/2$ and $k = -M_z/2, \dots, M_z/2$, except in the particular case $i = j = k = 0$ where one has $\widehat{\Phi^*}_{000} = 0$.

It is worth noting that equation (46) is equivalent to equation (30), so the formulation in terms of the Green operator (35) will naturally follow exactly the same way. The standard strong formulation and the alternative strong formulation with truncation will therefore lead to the same exact algorithm. The two approaches simply differ from their derivation:

- In the classical approach, one starts with full Fourier series, the truncation of the Fourier series is a necessary consequence of the numerical estimation of the Fourier coefficients of τ .
- In the alternative approach, the truncation of the Fourier series is a starting discretization approximation, leading to a second necessary truncation of higher-order modes associated with higher-order frequencies (related to aliasing effects) that need to be disregarded to solve the auxiliary problem;

In both cases, the evaluation of the Fourier coefficients is done using discrete Fourier transforms with at least the same number of integration points than the number of modes considered in the description of the fields. It should be noted that one can consider more integration points than the number of modes (i.e. in the case when $N_x > M_x$, $N_y > M_y$ and $N_z > M_z$); this is actually equivalent to disregard Fourier coefficients in the case of high frequencies (i.e. for $|i| > M_x/2$, $|j| > M_y/2$ and $|k| > M_z/2$). It is important to note that the natural choice, i.e., $N_x = M_x$, $N_y = M_y$, $N_z = M_z$, is *not necessarily* the optimal choice, in terms of quality of the solution. The influence of the ratio M/N will be investigated in Section 5 on a (specific) checkerboard microstructure.

3.4. Weak formulation with Galerkin approximation

We now consider the weak formulation of the auxiliary problem together with a Galerkin method. First, the unknown fluctuation field is described by a (partial) Fourier series,

$$\varphi^*(x, y, z) = \sum_{i=-M_x/2}^{M_x/2} \sum_{j=-M_y/2}^{M_y/2} \sum_{k=-M_z/2}^{M_z/2} \Phi_{ijk}^* \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z). \quad (47)$$

We are thus looking for the Fourier modes Φ_{ijk}^* for $i = -M_x/2, \dots, M_x/2$, $j = -M_y/2, \dots, M_y/2$ and $k = -M_z/2, \dots, M_z/2$, except in the particular case $i = j = k = 0$ where one has $\Phi_{000}^* = 0$ by construction.

The weak formulation of the local equilibrium (20) reads

$$\int_{\Omega} \Delta \tilde{\varphi}^*(\mathbf{x}) \nu(\mathbf{x}) d\Omega = \frac{1}{c_0} \int_{\Omega} \operatorname{div} \boldsymbol{\tau}(\mathbf{x}) \nu(\mathbf{x}) d\Omega, \quad (48)$$

where Δ is the Laplacian operator and $v(\mathbf{x})$ is a testing differentiable function. Following equation (47), these trial functions are the interpolation basis of φ^* , i.e.,

$$v_{lmn}(\mathbf{x}) = \exp(\iota \xi_l^x x) \exp(\iota \xi_m^y y) \exp(\iota \xi_n^z z) \quad (49)$$

for $l = -M_x/2, \dots, M_x/2$, $m = -M_y/2, \dots, M_y/2$ and $n = -M_z/2, \dots, M_z/2$. The integrals involved in the weak formulation (48), that need to be calculated, are denoted by

$$I_{lmn} = \int_{\Omega} \Delta \tilde{\varphi}^*(\mathbf{x}) v_{lmn}(\mathbf{x}) d\Omega, \quad J_{lmn} = \int_{\Omega} (\operatorname{div} \boldsymbol{\tau}(\mathbf{x})) v_{lmn}(\mathbf{x}) d\Omega. \quad (50)$$

Calculation of integral I_{lmn} . The calculation of I_{lmn} simply relies on the property of orthogonality of complex exponentials and is calculated analytically; one has

$$I_{lmn} = -L_x L_y L_z ((\xi_{-l}^x)^2 + (\xi_{-m}^y)^2 + (\xi_{-n}^z)^2) \Phi_{-l-m-n}^* \quad (51)$$

since

$$\int_0^{L_x} \exp(\iota \xi_i^x x) \exp(\iota \xi_l^x x) dx = \begin{cases} 0 & \text{if } i \neq l \\ L_x & \text{if } i = -l. \end{cases} \quad (52)$$

The calculation of J_{lmn} can then be done using two different ways.

Calculation of integral J_{lmn} . Assuming that the polarization field is a periodic function described by Fourier series as in equation (42), and using again the property of orthogonality of complex exponentials, one has

$$J_{lmn} = \iota L_x L_y L_z (\xi_{-l}^x T_{-l-m-n}^x + \xi_{-m}^y T_{-l-m-n}^y + \xi_{-n}^z T_{-l-m-n}^z). \quad (53)$$

This leads to

$$\Phi_{-l-m-n}^* = \iota \frac{\xi_{-l}^x T_{-l-m-n}^x + \xi_{-m}^y T_{-l-m-n}^y + \xi_{-n}^z T_{-l-m-n}^z}{c_0 ((\xi_{-l}^x)^2 + (\xi_{-m}^y)^2 + (\xi_{-n}^z)^2)} \quad (54)$$

or equivalently, using the relations $\xi_{-l}^x = -\xi_l^x$, $\xi_{-m}^y = -\xi_m^y$ and $\xi_{-n}^z = -\xi_n^z$,

$$\Phi_{lmn}^* = -\iota \frac{\xi_l^x T_{lmn}^x + \xi_m^y T_{lmn}^y + \xi_n^z T_{lmn}^z}{c_0 ((\xi_l^x)^2 + (\xi_m^y)^2 + (\xi_n^z)^2)}. \quad (55)$$

This leads to the exact same formulation than the strong forms derived previously (equations (28) and (44)). Therefore, as previously, one needs to evaluate these coefficients using discrete Fourier transforms (see equation (29)).

Alternative calculation of integral J_{lmn} . Integral J_{lmn} can be alternatively calculated without assuming that the polarization field is described by Fourier series. In that case, the integral is directly approximated using a trapezoidal rule and will thus lead to the same discrete Fourier transform and thus the exact same formulation.

3.5. Iterative scheme

The basic scheme of [1] then consists in a fixed point in order to determine $\boldsymbol{\tau}(\mathbf{x})$ solution of the problem. Irrespective of the method considered (strong form, alternative strong form or weak form), the resolution of the auxiliary problem leads to the same relation between the Fourier

coefficients of the unknown fluctuation and the Fourier coefficients of the polarization tensor. The following fixed-point algorithm then follows:

$$\left\{ \begin{array}{ll} \text{Initialization} & \mathbf{E}^0(\mathbf{x}) = \bar{\mathbf{E}} \\ & \mathbf{J}^0(\mathbf{x}) = c(\mathbf{x})\mathbf{E}^0(\mathbf{x}) \\ \text{Iterate } i+1 & \mathbf{E}^i \text{ and } \mathbf{J}^i \text{ being known} \\ & (a) \quad \boldsymbol{\tau}^i(\mathbf{x}) = \mathbf{J}^i(\mathbf{x}) - c^0\mathbf{E}^i(\mathbf{x}) \\ & (b) \quad \hat{\boldsymbol{\tau}}^i = \text{DFT3}(\boldsymbol{\tau}^i) \\ & (c) \quad \text{Convergence test} \\ & (d) \quad \hat{\mathbf{E}}^{i+1}(\boldsymbol{\xi}) = -\hat{\Gamma}^0(\boldsymbol{\xi}) : \hat{\boldsymbol{\tau}}^i(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \neq \mathbf{0} \quad \text{and} \quad \hat{\mathbf{E}}^{i+1}(\mathbf{0}) = \bar{\mathbf{E}} \\ & (e) \quad \mathbf{E}^{i+1} = \text{IDFT3}(\hat{\mathbf{E}}^{i+1}) \\ & \quad \mathbf{E}^{i+1}(\mathbf{x}) \text{ expressed using Fourier series} \\ & (f) \quad \mathbf{J}^{i+1}(\mathbf{x}) = c(\mathbf{x})\mathbf{E}^{i+1}(\mathbf{x}) \end{array} \right. \quad (56)$$

Convergence is reached when $\mathbf{J}^{i+1}$ is in equilibrium, i.e. its divergence is null. In practice, the iterative procedure is stopped when the error serving to check convergence [1],

$$e^i = \frac{\left(\left\langle \|\text{div}(\mathbf{J}^i)\|^2 \right\rangle \right)^{1/2}}{\|\langle \mathbf{J}^i \rangle\|}, \quad (57)$$

is smaller than a prescribed (small) value.

4. Representation of solution fields and non-uniform discretization

4.1. Representation of solution fields

The FFT-based method is based on the approximate calculation of Fourier series coefficients (equation (16)) using the discrete Fourier transform on a *regular* grid. The role of the (uniform) spatial discretization of the domain is therefore to calculate numerically the Fourier coefficients used in the description of the fields (equation (18)), based on a trapezoidal (and thus uniform) quadrature rule (with $N + 1$ equi-spaced points) [26], i.e.,

$$\hat{\mathbf{F}} = \text{DFT}(\mathbf{f}), \quad (58)$$

where $\mathbf{f}$ contains the nodal values of function f on a uniform grid. By taking advantage of the relation (13) between the discrete Fourier coefficients (i.e. $\hat{\mathbf{F}}$) and the Fourier series coefficients (i.e. $\mathbf{F}$), the reconstruction of function f then follows *in the whole domain*

$$f(x) = \sum_{i=-M/2}^{M/2} F_i \exp(i\xi_i x), \quad \forall x \in [0, L]. \quad (59)$$

Usually, the number of Fourier modes ($M + 1$) is taken equal to the number of spatial nodes ($N + 1$), i.e. $N = M$, and then the inverse Fourier transform can then be used to evaluate, *on the same spatial uniform grid used to approximate the integrals*, function f as,

$$\mathbf{f} = \text{IDFT}(\hat{\mathbf{F}}). \quad (60)$$

In the literature, the solution fields are then generally represented from these nodal values, either using uniform values in a voxel or a linear interpolation between nodal values (see e.g. [1]). However, as the solution fields are described, by construction, by Fourier series, they are defined in the whole domain as shown by equation (59). This implies that representing the field only at the nodes, or uniformly in the voxels, cannot capture variations between spatial nodes, and may therefore hide some spurious numerical effects such as Gibbs effects. This can be overcome by representing the solution fields on a finer grid, using directly equation (59). In the context of

homogenization, such a refined representation does not modify macroscopic quantities, which depend only on spatial averages. Representing the solution field on a finer grid is therefore not motivated by the computation of effective properties, but rather by situations where an accurate evaluation of the solution fields is required (e.g., evaluation of local criteria, stress-assisted diffusion, or phase-field models).

Let us denote by x_i^{fine} the grid points of a uniform finer grid (with $N^{\text{fine}} + 1$ points) and by $f_i^{\text{fine}} = f(x_i^{\text{fine}})$ the value of function f on this refined grid. We denote by $\mathbf{f}^{\text{fine}}$ the array of size $N^{\text{fine}} + 1$ containing the values f_i^{fine} . From equation (59) one simply has

$$f_i^{\text{fine}} = \sum_{j=-M/2}^{M/2} F_j \exp(i \xi_j x_i^{\text{fine}}) \quad (61)$$

which can be written as

$$\mathbf{f}^{\text{fine}} = \mathbf{A} \cdot \mathbf{F}, \quad (62)$$

where $\mathbf{A}$ is the $(N^{\text{fine}} + 1) \times (M + 1)$ array defined by

$$A_{ij} = \exp(i \xi_j x_i^{\text{fine}}). \quad (63)$$

The determination of $\mathbf{f}^{\text{fine}}$ then requires a matrix-vector operation involving matrix $\mathbf{A}$ for which the numerical complexity scales in $\mathcal{O}(N^{\text{fine}} \times M)$. This evaluation can be considerably reduced by using the inverse discrete Fourier transform with a modified vector for the Fourier coefficients, denoted by $\hat{\mathbf{F}}^{\text{fine}}$, taken null for the extra high frequency modes:

$$\hat{F}_i^{\text{fine}} = \begin{cases} \hat{F}_i & \text{if } i \leq M/2 \\ 0 & \text{if } i > M/2. \end{cases} \quad (64)$$

In that case the numerical complexity would scale in $\mathcal{O}(N^{\text{fine}} \log(N^{\text{fine}}))$.

It is important to note that representing the solution on a finer grid than the resolution grid is *not* a reinterpolation of the solution; it is simply the exact evaluation of the same spectral approximation at new points. No additional interpolation procedure is performed; the underlying trigonometric polynomial remains unchanged.

4.2. Non-uniform discretization

The representation of the local fields on a finer grid entails a non-negligible computational cost, scaling as $\mathcal{O}(N^{\text{fine}} \log N^{\text{fine}})$. In situations where only localized regions of the RVE are of interest, the computational cost can be reduced by resorting to non-uniform FFT evaluations on *selected points*.

Indeed, as function f is described by the Fourier series (59), it can be calculated in any grid point, even in an *non-uniform grid*. Let us denote by x_i^{nu} the grid points of a non-uniform grid (with $N^{\text{nu}} + 1$ points $x_i^{\text{nu}} \in [0, L]$) and by $f_i^{\text{nu}} = f(x_i^{\text{nu}})$ the value of function f on this non-uniform grid. We denote by $\mathbf{f}^{\text{nu}}$ the array of size $N^{\text{nu}} + 1$ containing the values f_i^{nu} . From equation (59) one simply has

$$f_i^{\text{nu}} = \sum_{j=-M/2}^{M/2} F_j \exp(i \xi_j x_i^{\text{nu}}) \quad (65)$$

which can be written as

$$\mathbf{f}^{\text{nu}} = \mathbf{A} \cdot \mathbf{F}, \quad (66)$$

where $\mathbf{A}$ is the $(N^{\text{nu}} + 1) \times (M + 1)$ array defined by

$$A_{ij} = \exp(i \xi_j x_i^{\text{nu}}). \quad (67)$$

As previously the numerical complexity of this operation is in $\mathcal{O}(N^{\text{nu}} \times M)$. However, in the case where $N^{\text{nu}} = M$, but still with a non-uniform grid, the evaluation of function f can be done with the same numerical complexity of fast Fourier transforms at nonequispaced points, which is referred to as NFFT, for nonequispaced fast Fourier transform, or nonuniform fast Fourier transform (NUFFT) [29,30].

Knowing the Fourier series coefficients of function f , i.e. F (or equivalently the discrete Fourier coefficients $\hat{F}$), the NUFFT defined by equation (66) permits to evaluate *fastly* function f at any points of an arbitrary grid. It thus can be used simply as a post-treatment to represent the solution field with a fine resolution in areas where strong gradients are expected instead of using a very fine mesh everywhere in the domain (which would entail a non-negligible computational cost).

In addition, as equation (66) provides the matrix expression of the NFFT, it could be inverted so as to provide the value of the Fourier coefficients as a function of the grid point values of function f and can be calculated efficiently as standard inverse FFT [31]. It thus defines an alternative way of calculating the Fourier coefficients on a non-uniform grid which can be interpreted as a modified quadrature rule. However it is expected not to improve the results obtained using the trapezoidal rule which is exponentially convergent for a periodic function [26].

5. Numerical experiments

The aim of this section is to study, for a specific problem admitting an analytical solution, the influence of the number of Fourier modes and grid discretization on the quality of the solution fields. We investigate (i) the influence of the grid discretization (N) for a given number of modes (M) and (ii) the influence of the number of modes (M) for a given grid discretization (N). Indeed, as explained in Section 3, the number of grid points N can be chosen independently of the number of mode M with the sole condition $N \geq M$; the relevance of the standard choice $M = N$ will be notably studied. A particular attention will be put on the representation of the fields as explained in Section 4.

5.1. Description of the problem

We consider the classical problem of conductivity for checkerboard-type microstructures which is of interest for the method since this problem has an analytical exact solution [32–34] (see also [35]). This problem was notably investigated in previous studies since it is highly challenging for the accurate description, without spurious oscillations of the local fields [18,24,36]. This is an extreme case due to the presence of sharp interfaces and corners; it is thus a good benchmark to assess the influence of the ratio M/N .

We consider a 2D periodic square domain $[0, 1] \times [0, 1]$ discretized with $(N + 1) \times (N + 1)$ grid points, using the notations of Section 3. We consider that the isotropic conductivity field $c(x_1, x_2)$ is heterogeneous and forms a checkerboard-type microstructure with values $c_1 = 1$ and $c_2 = 100$ as shown in Figure 1a and defined as

$$c(x_1, x_2) = \begin{cases} c_2 & \text{if } x_1 < 0.5 \text{ and } x_2 < 0.5 \\ c_1 & \text{elsewhere.} \end{cases} \quad (68)$$

Thus, the grid point at $(0, 0)$ is at the left bottom corner of the included square. One might remark that other choice of alignment between the grid and the microstructure might lead to slightly different results. This influence, sometimes referred as the “egg-box” effect (see, for instance, [37]), is not investigated here.

The values $\bar{E}_1 = 1$ and $\bar{E}_2 = 0$ are considered for the prescribed macroscopic electric field $\bar{\mathbf{E}}$. Convergence of the iterative procedure is reached when the error defined by equation (57) is smaller than the prescribed value 10^{-8} . The normalized electric field $E_1 / \|\bar{\mathbf{E}}\|$ obtained from the theoretical solution of [34] is represented in Figure 1b.

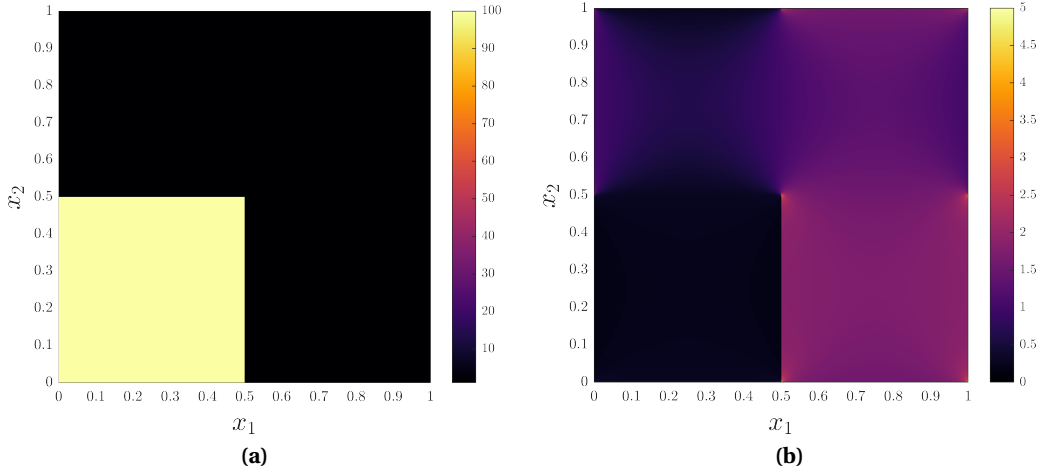


Figure 1. Checkerboard problem considered. (a) Distribution of the conductivity field $c(x_1, x_2)$ field and (b) Distribution of the theoretical normalized electric field $E_1 / \|\bar{\mathbf{E}}\|$.

For the numerical calculations, several values for the number of grid points will be considered, $N = [1024, 512, 256, 128]$, and several values for the number of Fourier modes will also be considered, $M = [1024, 512, 256, 128]$. The field will be represented on both their “original” grid (of size $(N + 1) \times (N + 1)$) as well as on a refined grid as explained in Section 4, of size $(N^{\text{fine}} + 1) \times (N^{\text{fine}} + 1)$ with $N^{\text{fine}} = 8192$.

Qualitative results are first presented (Section 5.2 and 5.3) to illustrate separately the effect of mesh refinement and Fourier modes refinement on the local fields. Quantitative comparisons are then provided in Section 5.4.

5.2. Influence of the grid discretization (N) for a given number of modes (M)

We consider the case $M = 128$ Fourier modes and we vary the number of grid nodes ($N = [1024, 512, 256, 128]$), which therefore consists in modifying the number of integration points for calculating numerically the Fourier coefficients. First, the distribution of the component E_1 of the electric field is represented in Figure 2 for several values of the grid nodes. In each case, the distribution is represented on both the calculation grid (i.e. the coarse grid of size $(N + 1) \times (N + 1)$) and the fine grid (of size $(N^{\text{fine}} + 1) \times (N^{\text{fine}} + 1)$).

Overall, the results in all cases are close to the analytical solution, but spurious oscillations are observed. Interestingly, these oscillations are not visible in the case ($M = 128, N = 128$) when the solution is represented in the coarse grid, i.e. on the grid used for the calculation of the FFT (Figure 2a), but they are visible when the solution is represented in the fine grid (Figure 2b); this means that, when the solution is only plotted on the nodes of calculation, oscillations between these nodes may not be seen, but between grid points these oscillations do exist. When increasing the number of integration points, i.e. N , the solution seems to converge, and the solutions represented on the discretization grid and the fine grid coincide.

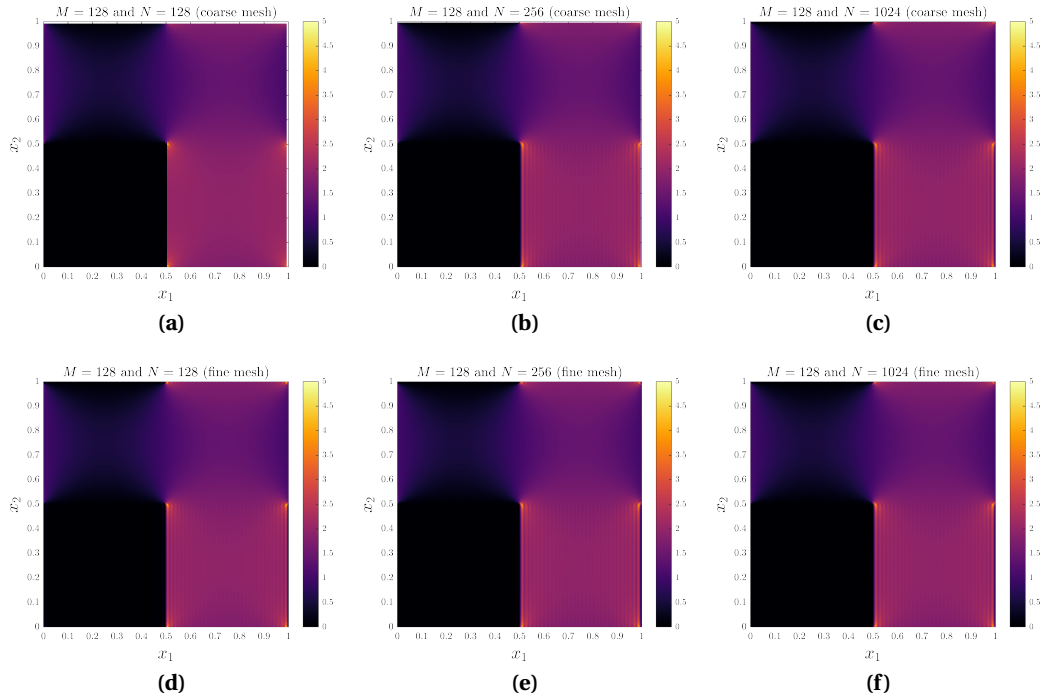


Figure 2. Influence of the grid discretization (N) for a given number of modes (M) on the distribution of the normalized electric field $E_1 / \|\bar{\mathbf{E}}\|$. The solution is represented on both the discretization grid (coarse mesh) and the fine grid (fine mesh).

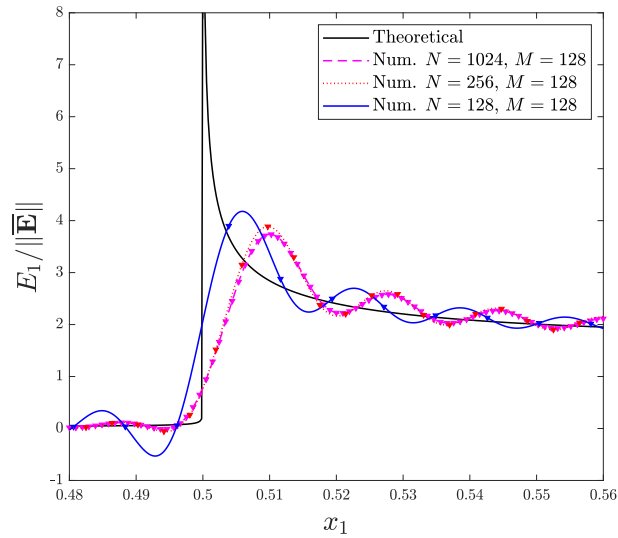


Figure 3. Line plot of the normalized electric field $E_1 / \|\bar{\mathbf{E}}\|$ interface between the two phases ($x_2 = 0$) for a fixed number of modes ($M = 128$). The continuous lines corresponds to the solution on the fine grid and the triangles corresponds to the solution on the grid used for the FFT calculations.

Additionally, an enlarged view, in the vicinity of the interface between the two phases, is plotted in Figure 3, on the line $x_2 = 0$. In the case ($M = 128$, $N = 128$), the oscillations are not visible if one only represents the gridpoint values of the field (represented by blue triangles) but they become visible when the solution is represented in the fine grid. It is interesting to note that the solution field stabilizes when N increases. This is consistent with the role played by N , which is to calculate approximately integrals: for a fixed number of modes M , the increase of N allows a better approximation of the integrals, but this evaluation is expected to reach a stationary value when increasing the number of integration points.

5.3. Influence of the number of modes (M) for a given grid discretization (N)

We now consider the case $N = 1024$ and we vary the number of Fourier modes ($M = [1024, 256, 128]$), which therefore consists in removing some high frequencies in the solution. First, the distribution of the component E_1 of the electric field is represented in Figure 4 for several values of the grid nodes. In each case, the distribution is represented on both the calculation grid (i.e. the coarse grid of size $(N + 1) \times (N + 1)$) and the fine grid (of size $(N^{\text{fine}} + 1) \times (N^{\text{fine}} + 1)$).

First, the distribution seems quite similar irrespective of the resolution considered for the representation. As previously, spurious oscillations are observed on the solution fields, which tend to decrease when the number of Fourier modes increases.

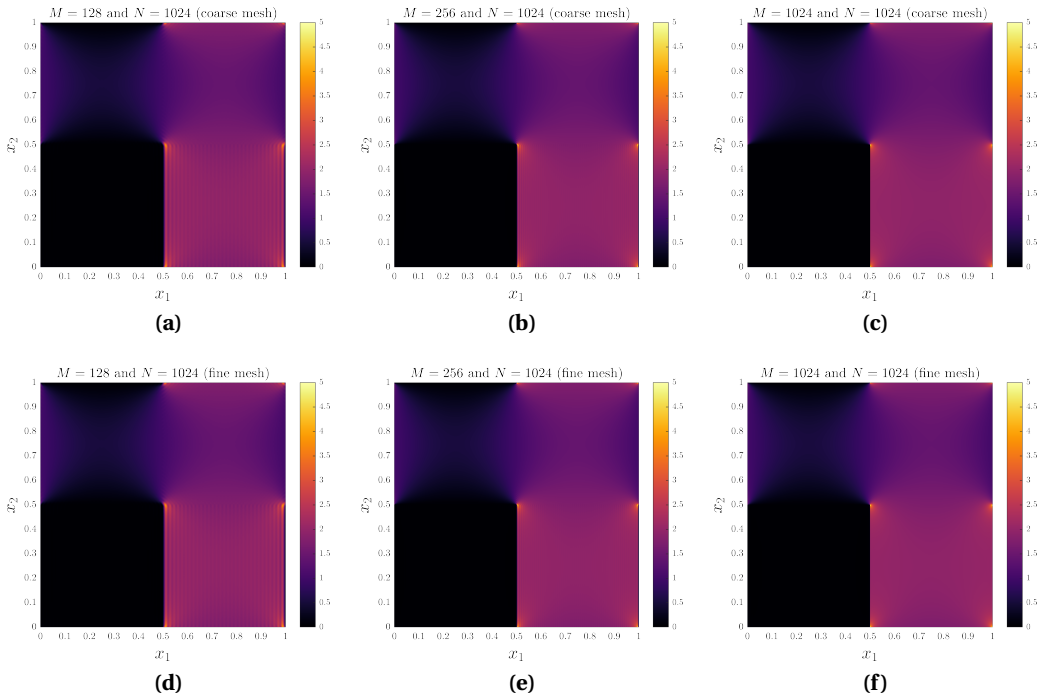


Figure 4. Influence of number of modes (M) for a given grid discretization (N) on the distribution of the normalized electric field $E_1 / \|\vec{E}\|$. The solution is represented on both the discretization grid (coarse mesh) and the fine grid (fine mesh).

An enlarged view of the solution field, at the vicinity of the interface between the two phases, is plotted in Figure 5, on the line $x_2 = 0$. This representation allows to see the oscillatory

nature of all solutions, with a period that depends on the number of Fourier modes (M). This is consistent with the interpretation of the Fourier modes which are associated with increasing frequencies. Increasing the number of modes allows to restrict the Gibbs phenomenon close to the interface but it may increase the amplitude of the oscillations locally. Therefore, even though the numerical solution seems to get closer to the analytical one when increasing M , oscillations with higher frequencies appear, so one cannot conclude on the improvement of quality of the solution field with the increase of the modes.

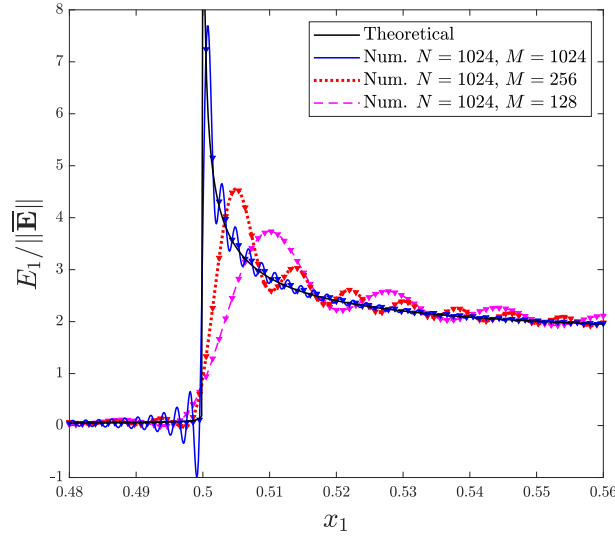


Figure 5. Line plot of the normalized electric field $E_1/||\vec{E}||$ interface between the two phases ($x_2 = 0$) for a fixed grid discretization ($N = 1024$). The continuous lines corresponds to the solution on the fine grid and the triangles corresponds to the solution on the grid used for the FFT calculations.

5.4. Quantitative comparison with the analytical solution

A quantitative comparison between numerical solutions and the analytical one is made by studying the relative error in energy norm defined as [38]

$$\text{err} = \frac{\int_{\Omega} (\mathbf{E}^{\text{th}} - \mathbf{E}^{\text{num}}) : \mathbf{c} : (\mathbf{E}^{\text{th}} - \mathbf{E}^{\text{num}}) d\Omega}{\int_{\Omega} \mathbf{E}^{\text{th}} : \mathbf{c} : \mathbf{E}^{\text{th}} d\Omega}, \quad (69)$$

where $\mathbf{E}^{\text{th}}$ is the theoretical electric field and $\mathbf{E}^{\text{num}}$ is the numerical electric field obtained by FFT. Following Galerkin's framework, this error should be minimal in the chosen function subspace, i.e. for the number of modes M .

The error is represented in Figure 6a as a function of the number of grid nodes N (for a given number of modes M). For a given number of modes M , the error decreases with the number of nodes (which plays here the role of integration points) and tends to an asymptotic value, in agreement with the results of Figure 3. This confirms that the use of FFT (approximate calculation of integrals using a trapezoidal rule, which can be seen as a reduced integration) induces an error on the solution fields. The stabilization of the error indicates that the reduced-integration-induced error has been suppressed. Residual error comes from the truncated Fourier series approximation, and decreases with the number of modes M . This confirms that the approximate

calculation of integrals using a trapezoidal rule, for a given function (imposed by the number of modes), reaches a stationary value when increasing the number of integration points N .

Then, the error is represented in Figure 6b as a function of the number of modes M (for a given number of nodes N). Overall, the error tends to decrease when the number of modes M increases (for a fixed N) until a limit value M_{lim} above which the error increases. This value M_{lim} depends on the number of nodes N and its order of magnitude is of about $M_{\text{lim}} \sim 3N/4$. This peculiar result indicates that, for this problem, the optimal number of modes M is not necessarily equal to the number of integration points N : a truncation of the Fourier series (i.e. $M < N$), which corresponds to disregard Fourier coefficients for high frequencies, can lead to a better solution field. This is equivalent to filtering high frequency components known to degrade solution fields near regions of discontinuity (see e.g. [39]).

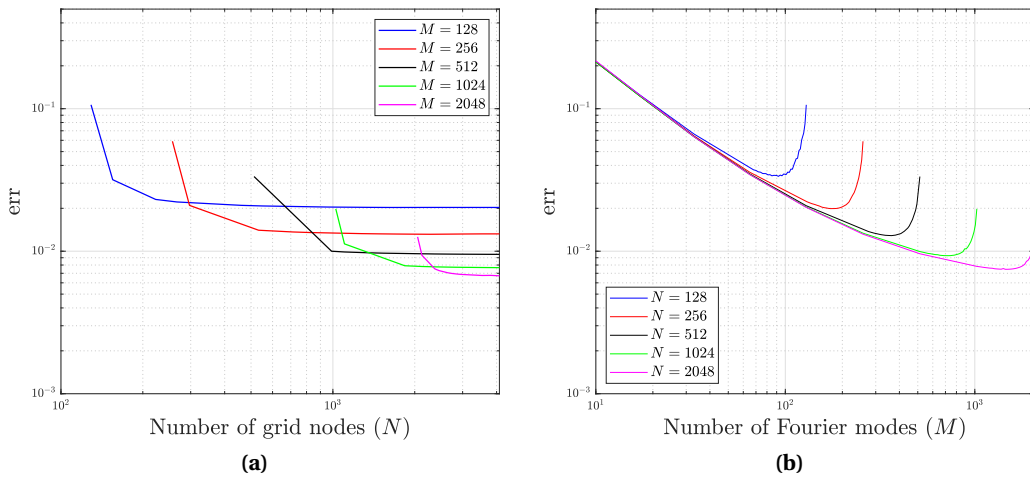


Figure 6. Error analysis on the local fields. (a) Influence of the number of integration points N and (b) Influence of the number of modes M .

In addition, since FFT-based methods are generally used to calculate the overall properties of composites, it is also interesting to study the influence of the discretization on the macroscopic conductivity $\bar{c}$ which arises in the relation

$$\langle \mathbf{J} \rangle = \bar{c} \langle \mathbf{E} \rangle, \quad (70)$$

where $\langle \mathbf{J} \rangle$ and $\langle \mathbf{E} \rangle$ are the volume averages of the electric current and electric field, respectively. The macroscopic conductivity $\bar{c}$, obtained from the numerical calculation, denoted by $\bar{c}^{\text{num}}$, is simply obtained using the Fourier coefficients of the electric current and electric field associated to the first Fourier frequencies, since the first Fourier frequencies are associated to the average of the considered field. For the loading considered, one thus has (using the notation of equation (32))

$$\bar{c}^{\text{num}} = \frac{\hat{J}_{00}^x}{\hat{E}_{00}^x}. \quad (71)$$

The theoretical macroscopic conductivity, denoted by $\bar{c}^{\text{th}}$, is given by [34]:

$$\bar{c}^{\text{th}} = \sqrt{\frac{c_1 + 3c_2}{3c_1 + c_2}}. \quad (72)$$

Then we define the relative error for the macroscopic conductivity as

$$\text{err}_{\text{macro}} = \frac{|\bar{c}^{\text{th}} - \bar{c}^{\text{num}}|}{\bar{c}^{\text{th}}}. \quad (73)$$

First, the error for the macroscopic conductivity is represented in Figure 7a as a function of the number of grid nodes N (for a fixed number of modes M). A stationary value of the error is observed when increasing the number of nodes, which act here as integration points. This behaviour is consistent with the convergence of the numerical integration as the number of grid points increases. For the considered test case, the macroscopic error is minimal when the integration grid matches the spectral truncation, i.e. ($N = M$). This behaviour is attributed to a partial compensation between truncation and numerical integration errors. It is therefore problem-dependent and not expected to hold in general.

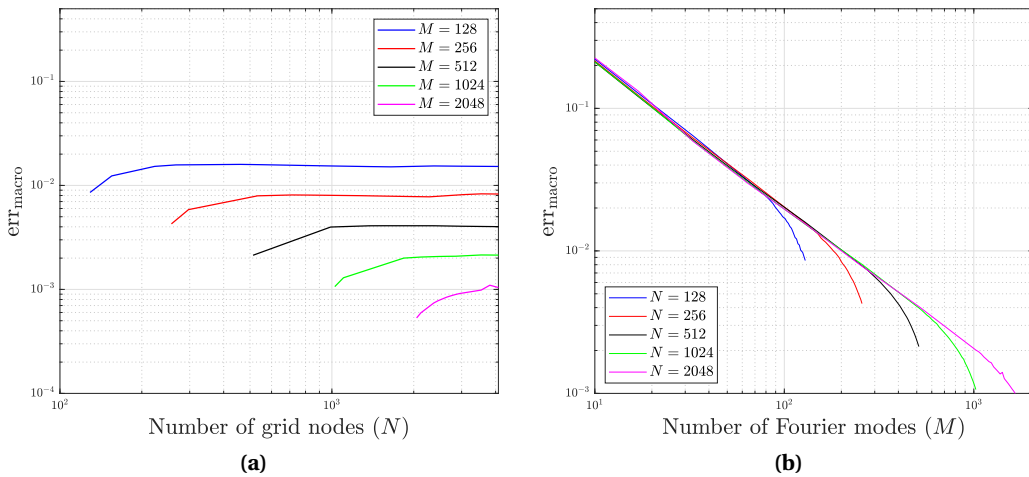


Figure 7. Error analysis on the macroscopic conductivity. (a) Influence of the number of integration points N and (b) Influence of the number of modes M .

Then, the error for the macroscopic conductivity is represented in Figure 7b as a function of the number of modes M (for a given number of nodes N). Contrary to the error related to the local fields, the error for the macroscopic conductivity keeps decreasing while increasing the number of modes. Therefore, the estimation of macroscopic properties is not affected by the high frequencies, in contrast to the local fields, which may be affected by Gibbs phenomenon, while macroscopic properties correspond to the zero Fourier mode of the solution, as shown by equation (71). Nevertheless, the zero Fourier mode results from the solution of the Lippmann–Schwinger equation, which involves a convolution in Fourier space and therefore depends on the full spectral content of the fields. Thus, this result is a numerical observation, which might be case-dependent.

5.5. Influence of the contrast on the optimal ratio M/N

The above results show that the choice $M = N$ is not necessarily the optimal choice in terms of quality of the local fields, which is a peculiar and unexpected emerging result. For this specific case, i.e., checkerboard microstructure with a contrast $c_2/c_1 = 100$, the optimal value M/N associated to the relative error in energy norm is almost constant and lies between 0.67 and 0.72.

This emerging result is not universal as it is expected to depend on both the microstructure and the contrast. Indeed, in the limit case of a homogeneous material, i.e., $c_2/c_1 = 1$, no oscillations are expected and thus the improvement of the solution field obtained by taking M lower than N is not expected to hold in that case. To assess quantitatively the effect of the contrast on the optimal value M/N , we consider several values of the contrast $c_2/c_1 = [2, 10, 100, 1000]$ and only one value of the number of grid points $N = 512$ is considered.

The error is represented in Figure 8 as a function of the number of modes M (for a given number of nodes $N = 512$). The error decreases when the number of modes M increases and, depending on the value of the contrast, it can reach a limit value M_{lim} above which the error increases. As expected the optimal value of the ratio M/N is close to 1 for low-contrasted materials, and then it decreases when the contrast increases. For this specific problem this is attributed to the presence of oscillations, whose occurrence depends on the contrast.

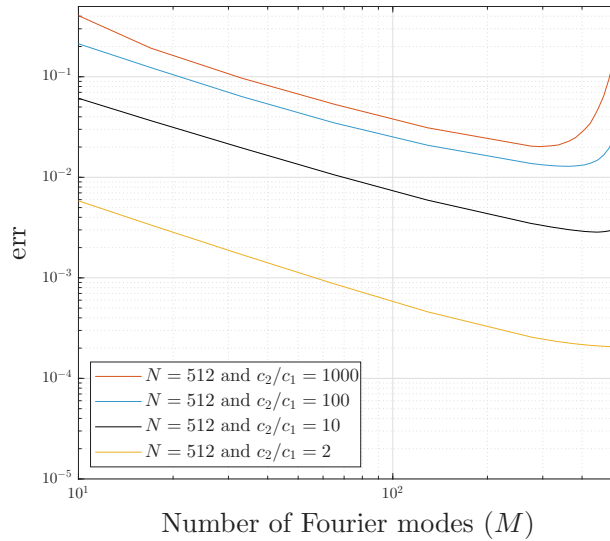


Figure 8. Influence of the contrast upon the error analysis on the local fields.

The emergence of this optimal value M/N cannot be seen as a universal result and is expected to be problem-dependent: the only conclusion that can be drawn with this analysis is that the choice $M = N$ is not necessarily the optimal choice for local fields.

6. Discussion on spurious oscillations

In this section, we discuss the occurrence of spurious oscillations, which are the main source of discrepancies between the numerical and analytical solutions, and present remedies based either on discrete Green operators or on smoothing techniques.

6.1. On generalized Green operators

The three approaches considered in Section 3 all rely on Fourier series together with standard derivative rules which leads to the so-called “continuous” Green operator defined by equation (35). It has been shown in the literature that modified “discrete” Green operators, obtained using finite

differences or finite elements approximations, generally improve the convergence rate of the iterative procedure as well as the quality of the local fields, by suppressing most of the spurious oscillations [18,20,21]. *It should be noted that the developments of Green operators are thus not compatible as is with the description of fields using Fourier series*; the use of discrete Fourier transforms simply arise in the resolution of the linear system obtained using either finite differences or finite elements (which defines a circulant matrix in the case of periodic boundary conditions), leading to the so-called preconditioned FFT-accelerated finite element solvers.

In this section we revisit the construction of the discrete Green operator by introducing *modified derivative rules* into the framework of Section 3, still in the framework of Fourier series. Indeed, instead of using the standard spectral derivative rules, we will consider a difference (continuous) operator with a given spacing (not necessarily equal to the grid size), based on difference quotient. In order to retrieve [18]'s scheme and generalize it, we will consider a forward-type derivative scheme for the electric potential and a backward-type derivative scheme for the current. Other choice for these derivative rules would lead to another modified Green operators. This enables the use of discrete Green operators within the Fourier series framework (i.e. using spectral interpolation) and allows to determine whether oscillations originate from the derivative rule or the interpolation (spectral) functions.

Modified derivative rules for the electric potential. We consider the following *continuous* forward derivative scheme for the electric potential

$$\begin{aligned}\delta_{h_x}^{x+} [\varphi^*] (x, y, z) &= \frac{\varphi^*(x + h_x, y, z) - \varphi^*(x, y, z)}{h_x} \\ \delta_{h_y}^{y+} [\varphi^*] (x, y, z) &= \frac{\varphi^*(x, y + h_y, z) - \varphi^*(x, y, z)}{h_y} \\ \delta_{h_z}^{z+} [\varphi^*] (x, y, z) &= \frac{\varphi^*(x, y, z + h_z) - \varphi^*(x, y, z)}{h_z}\end{aligned}\quad (74)$$

where h_x , h_y and h_z are the spatial steps considered in each spatial directions.

Modified derivative rules for the electric current. We then introduce a *continuous* backward derivative scheme

$$\begin{aligned}\delta_{h_x}^{x-} [J_x] (x, y, z) &= \frac{J_x(x, y, z) - J_x(x - h_x, y, z)}{h_x} \\ \delta_{h_y}^{y-} [J_y] (x, y, z) &= \frac{J_y(x, y, z) - J_y(x, y - h_y, z)}{h_y} \\ \delta_{h_z}^{z-} [J_z] (x, y, z) &= \frac{J_z(x, y, z) - J_z(x, y, z - h_z)}{h_z}.\end{aligned}\quad (75)$$

Some comments are in order:

- In the case $h_x = \Delta x$, $h_y = \Delta y$ and $h_z = \Delta z$, the derivative rules are very similar to the finite difference scheme of [18], but the difference is that *it is defined even between nodal values* in the present formulation.
- In the case $h_x \rightarrow 0$, $h_y \rightarrow 0$, $h_z \rightarrow 0$ the derivative rules considered correspond to the standard derivative rules considered previously.

Assuming that φ^* is described by a Full Fourier series, the forward derivative scheme (74) in the x -direction reads

$$\begin{aligned}\delta_{h_x}^{x+} [\varphi^*] (x, y, z) &= \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \Phi_{ijk}^* \frac{\exp(i\xi_i^x(x + h_x)) - \exp(i\xi_i^x x)}{h_x} \exp(i\xi_j^y y) \exp(i\xi_k^z z) \\ &= \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \Phi_{ijk}^* \frac{\exp(i\xi_i^x h_x) - 1}{h_x} \exp(i\xi_i^x x) \exp(i\xi_j^y y) \exp(i\xi_k^z z).\end{aligned}\quad (76)$$

Assuming that τ_x is described by a full Fourier series, the backward derivative scheme (75) in the x -direction reads

$$\begin{aligned}\delta_{h_x}^{x-}[\tau_x](x, y, z) &= \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} T_{ijk}^x \frac{\exp(\iota \xi_i^x x) - \exp(\iota \xi_i^x (x - h_x))}{h_x} \exp(\iota \xi_j^y y) \exp(\iota \xi_k^z z) \\ &= \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} T_{ijk}^x \frac{1 - \exp(-\iota \xi_i^x h_x)}{h_x} \exp(\iota \xi_i^x x) \exp(\iota \xi_j^y y) \exp(\iota \xi_k^z z).\end{aligned}\quad (77)$$

Then, considering these derivative rules and the description of the fields using full Fourier series as in Section 3, the auxiliary problem reduces, after a few straightforward calculations involving trigonometric relations, to

$$\begin{aligned}c_0 \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \left((\alpha_i^x)^2 + (\alpha_j^y)^2 + (\alpha_k^z)^2 \right) \Phi_{ijk}^* \exp(\iota \xi_i^x x) \exp(\iota \xi_j^y y) \exp(\iota \xi_k^z z) \\ = - \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \iota \left(\beta_i^x T_{ijk}^x + \beta_j^y T_{ijk}^y + \beta_k^z T_{ijk}^z \right) \exp(\iota \xi_i^x x) \exp(\iota \xi_j^y y) \exp(\iota \xi_k^z z)\end{aligned}\quad (78)$$

where α_i^x , α_j^y , and α_k^z are given by

$$\alpha_i^x = \frac{2 \sin\left(\frac{\xi_i^x h_x}{2}\right)}{h_x}, \quad \alpha_j^y = \frac{2 \sin\left(\frac{\xi_j^y h_y}{2}\right)}{h_y}, \quad \alpha_k^z = \frac{2 \sin\left(\frac{\xi_k^z h_z}{2}\right)}{h_z}, \quad (79)$$

and β_i^x , β_j^y , and β_k^z are given by

$$\beta_i^x = \frac{1 - \exp(-\iota \xi_i^x h_x)}{\iota h_x}, \quad \beta_j^y = \frac{1 - \exp(-\iota \xi_j^y h_y)}{\iota h_y}, \quad \beta_k^z = \frac{1 - \exp(-\iota \xi_k^z h_z)}{\iota h_z}. \quad (80)$$

The relation between the Fourier series coefficients of the fluctuation and the components of the polarization then follows

$$\Phi_{ijk}^* = -\iota \frac{\beta_i^x T_{ijk}^x + \beta_j^y T_{ijk}^y + \beta_k^z T_{ijk}^z}{c_0 \left((\alpha_i^x)^2 + (\alpha_j^y)^2 + (\alpha_k^z)^2 \right)}. \quad (81)$$

As previously this relation holds for all frequencies, i.e. for $i = -\infty, \dots, +\infty$, $j = -\infty, \dots, +\infty$ and $k = -\infty, \dots, +\infty$ (except in the particular case $i = j = k = 0$ where one has $\Phi_{000}^* = 0$), and is expressed in terms of the Fourier series coefficients which need to be evaluated. Using discrete Fourier transforms to evaluate these coefficients with $(N_x + 1) \times (N_y + 1) \times (N_z + 1)$ voxels (so that the spatial scales are $\Delta x = L_x/(N_x + 1)$, $\Delta y = L_y/(N_y + 1)$ and $\Delta z = L_z/(N_z + 1)$), one has

$$\widehat{\Phi}_{ijk}^* = -\iota \frac{\beta_i^x \widehat{T}_{ijk}^x + \beta_j^y \widehat{T}_{ijk}^y + \beta_k^z \widehat{T}_{ijk}^z}{c_0 \left((\alpha_i^x)^2 + (\alpha_j^y)^2 + (\alpha_k^z)^2 \right)}, \quad (82)$$

but now for $i = -N_x/2, \dots, N_x/2$, $j = -N_y/2, \dots, N_y/2$ and $k = -N_z/2, \dots, N_z/2$, except in the particular case $i = j = k = 0$ where one has $\widehat{\Phi}_{000}^* = 0$. Equation (82) provides the coefficients of the discrete Fourier transform of φ^* as a function of the coefficients of the discrete Fourier

transform of τ_x , τ_y and τ_z . It then can be used to define a *generalized Green operator* on the electric field as

$$\begin{cases} \hat{E}_{ijk}^x = -\beta_i^{x*} \frac{\beta_i^x \hat{T}_{ijk}^x + \beta_j^y \hat{T}_{ijk}^y + \beta_k^z \hat{T}_{ijk}^z}{c_0 \left((\alpha_i^x)^2 + (\alpha_j^y)^2 + (\alpha_k^z)^2 \right)} \\ \hat{E}_{ijk}^y = -\beta_j^{y*} \frac{\beta_i^x \hat{T}_{ijk}^x + \beta_j^y \hat{T}_{ijk}^y + \beta_k^z \hat{T}_{ijk}^z}{c_0 \left((\alpha_i^x)^2 + (\alpha_j^y)^2 + (\alpha_k^z)^2 \right)} \\ \hat{E}_{ijk}^z = -\beta_k^{z*} \frac{\beta_i^x \hat{T}_{ijk}^x + \beta_j^y \hat{T}_{ijk}^y + \beta_k^z \hat{T}_{ijk}^z}{c_0 \left((\alpha_i^x)^2 + (\alpha_j^y)^2 + (\alpha_k^z)^2 \right)}, \end{cases} \quad (83)$$

for $i = -N_x/2, \dots, N_x/2$, $j = -N_y/2, \dots, N_y/2$ and $k = -N_z/2, \dots, N_z/2$ (except in the case $i = j = k = 0$). In the case $i = j = k = 0$, one has

$$\hat{E}_{000}^x = \bar{E}_x, \quad \hat{E}_{000}^y = \bar{E}_y, \quad \hat{E}_{000}^z = \bar{E}_z, \quad (84)$$

where $\bar{E}_x$, $\bar{E}_y$ and $\bar{E}_z$ are the components of the average of the electric field $\bar{\mathbf{E}}$. In equation (83), β_i^{x*} , β_j^{y*} and β_k^{z*} are respectively the complex conjugate of β_i^x , β_j^y and β_k^z and $\hat{E}_{ijk}^x$, $\hat{E}_{ijk}^y$ and $\hat{E}_{ijk}^z$ are the coefficients of the discrete Fourier transforms of the components of the local electric field $\mathbf{E}$. In addition, it is worth noting that

$$\beta_i^{x*} \beta_i^x = (\alpha_i^x)^2, \quad (85)$$

so one finally has

$$\hat{\mathbf{E}}(\boldsymbol{\xi}) = -\hat{\Gamma}_G^0(\boldsymbol{\xi}) : \hat{\boldsymbol{\tau}}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \neq \mathbf{0}, \quad \hat{\mathbf{E}}(\mathbf{0}) = \bar{\mathbf{E}}, \quad (86)$$

where $\hat{\Gamma}_G^0$ is the generalized Green operator which has an explicit form in Fourier space:

$$\hat{\Gamma}_G^0(\boldsymbol{\xi}) = \frac{\boldsymbol{\beta}^*(\boldsymbol{\xi}) \otimes \boldsymbol{\beta}(\boldsymbol{\xi})}{c^0 |\boldsymbol{\beta}^*(\boldsymbol{\xi}) \cdot \boldsymbol{\beta}(\boldsymbol{\xi})|^2}. \quad (87)$$

Some comments are in order:

- In the case $h_x \rightarrow 0$, $h_y \rightarrow 0$, $h_z \rightarrow 0$, it is easy to show that $\boldsymbol{\beta}(\boldsymbol{\xi}) \rightarrow \boldsymbol{\xi}$ and therefore $\hat{\Gamma}_G^0(\boldsymbol{\xi}) \rightarrow \hat{\Gamma}^0(\boldsymbol{\xi})$. This result is expected as in this case, the modified derivative rules reduce to standard derivative rules, so the generalized Green operator reduces to the standard “continuous” Green operator.
- In the case $h_x = \Delta x$, $h_y = \Delta y$ and $h_z = \Delta z$, the generalized Green operator reduces to the so-called “discrete” Green operator of [18], which is also expected as the derivation rules are the same. The main difference is that [18]’s scheme is based on finite differences, and therefore the solution field is only defined as the nodes, while the present scheme provides the same nodal values but the solution fields are described by Fourier series and therefore are defined in the whole domain.
- In the cases $0 < h_x < \Delta x$, $0 < h_y < \Delta y$ and $0 < h_z < \Delta z$, the associated Green operator generalizes the “continuous” and “discrete” Green operators.

The effect of the Green operator, obtained using modified derivation rule, is now investigated on the checkerboard problem considered in Section 5. We consider the case $N = M = 1024$ and several values for the spatial steps used for the derivative rules:

- Case 1: $h_x = h_y \rightarrow 0$, which corresponds to the (standard) continuous Green operator.
- Case 2: $h_x = h_y = \Delta x/2 = \Delta y/2$.
- Case 3: $h_x = h_y = \Delta x = \Delta y$, which corresponds to the discrete Green operator of [18];

The solution field at the vicinity of the interface between the two phases is plotted in Figure 9, on the line $x_2 = 0$.

The very first observation is that the predictions obtained with the three considered Green operators lead to the almost exact same results when the solution is represented on the fine grid

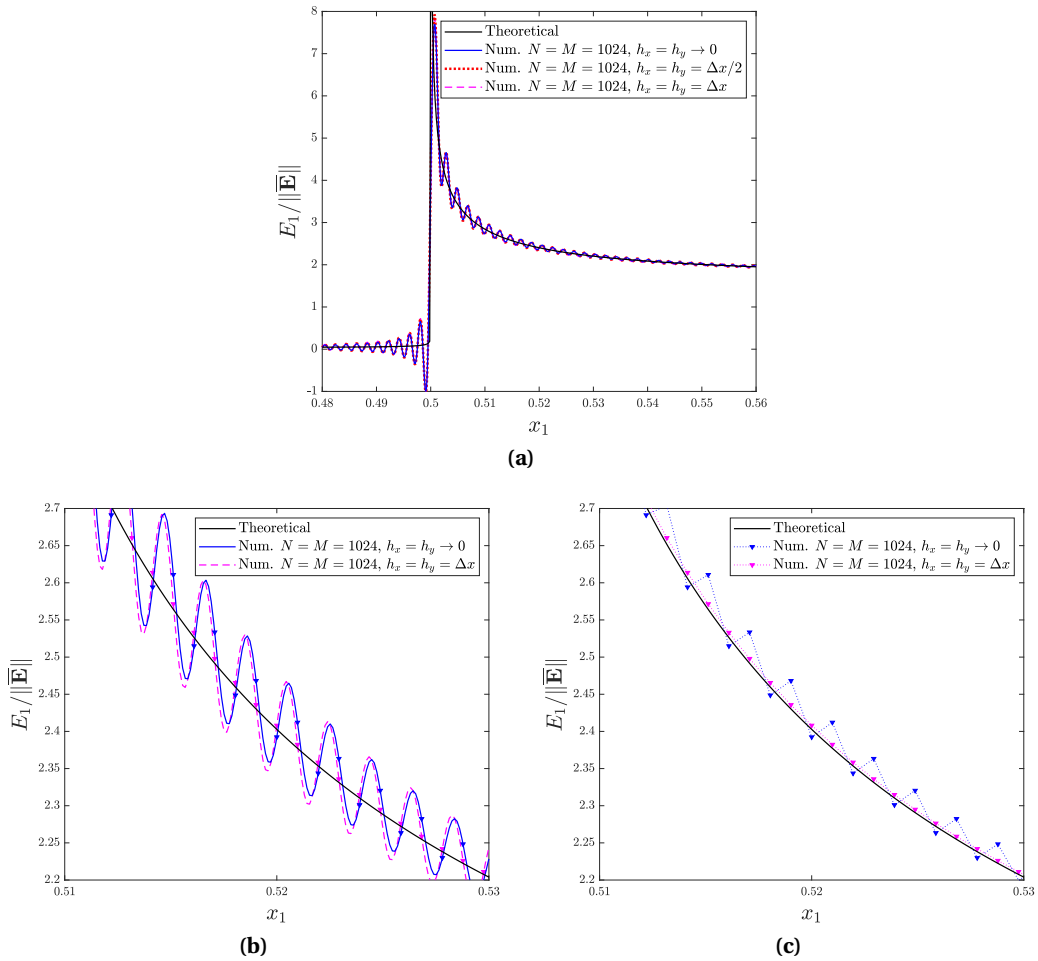


Figure 9. Effect of the Green operator on the normalized electric field $E_1 / \|\bar{\mathbf{E}}\|$. (a) Line plot at the interface between the two phases ($x_2 = 0$), (b) Enlarged view, (c) Representation with a linear interpolation between grid node values.

(Figure 9a). Interestingly, an enlarged view (Figure 9b) shows that the grid nodes values calculated with the continuous and discrete Green operators are not the same, although the continuous functions (projected on the fine grid) almost coincide. It is worth noting that, if the solution is interpolated linearly between grid nodes (Figure 9c), the oscillatory nature of the solution is very different using the continuous or discrete operator: the solution calculated with the discrete operator does not exhibit oscillations using such representation. It is important to remark that this representation (using linear interpolation) is not consistent with the framework considered in this work but it is consistent with a finite element discretization of the fields.

Therefore, the partial conclusion that can be made is that discrete Green operators do not improve intrinsically the local fields; it is the underlying discretization, based on finite elements and leading to discrete Green operators, which may improve the quality of the local fields.

6.2. Discussion on smoothing techniques

Another solution for removing oscillation is to consider a smoother. Indeed, the presentation of the results between spatial nodes clearly shows the oscillatory nature of functions described by Fourier series in presence of discontinuities, which is related to the Gibbs phenomenon. We explore in this section the effect of material properties smoothing using the periodic smoother proposed by [24]. This smoother is based on a modification of the Fourier coefficients of the function to be smoothed, which is briefly recalled hereafter in 1-d. We consider periodic unidimensional function f with period L , defined in the interval $[0, L]$ discretized with $N + 1$ segments so the spatial scale is $\Delta x = L/(N + 1)$. The discrete values of function f denoted $f_j = f(x_j)$ may thus be written as an array $\mathbf{f}$ of size N . The smooth estimate is denoted by $\tilde{f}$, and we are thus looking for its discrete values $\tilde{f}_i$ which consists in finding the array $\tilde{\mathbf{f}} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_N)$. The smoother proposed by [24] is based on smoothing splines solution of the minimization of a functional which balances the fidelity to the data, through the residual sum-of-squares (RSS), and the smoothness of the estimate, through some penalty term. In the particular case of the first derivative considered in the penalty term, the smoother reads

$$\tilde{\mathbf{f}} = \text{IDFT}(\mathbf{\Gamma}(\bar{s}) \circ \text{DFT}(\mathbf{f})), \quad (88)$$

where $\circ$ denoted Hadamard product (pointwise product) and $\mathbf{\Gamma}(\bar{s})$ is an array of size N , whose components are given by

$$\Gamma_i = \frac{1}{1 + \bar{s} L^2 \xi_i^2}. \quad (89)$$

In this equation, $\bar{s}$ is a parameter that controls the smoothing, i.e. the tradeoff between the “smoothness” of the solution and the fidelity to the data. A given value of $\bar{s}$ will create some interphase between material discontinuities: the increase of $\bar{s}$ increases the thickness of this interphase.

The objective is now to study the effect of this smoother on the local fields, especially on the occurrence of oscillations. Several spatial discretizations are considered; in order to assess the ability of the smoother to reproduce and get closer to the analytical solution, the thickness of the interphase is chosen so as to decrease with the spatial discretization. We thus take the following form for the smoothing parameter

$$\bar{s} = \frac{a}{N^2}. \quad (90)$$

This allows to have a finite interphase for small values of N and the size of this interphase tends to zero when the spatial discretization increases. The microstructure obtained increasing the number of nodes N tends to the initial microstructure (with discontinuities): the local fields are thus expected to get closer to the analytical solution, when N increases, without numerical artifacts. Several values for the parameter a will be considered.

We begin with the case $\bar{s} = 1/N^2$ (i.e. $a = 1$). First, the microstructure is represented in Figure 10a for several values of the discretization (and thus the smoothing parameter); the node values are represented by discrete symbols and the associated continuous representation is also shown (based on a Fourier interpolation using the node values to calculate the Fourier coefficients and then the Fourier coefficients to reconstruct the function everywhere). Since the smoothing parameter has been taken to decrease when the number of nodes N increases, the thickness of the interphase decreases when N increases and the smooth microstructure tends to the theoretical one.

The line plot of the solution field $E_1 / \|\bar{\mathbf{E}}\|$ at the interface between the two phases ($x_2 = 0$) is then represented in Figure 10b. As expected by the effect of smoothing, (i) the solution is spread due to the presence of an interphase of properties and (ii) spurious oscillations are considerably decreased by comparison with the solution calculated without smoothing (see e.g. Figure 3).

The smoothing parameter has been chosen so as to remove most of the oscillations without introducing too much spreading, for a given value of N ; an increase of the smoothing parameter would necessarily decrease the magnitude of the oscillations but increase the spreading of the solution: a compromise between oscillations and spreading is inevitable. It is worth noting that the numerical solution seems to tend to the analytical one, without numerical artifacts, with an increase of the spatial resolution as well as the introduction of a sufficient (but small enough) smoothing parameter.

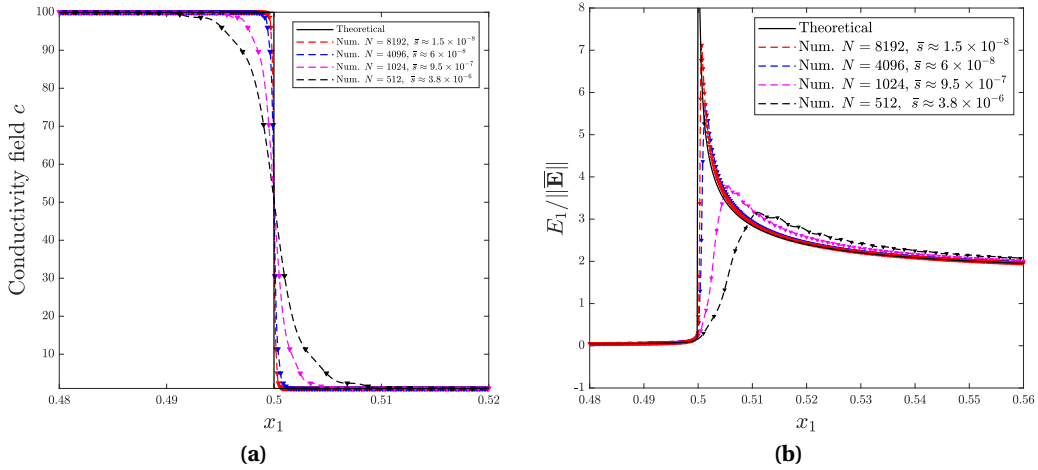


Figure 10. Effect of smoothing on the local fields. (a) Line plot of the conductivity field c at the center of the cell ($x_2 = 0.5$). (b) Line plot of the normalized electric field $E_1 / \|\vec{E}\|$ at the interface between the two phases ($x_2 = 0$). The continuous lines corresponds to the solution on the fine grid and the triangles corresponds to the solution on the grid used for the FFT calculations.

A quantitative comparison between the numerical and analytical solutions is then made but studying the error defined by equation (69), for $\bar{s} = 1/N^2$ and $\bar{s} = 1/(200 \times N^2)$. The error is represented in Figure 11a as a function of the number of modes M (for a given number of nodes N) and is compared to that obtained without smooth. In contrast to the original case where the error starts increasing for high frequencies (when $M > M_{\text{lim}}$), the error decreases and reaches a stationary value for a smooth microstructure. This confirms that oscillations due to high frequencies do not degrade the solution field because they are greatly decreased in the smooth case. It is worth noting that the error, in this case, is greater than the error of the non-smooth case. This is due to the fact that the smooth microstructure is not strictly equivalent to the non-smooth case because of the thin interface induced by smoothing; therefore the problem solved is slightly different which necessarily induces errors in comparison with the analytical solution. Moreover, this result was expected as the non-smoothed solution, without integration error ($N \gg M$), is optimal in the sense of the error defined by (69). In a way, the non-smoothed solution optimized this error while disregarding the local error, therefore the spurious oscillation. This is confirmed by considering a smaller value for the smoothing parameter, $\bar{s} = 1/(200 \times N^2)$. In this case, the error is represented in Figure 11b, again as a function of the number of modes M (for a given number of nodes N). For this smaller value of the smoothing parameter, the error keeps decreasing when increasing the number of modes and is smaller than in the case without smoothing. This indicates that decreasing the interphase improves the local fields because (i) the microstructure is closer to the analytical one and (ii) oscillations are still removed by the smoother. It should be noted that, compared to the truncation of high frequencies

considered previously (i.e. for $M \sim M_{\text{lim}}$), the error defined by (69) is not significantly improved by the smoother; this can be explained by the fact that filtering high frequencies and smoothing material discontinuities play somehow a similar role, as both acts as a regularization of spurious oscillations due to high frequencies in presence of discontinuities.

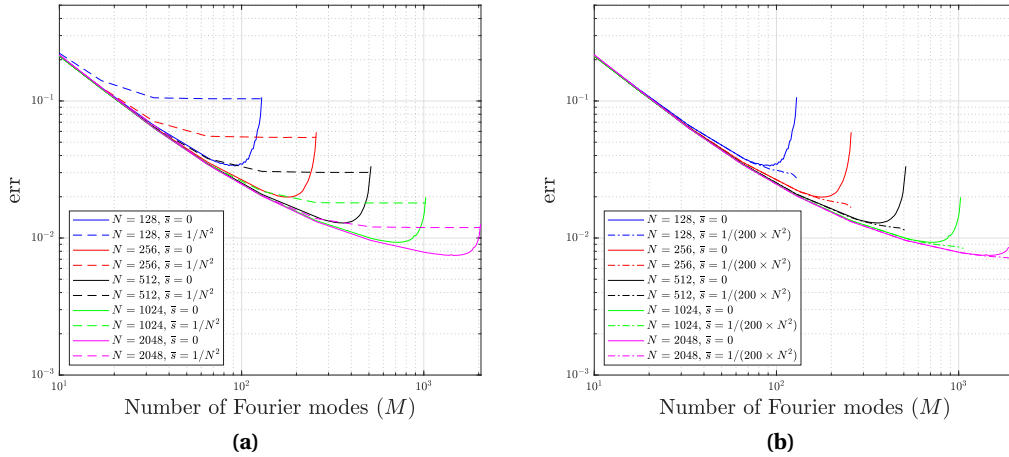


Figure 11. Influence of material properties smoothing on error analysis on the local fields. (a) Case $\bar{s} = 1/N^2$ and (b) Case $\bar{s} = 1/(200 \times N^2)$.

For completeness, the influence of the smoother on the macroscopic conductivity is represented in Figure 12. As for the error on the local fields, the error on the macroscopic conductivity decreases with the increase of M and reaches a stationary value for a smooth microstructure. Overall, independently of the value of the smoothing parameter, the use of the smoother slightly alters the predicted effective properties. This behaviour is attributed to the modification of the microstructure induced by the smoothing operation, meaning that the resulting computation corresponds to a slightly different physical problem.

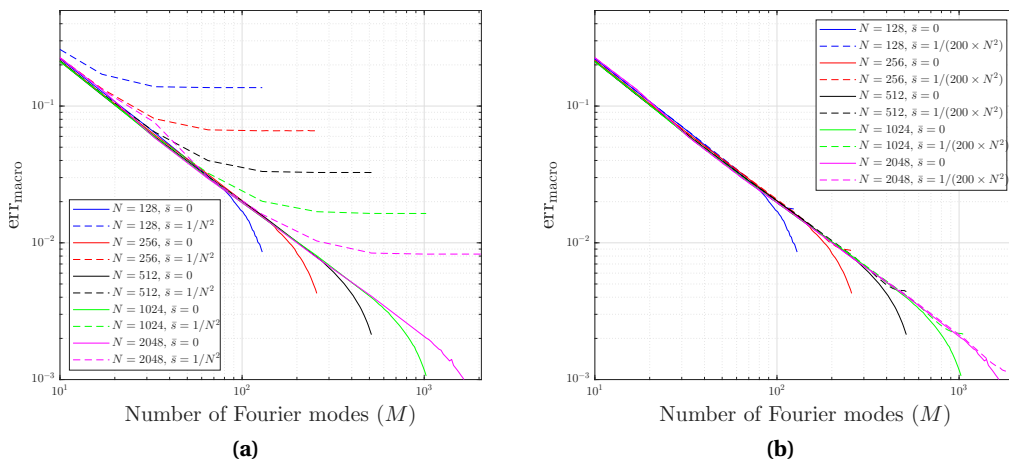


Figure 12. Influence of material properties smoothing on error analysis on the macroscopic conductivity. (a) Case $\bar{s} = 1/N^2$ and (b) Case $\bar{s} = 1/(200 \times N^2)$.

7. Conclusion

The aim of this work was to revisit the formulation of FFT-based methods for heterogeneous materials in the periodic setting, originally proposed by the pioneering work of [1]. This class of methods is based on the iterative resolution of an auxiliary problem involving a reference homogeneous material and a polarization tensor. Assuming a description of the fields by (infinite or partial) Fourier series, we show the equivalence between several discretization approaches including strong and weak formulations and different numerical approximations, following ideas from [12]. The formulation proposed in this work notably decouples the number of Fourier modes to the spatial discretization grid which is shown to be used to calculate approximately integrals with the trapezoidal rule. The relation between the discretization grid and the representation of the continuous fields is discussed including the possibility of using non-uniform grids. Several numerical experiments are conducted on the checkerboard problem for which an analytical solution is used; we notably investigate the effect of the number of Fourier modes together with the grid discretization. The importance of representing the fields between grid nodes is emphasized as spurious oscillations can emerge between grid nodes. Increasing the number of grid nodes, for a given number of Fourier modes, has only a moderate effect as it simply improves the calculations of the Fourier coefficients. Conversely, for this specific checkerboard problem, increasing the number of Fourier modes, for a given grid discretization, improves the solution up to a limit value for the number of modes, above which the quality of the solution is degraded due to the occurrence of spurious oscillations at high frequencies. In this continuous setting, i.e. by considering Fourier series to describe fields, generalized Green operators are derived using modified derivative rules, which leads to a generalization of both continuous and discrete Green operators relying on a single spatial parameter. It was shown that this does not improve the local fields, as oscillations are mainly due to the description of solution fields using Fourier series. On the other hand, the effect of material discontinuity smoothing is studied by using [24]'s interface smoother: it is shown that a combination of mesh refinement and suitable smoothing improves the numerical solution without numerical artifacts even between grid nodes. Overall, the standard choice $M = N$ is not necessarily optimal.

It is important to remark that this work was concerned with a description of fields based on Fourier series, even when (generalized) discrete Green operators were used; this is mainly why spurious oscillations between nodes were observed in that case. Standard approaches making use of discrete Green operators are generally based either on finite differences [18,19] or a finite element discretization, generally using linear elements [20,21]. Therefore in that case, the obtained fields should be post-treated according to the interpolation basis functions used and may not show any oscillations between the node grid used for evaluating the Fourier transforms as shown using a linear interpolation between grid node values. Oscillations are thus expected to be less important using these approaches because of the representation of the solution field.

This work provides a new perspective on solution fields obtained by FFT-based solvers, especially on the quality of the local fields. In presence of material discontinuities, spurious oscillations inherent of the Fourier framework degrade the solution field. Simple remedies consist in (i) filtering the high frequencies, by simply removing some Fourier coefficients as a post-treatment to reconstruct the solution field or (ii) removing material discontinuities using an interface spreading smoother.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

- [1] H. Moulinec and P. Suquet, "A numerical method for computing the overall response of nonlinear composites with complex microstructure", *Comput. Methods Appl. Mech. Eng.* **157** (1998), no. 1–2, pp. 69–94.
- [2] M. Schneider, "A review of nonlinear FFT-based computational homogenization methods", *Acta Mech.* **232** (2021), no. 6, pp. 2051–2100.
- [3] S. Lucarini, M. V. Upadhyay and J. Segurado, "FFT based approaches in micromechanics: fundamentals, methods and applications", *Model. Simul. Mat. Sci. Eng.* **30** (2021), no. 2, article no. 023002.
- [4] L. Gélébart, *AMITEX*, 2022. Online at <https://amitexfft.github.io/AMITEX/index.html>.
- [5] R. Lebensohn, "N-site modeling of a 3D viscoplastic polycrystal using Fast Fourier Transform", *Acta Mater.* **49** (2001), no. 14, pp. 2723–2737.
- [6] R. Brenner, A. J. Beaudoin, P. Suquet and A. Acharya, "Numerical implementation of static Field Dislocation Mechanics theory for periodic media", *Philos. Mag.* **94** (2014), no. 16, pp. 1764–1787.
- [7] N. Bertin, M. V. Upadhyay, C. Pradalier and L. Capolungo, "A FFT-based formulation for efficient mechanical fields computation in isotropic and anisotropic periodic discrete dislocation dynamics", *Model. Simul. Mat. Sci. Eng.* **23** (2015), no. 6, article no. 065009.
- [8] D. J. Eyre and G. W. Milton, "A fast numerical scheme for computing the response of composites using grid refinement", *Eur. Phys. J. AP* **6** (1999), no. 1, pp. 41–47.
- [9] R. Brenner, "Numerical computation of the response of piezoelectric composites using Fourier transform", *Phys. Rev. B* **79** (2009), no. 18, article no. 184106.
- [10] N. Bilger, F. Auslender, M. Bornert, J.-C. Michel, H. Moulinec, P. Suquet and A. Zaoui, "Effect of a nonuniform distribution of voids on the plastic response of voided materials: a computational and statistical analysis", *Int. J. Solids Struct.* **42** (2005), no. 2, pp. 517–538.
- [11] J. Paux, R. Brenner and D. Kondo, "Plastic yield criterion and hardening of porous single crystals", *Int. J. Solids Struct.* **132–133** (2018), pp. 80–95.
- [12] J. Vondřejc, J. Zeman and I. Marek, "An FFT-based Galerkin method for homogenization of periodic media", *Comput. Math. Appl.* **68** (2014), no. 3, pp. 156–173.
- [13] L. Risthaus and M. Schneider, "Imposing Dirichlet boundary conditions directly for FFT-based computational micromechanics", *Comput. Mech.* **74** (2024), no. 5, pp. 1089–1113.
- [14] L. Morin and J. Paux, "A fast numerical method for the conductivity of heterogeneous media with Dirichlet boundary conditions based on discrete sine–cosine transforms", *Comput. Methods Appl. Mech. Eng.* **421** (2024), article no. 116772.
- [15] L. Risthaus and M. Schneider, "FFT-based computational micromechanics with Dirichlet boundary conditions on the rotated staggered grid", *Int. J. Numer. Methods Eng.* **125** (2024), no. 21, article no. e7569.
- [16] J. Paux, L. Morin and L. Gélébart, "A Discrete Sine-Cosine Transforms Galerkin Method for the Conductivity of Heterogeneous Materials With Mixed Dirichlet/Neumann Boundary Conditions", *Int. J. Numer. Methods Eng.* **126** (2025), no. 1, article no. e7615.
- [17] J. Paux, L. Morin, L. Gélébart and A. M. Amadou Sanoko, "A discrete sine–cosine based method for the elasticity of heterogeneous materials with arbitrary boundary conditions", *Comput. Methods Appl. Mech. Eng.* **433** (2025), article no. 117488.
- [18] F. Willot, B. Abdallah and Y.-P. Pellegrini, "Fourier-based schemes with modified Green operator for computing the electrical response of heterogeneous media with accurate local fields", *Int. J. Numer. Methods Eng.* **98** (2014), no. 7, pp. 518–533.
- [19] F. Willot, "Fourier-based schemes for computing the mechanical response of composites with accurate local fields", *Comptes Rendus. Mécanique* **343** (2015), no. 3, pp. 232–245.
- [20] J. Zeman, T. W. J. de Geus, J. Vondřejc, R. H. J. Peerlings and M. G. D. Geers, "A finite element perspective on nonlinear FFT-based micromechanical simulations", *Int. J. Numer. Methods Eng.* **111** (2017), no. 10, pp. 903–926.
- [21] M. Schneider, D. Merkert and M. Kabel, "FFT-based homogenization for microstructures discretized by linear hexahedral elements", *Int. J. Numer. Methods Eng.* **109** (2017), no. 10, pp. 1461–1489.
- [22] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*, Springer, 2006.
- [23] L. N. Trefethen, *Spectral methods in MATLAB*, Society for Industrial and Applied Mathematics, 2000.
- [24] L. Morin, R. Brenner, K. Derrien and K. Dorhmi, "Periodic smoothing splines for FFT-based solvers", *Comput. Methods Appl. Mech. Eng.* **373** (2021), article no. 113549.
- [25] M. Frigo and S. G. Johnson, "FFTW: an adaptive software architecture for the FFT", in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, vol. 3, IEEE Press, 1998, pp. 1381–1384.
- [26] L. N. Trefethen and J. A. C. Weideman, "The exponentially convergent trapezoidal rule", *SIAM Rev.* **56** (2014), no. 3, pp. 385–458.

- [27] S. Lucarini and J. Segurado, "DBFFT: A displacement based FFT approach for non-linear homogenization of the mechanical behavior", *Int. J. Eng. Sci.* **144** (2019), article no. 103131.
- [28] M. Tasche and H. Zeuner, "Worst and average case roundoff error analysis for FFT", *BIT Numer. Math.* **41** (2001), no. 3, pp. 563–581.
- [29] A. J. W. Duijndam and M. A. Schonewille, "Nonuniform fast Fourier transform", *Geophysics* **64** (1999), no. 2, pp. 539–551.
- [30] L. Greengard and J.-Y. Lee, "Accelerating the nonuniform fast Fourier transform", *SIAM Rev.* **46** (2004), no. 3, pp. 443–454.
- [31] M. Kircheis and D. Potts, "Direct inversion of the nonequispaced fast Fourier transform", *Linear Algebra Appl.* **575** (2019), pp. 106–140.
- [32] A. M. Dykhne, "Conductivity of a two-dimensional two-phase system", *Sov. Phys. JETP* **32** (1971), no. 1, pp. 63–65.
- [33] Y. V. Obnosov, "Periodic Heterogeneous Structures: New Explicit Solutions and Effective Characteristics of Refraction of an Imposed Field", *SIAM J. Appl. Math.* **59** (1999), no. 4, pp. 1267–1287.
- [34] R. V. Craster and Y. V. Obnosov, "Four-Phase Checkerboard Composites", *SIAM J. Appl. Math.* **61** (2001), no. 6, pp. 1839–1856.
- [35] C. Bellis, H. Moulinec and P. Suquet, "Eigendecomposition-based convergence analysis of the Neumann series for laminated composites and discretization error estimation", *Int. J. Numer. Methods Eng.* **121** (2020), no. 2, pp. 201–232.
- [36] C. Dorn and M. Schneider, "Lippmann-Schwinger solvers for the explicit jump discretization for thermal computational homogenization problems", *Int. J. Numer. Methods Eng.* **118** (2019), no. 11, pp. 631–653.
- [37] E. L. Briggs, D. J. Sullivan and J. Bernholc, "Real-space multigrid-based approach to large-scale electronic structure calculations", *Phys. Rev. B* **54** (1996), no. 20, article no. 14362.
- [38] D. W. Kelly, J. P. de S. R. Gago, O. C. Zienkiewicz and I. Babuska, "A posteriori error analysis and adaptive processes in the finite element method: Part I—error analysis", *Int. J. Numer. Methods Eng.* **19** (1983), no. 11, pp. 1593–1619.
- [39] K. S. Djaka, V. Taupin, S. Berbenni and C. Fressengeas, "A numerical spectral approach to solve the dislocation density transport equation", *Model. Simul. Mat. Sci. Eng.* **23** (2015), no. 6, article no. 065008.