



HAL
open science

Numerical simulation of seismic wave propagation in complex geological media

Paul Cupillard

► **To cite this version:**

Paul Cupillard. Numerical simulation of seismic wave propagation in complex geological media. Geophysics [physics.geo-ph]. Université de Lorraine, 2022. tel-03680080

HAL Id: tel-03680080

<https://hal.science/tel-03680080v1>

Submitted on 27 May 2022

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

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



UNIVERSITÉ
DE LORRAINE

SIReNa



Numerical simulation of seismic wave propagation in complex geological media

Paul CUPILLARD

Maître de conférences à l'Université de Lorraine

Mémoire présenté publiquement le 22 mars 2022 pour l'obtention d'une
HABILITATION À DIRIGER DES RECHERCHES
de l'UNIVERSITÉ DE LORRAINE

préparé au laboratoire GeoResources (UMR 7359 UL-CNRS)
dans le cadre de l'École Doctorale SIReNa
mention Géosciences

Rapporteurs :	Heiner IGEL	Professeur à LMU Munich
	Anne PAUL	Directrice de Recherche au CNRS Université Grenoble Alpes
	Johan ROBERTSSON	Professeur à ETH Zürich
Examineurs :	Hélène BARUCQ	Directrice de Recherche à l'INRIA Université de Bordeaux
	Guillaume CAUMON	Professeur à l'Université de Lorraine

Contents

Foreword / Avant-propos	1
Seismic waves and earth models / Ondes sismiques et modèles de terre	1
Models and reality / Modèles et réalité	2
Modeling seismic waves / Modélisation des ondes sismiques	5
1 Research activities	9
Introduction	9
1.1 The spectral element method	10
1.1.1 Basics	11
1.1.2 Contributions	13
1.2 Non-periodic homogenization of the elastic wave equation	15
1.2.1 Basics	15
1.2.2 Contributions	17
1.2.3 Perspective №1: Homogenization of multiscale fractured media	20
1.2.4 Perspective №2: Effective media for ground motion modeling	26
1.3 Inferring the earth interior and seismic source features from full waveforms	29
1.3.1 Basics	29
1.3.2 Contributions	31
1.3.3 Perspective №1: Bayesian inversion of FWI images for reducing structural un- certainties	32
1.3.4 Perspective №2: Inverse homogenization of time-reversal focal spot	34
1.4 Ambient noise correlations	35
1.4.1 Basics	35
1.4.2 Contributions	38
1.4.3 Perspective: Modeling full waveforms of ambient noise correlations	42
References	44
2 Personal records	59
2.1 Curriculum vitae	59
2.1.1 Contact	59
2.1.2 Biodata	59
2.1.3 Appointments	59
2.2 Teaching data	60
2.2.1 Courses	60
2.2.2 Supervision	60
2.2.3 Responsibilities	60
2.3 Research data	60
2.3.1 Indicators	60
2.3.2 Supervision	61
2.3.3 Responsibilities	61
2.3.4 Collaborations	62

2.3.5	Mobility	62
2.3.6	Miscellaneous	63
2.4	List of publications and communications	64
3	Selected articles	69
	Amplitude of surface waves obtained by noise correlation	69
	On the one-bit noise correlation	84
	RegSEM: a regional spectral element code	102
	Homogenization of 3-D complex elastic media	120
	Interpretation of time reversal focal spot based on point-source homogenization	139
	Remerciements	163

Foreword / Avant-propos

Seismic waves and earth models

Like an emotion, a mechanical wave can only occur within a host body. This is a thrill of the latter. Disturbed by an excitation on its surface or in its volume, the body will propagate the disturbance within itself. The wave precisely is this propagated disturbance. The excitation and the host body fully characterize it. The wave equation, established in its simplest form by Jean Le Rond d'Alembert in 1746, translates this complete dependence of the wave u on the excitation f and on the mechanical properties (e.g. the Young modulus E and the density ρ) of the host body:

$$\rho \frac{\partial^2 u}{\partial t^2} = E \frac{\partial^2 u}{\partial x^2} + f. \quad (1)$$

Written in this way, the equation is valid for a one-dimensional homogeneous body. x and t are space and time, respectively. At any point x where the excitation f does not act directly, a solution of the equation is

$$u(x, t) = f(x - ct), \quad (2)$$

where $c = \sqrt{E/\rho}$ is the wave propagation velocity. Thus it is clear that the wave is fully determined by the source excitation and the medium in which it propagates. In underwater acoustics, ultrasonic imaging and seismology, we can therefore use the recording of waves to estimate the source that generated them and the properties of the medium through which they traveled.

Seismic waves will be the main object of interest in the present work. Their analysis makes it possible to build seismic source models and earth models at different scales. Seismic source models describe slip on a fault plane characterized by its geographical position and its orientation. Earth models describe the spatial distribution of mechanical properties within the globe, lithosphere, crust,

Ondes sismiques et modèles de terre

A la manière d'une émotion, une onde mécanique ne peut survenir qu'au sein d'un corps hôte. Il s'agit d'un tressaillement de ce dernier. Perturbé par une excitation à sa surface ou dans son volume, le corps va propager la perturbation en son sein. L'onde est précisément cette perturbation propagée. L'excitation et le corps hôte la caractérisent complètement. L'équation d'onde, établie sous sa forme la plus simple par Jean Le Rond d'Alembert en 1746, traduit cette complète dépendance de l'onde u à l'excitation f et aux propriétés mécaniques (e.g. au module d'Young E et à la masse volumique ρ) du corps hôte :

Ainsi écrite, cette équation vaut pour un corps homogène unidimensionnel. x et t y sont l'espace et le temps, respectivement. En tout point x où l'excitation f ne s'exerce pas directement, une solution de l'équation est

où $c = \sqrt{E/\rho}$ est la vitesse de propagation de l'onde. Ainsi est-il clair que l'onde est totalement déterminée par l'excitation source et par le milieu dans lequel elle se propage. En acoustique sous-marine, en imagerie ultrasonore, en sismologie, on pourra donc utiliser l'enregistrement des ondes pour estimer la source qui les a générées et les propriétés du milieu qu'elles ont traversé.

Ce sont les ondes sismiques qui vont nous intéresser dans cet ouvrage. Leur analyse permet de construire des modèles de sources sismiques et des modèles de terre à différentes échelles. Les modèles de sources sismiques décrivent le glissement sur un plan de faille caractérisé par sa position géographique et son orientation. Les modèles de terre décrivent la distribution spa-

sedimentary basins, etc., depending on the scale of study (figure 1). The properties represented in these models often are the velocities of the seismic waves. The discontinuities of these properties point out the boundaries of geological layers, giving access to the geometric and topological structure of the underground. In addition, the value of the seismic velocities makes it possible to deduce other properties with varying degrees of accuracy, such as pressure, temperature, mineralogical composition, fluid content, etc. Therefore, the earth models provided by the seismological community find applications i) in fundamental earth sciences for understanding the structure and dynamics of our planet at different scales and ii) in engineering for the exploration and exploitation of underground resources (hydrocarbons, ores, geothermal energy, etc.), geological storage (CO₂, hydrogen, etc.) or the monitoring of damage to structures (tunnels, mines, etc.).

Like any material object, our planet is made of atoms. However, geologists do not always aim at describing the Earth by its elementary constituents. They can indeed consider minerals, rocks, geological units, major provinces, global envelopes... all these scales being nested, from the smallest (i.e. the atom) to the largest (i.e. the globe). As a consequence, the question of scale will also concern seismologists. When building models, they will wonder what their spatial resolution is, and how the values obtained at this resolution connect to the values at a smaller scale. Since seismic waves are the data from which the models are built, it is by studying their physics that seismologists will be able to answer these questions. Over the past decade, the development of non-periodic homogenization has shed new light on the interactions between seismic waves and the multiple scales involved in geology; this will be extensively discussed in this work.

Models and reality

I have heard many scientists declare, with the apparent humility of a reflexive conclusion, that models are an approximation of reality. This means that an earth model like PREM (figure 1)

tiale des propriétés mécaniques au sein du globe, de la lithosphère, de la croûte, des bassins sédimentaires, etc., selon l'échelle d'étude (figure 1). Les propriétés représentées dans ces modèles sont très souvent les vitesses des ondes sismiques. Leurs discontinuités indiquent des limites d'unités géologiques et donnent ainsi accès à la structure géométrique et topologique du sous-sol. De plus, la valeur des vitesses sismiques permet de déduire, de manière plus ou moins précise, d'autres propriétés comme la pression, la température, la composition minéralogique, le contenu fluide... Ainsi les modèles de terre fournis par les sismologues trouvent-ils leur utilité i) en sciences de la terre fondamentales pour la compréhension de la structure et de la dynamique de notre planète à différentes échelles et ii) en géo-ingénierie pour l'exploitation de ressources souterraines (hydrocarbures, minerais, énergie géothermique...), le stockage (de CO₂, d'hydrogène...) ou le suivi d'endommagement d'ouvrages (tunnels, mines...).

Comme tout objet matériel, notre planète est constituée d'atomes. Cependant, les géologues ne vont pas toujours chercher à décrire la Terre par ses constituants élémentaires. L'on peut en effet considérer les minéraux, les roches, les unités géologiques, les grandes provinces, les enveloppes globales... toutes ces échelles étant imbriquées, de la plus petite (i.e. l'atome) à la plus grande (i.e. le globe). *De facto*, la question de l'échelle va aussi concerner le sismologue. En produisant ses modèles, il se demandera quelle est leur résolution spatiale, et comment les valeurs obtenues à cette résolution se connectent aux valeurs à plus petite échelle. Les ondes sismiques étant la donnée grâce à laquelle sont construits les modèles, c'est en étudiant la physique de celles-ci que le sismologue pourra répondre à ses questions. Au cours de la dernière décennie, le développement de l'homogénéisation non-périodique a jeté un nouvel éclairage sur les interactions entre les ondes sismiques et les multiples échelles mises en jeu en géologie ; il en sera longuement question dans cet ouvrage.

Modèles et réalité

J'ai entendu de nombreux scientifiques déclarer, avec l'apparente humilité d'une conclusion reflexive, que les modèles sont une approximation de la réalité. Cela signifie qu'un modèle de

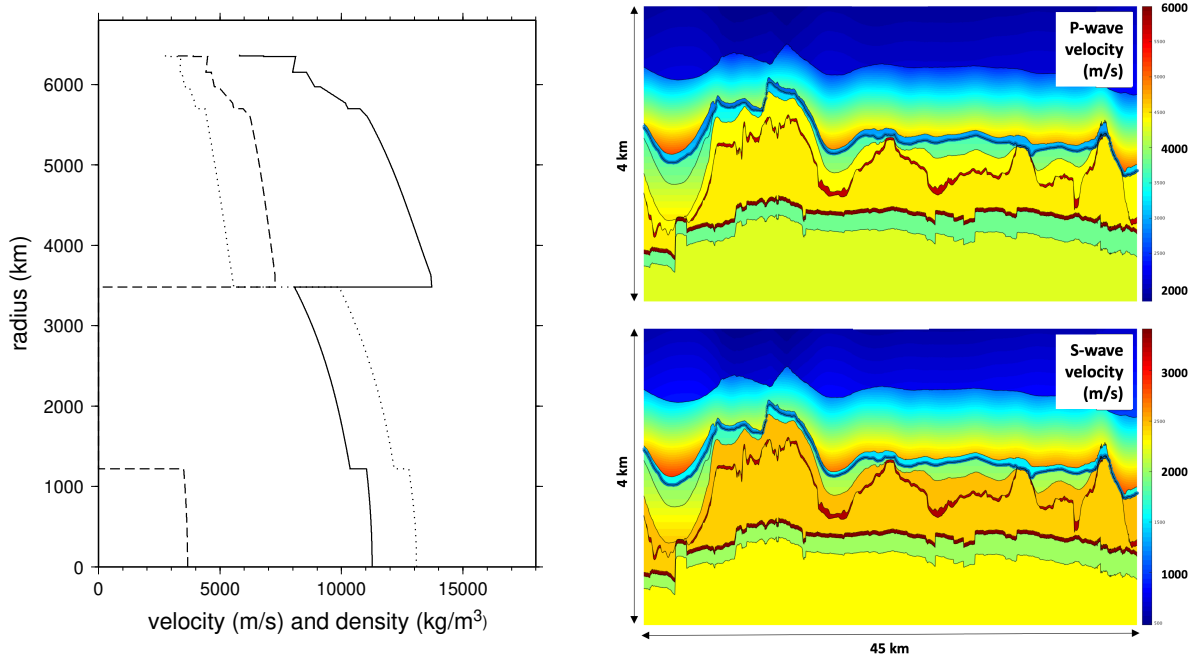


Figure 1: **Left:** An isotropic version of the PREM (Dziewonski, A. and D. Anderson, 1981. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* 25, 297–356). P-wave velocity (plain line), S-wave velocity (dashed line) and density (dotted line) are represented. The water layer in the shallower part of the original model has been removed. **Right:** A section across a subsurface model of the Groningen region in The Netherlands. Vertical exaggeration is $8\times$. Modified from Renat et al., 2018 (Homogenized elastic model of the Groningen gas field. MSc thesis, Université de Lorraine).

would be an approximation of our planet. Isn't it absurd ? PREM, like any earth model, is a mathematical abstraction¹ made of numbers with units. It is not made of atoms, minerals or rocks. How could it be something close to our planet? Is a picture of a car an approximation of a car? Obviously not. A soapbox is, because you can get into it and drive. Would you drive a picture?

Models are a representation of reality. When some claim that they approximate reality, I guess they just intend to point out that models are not perfect, i.e. they are not able to represent the whole complexity of the reality we experience and observe. An earth model, for instance, will probably never consist of a description of the nature and position of each atom that makes up the underground. And even so, we already know that such a model would remain limited, because it would not be able to describe all the physical and biological phenomena observed in geosciences. In particular, certain scaling laws would be lacking in order to account for macroscopic phenomena. Moreover, reality is not only physical and biological. When Paul Elu-

terre comme PREM (figure 1) serait une approximation de notre planète. N'est-ce pas absurde ? PREM, comme tout modèle de terre, est une abstraction mathématique¹. Il est fait de nombres, non d'atomes, de minéraux ou de roches. Comment pourrait-il approximer notre planète ? Le dessin d'une voiture approxime-t-il une voiture ? Bien sûr que non ! Une caisse à savon, oui, car nous pouvons monter à l'intérieur et la conduire. En revanche, conduirions-nous un dessin ?

Les modèles sont des représentations de la réalité. Quand un scientifique déclare qu'ils en sont une approximation, je devine qu'il entend souligner leur caractère limité, en ce sens qu'ils ne peuvent rendre compte de toute la complexité du réel que nous vivons et observons. Un modèle de terre, par exemple, ne consistera probablement jamais en une description de la nature et de la position de chaque atome qui compose le sous-sol. Et quand bien même, nous savons déjà qu'un tel modèle resterait limité, car il ne serait pas en mesure de décrire tous les phénomènes physiques et biologiques observés en géosciences. Manqueraient

¹Backus, G. and F. Gilbert (1968). The resolving power of gross earth data. *Geophys. J. Int.* 16(2), 169–205.

ard writes that "The Earth is blue like an orange", does he not express the reality of the emotion we can feel when thinking of the spherical fullness and the bluish softness of our planet?

In short, a model is the representation of one or more particular objects or phenomena, from a particular angle: that of science. A fundamental question then is that of the quality of the model: within the strict limits of what it is supposed to represent, how good is it? To continue with the example of an earth model, the latter will be all the better if it satisfies a large number and different types of observations (geophysical, geological, logs, etc.). In doing so, it will be all the more useful and illuminating². *Useful* refers to the ability of a model to impact concrete reality. A useful model can in particular give rise to a technology which will modify the concrete conditions of our existence. An *illuminating* model will play on the abstract reality of objects and phenomena, that is to say on the idea that we have of them. Indeed, the new light suddenly thrown by a model on a thing (object or phenomenon) instantly modifies our conception of this thing. Therefore, models are both tools and materials for the production of reality. As fruits of the "collaboration between the outside world and our spirit"³, they participate in the incessant construction and deconstruction of reality.

This vision of reality, changing and at human level, is opposed to that implied in the statement that we criticized above and which asserts that a model is an approximation of reality. In this statement, the idea of a transcendent and authentic entity indeed emerges. Such an *undermining view*⁴ inherits the tradition initiated by Plato in *Phaedo*, the *Republic*, the *Symposium* and *Timaeus*. In these works, the philosopher distinguishes a superior and authentic reality (the world of the *eidōs*, i.e. the concepts, ideas and abstractions) and a sensible reality which would be only the *idea* (i.e. the manifestation, or appearance) of the primary one. While it is true that philosophers and scientists produce abstractions, it does not seem to me necessary to assert that they belong to or tend towards a superior reality which would govern a lower one. Our products are characterized, as we have already said, by the ability to explain some observations.

notamment certaines lois de mise à l'échelle pour parvenir à rendre compte de phénomènes macroscopiques. De plus, la réalité n'est pas uniquement physique et biologique. Quand Eluard écrit que "La Terre est bleue comme une orange", n'exprime-t-il pas la réalité de l'émotion que nous pouvons ressentir à l'idée de la plénitude sphérique et de la douceur bleutée de notre planète ?

En somme, un modèle est la représentation d'un ou plusieurs objets ou phénomènes particuliers, sous un angle particulier : celui de la science. Une question fondamentale est alors celle de la qualité du modèle : dans les strictes limites de ce qu'il est censé représenter, dans quelle mesure celui-ci est-il bon ? Pour poursuivre avec l'exemple d'un modèle de terre, ce dernier sera d'autant meilleur qu'il satisfera à un grand nombre et à différents types d'observations (géophysiques, géologiques, diagraphiques, etc.). Ce faisant, il sera d'autant plus utile et éclairant². *Utile* renvoie à la capacité d'un modèle à impacter la réalité concrète. Un modèle utile peut notamment donner lieu à une technologie qui modifiera les conditions concrètes de nos existences. Un modèle *éclairant* jouera quant à lui sur la réalité abstraite des objets et des phénomènes, c'est-à-dire sur l'idée que nous nous en faisons. En effet, la lumière nouvelle jetée soudain par un modèle sur une chose (objet ou phénomène) modifie instantanément notre conception de cette chose. Ainsi les modèles sont-ils à la fois des outils et des matériaux de la fabrication du réel. Fruits de la "collaboration entre le monde extérieur et notre esprit"³, ils participent à l'incessante construction et déconstruction de la réalité.

Cette vision du réel, mouvante et à hauteur d'homme, s'oppose à celle sous-entendue dans le poncif que nous avons critiqué plus haut et qui affirme qu'un modèle est une approximation de la réalité. Dans cet énoncé sourd en effet l'idée d'une entité transcendante et authentique. Cette *undermining view*⁴ hérite de la tradition initiée par Platon dans le *Phédon*, la *République*, le *Banquet* et le *Timée*, ouvrages dans lesquels le philosophe distingue une réalité supérieure et authentique (le monde des *eidōs*, c'est-à-dire des concepts, idées ou abstractions) et une réalité sensible qui n'en

²Box, G. (1979). Robustness in the strategy of scientific model building. In R. L. Launer and G. N. Wilkinson (Eds.), *Robustness in Statistics*, pp. 201–236. Academic Press.

³Morin, Edgar. "Science et Philosophie", interview by D. Bougnoux et B. Engelbach. *Nonfiction*, April 10, 2008.

⁴Harman, G. (2018). *Object-oriented ontology: A new theory of everything*. Penguin UK.

One can marvel at the remarkable explanatory or even predictive power of some of these products, but what need is there to give their abstract nature a primary character? I often marvel at the central limit theorem, but it is not necessary for me to grant the Gaussian a superiority over the multiple phenomena it accounts for. Does it precede them? Is it its essence, or the outline of its essence? No one knows, because the very idea of an essence that would precede existence is only speculation. To the Platonic tradition, I therefore prefer the parsimonious wisdom of Edgar Morin. But the notion of reality being able to feed infinite debates, I will stop there my considerations on the subject and go back to seismology.

Modeling seismic waves

We have seen that earth models and seismic sources models are deduced from the waves that can be recorded on the surface of the globe or sometimes even at depth. Whatever the method used for this, it will be necessary, at a recurring stage of the process, to model these waves, i.e. to solve a wave propagation equation for a given source and a given medium. In some very simple cases (e.g. equation 1), analytical solutions exist (e.g. equation 2), but these cases rarely correspond to the geological complexity that we want to take into account. Geological environments are indeed heterogeneous and 3D. Moreover, they can present behaviors much more complex than the simple linear elasticity involved in equation (1). For lack of analytical solutions in such media, the resolution of the wave equation will be done numerically, using simulation methods.

In addition to their implication in the construction of models, the numerical methods for modeling seismic wave propagation are essential to the evaluation of the seismic risk, in particular the site effects which can amplify the energy of the waves on the surface and cause significant damage. These methods are also involved in seismic survey design in order to monitor or better image a given geolog-

serait que le produit ou l'apparence (*idea*). Or s'il est vrai que philosophes et scientifiques produisent des abstractions, il ne me semble pas nécessaire de les prétendre appartenir à ou tendre vers une réalité supérieure qui présiderait à une réalité inférieure. Nos productions se caractérisent, comme nous l'avons déjà dit, par une aptitude à expliquer des observations. L'on peut s'émerveiller du remarquable pouvoir explicatif voire prédictif de certaines de ces productions, mais quel besoin y a-t-il de conférer à leur nature abstraite un caractère premier ? Je m'émerveille souvent du théorème central limite, mais il ne m'est pas nécessaire d'accorder à la gaussienne une supériorité sur les multiples phénomènes dont elle rend compte. Les précède-t-elle ? En est-elle l'essence, ou l'ébauche d'une essence ? Nul ne sait, car l'idée même d'une essence qui précéderait l'existence n'est que spéculation. A la tradition platonicienne je préfère donc la sagesse parcimonieuse d'Edgar Morin. Mais la notion de réalité pouvant nourrir d'infinis débats, j'arrête là mes considérations sur le sujet pour en revenir à la sismologie.

Modélisation des ondes sismiques

Nous avons vu que les modèles de terre et de sources sismiques se déduisent des ondes que l'on peut enregistrer à la surface du globe ou parfois même en profondeur. Quelle que soit la méthode employée pour cela, il sera nécessaire, à une étape récurrente du procédé, de modéliser ces ondes, c'est-à-dire de résoudre une équation de propagation pour une source et un milieu donnés. Dans certains cas très simples (e.g. équation 1), la résolution peut se faire analytiquement (e.g. équation 2), mais ces cas correspondent rarement à la complexité géologique que l'on désire prendre en compte. Les milieux géologiques sont en effet hétérogènes et 3D. De plus, ils peuvent présenter des comportements bien plus complexes que la simple élasticité linéaire mise en jeu dans l'équation (1). Faute de solutions analytiques dans de tels milieux, la résolution de l'équation d'onde se fera numériquement, via des méthodes de simulation.

Outre leur implication dans la construction de modèles, les méthodes de modélisation numérique de la propagation des ondes sismiques sont indispensables à l'évaluation du risque sismique, notamment des effets de site qui peuvent amplifier l'énergie des ondes en surface et causer

ical medium.

Whatever the application, most numerical modeling methods require a mesh, i.e. a discretization of the geometry of the medium. However, we have seen that geological environments are complex by nature: they are heterogeneous, 3D, and above all they are discontinuous and multiscale. These last two features pose major issues when modeling waves numerically, because the required mesh can be very difficult to generate and the computational cost of the simulation can increase dramatically (figure 2). In many applications, these two features are set aside. This is the case in full waveform inversion (FWI) for instance. To remain above the seismic resolution and mitigate the non-uniqueness of the problem, FWI results are restricted to smooth models without any small scales. However, geological environments contain some, and geological complexity should remain one of the guide of seismological research.

Forgetting the complexity of geology, some seismologists say that the resolution of the seismic wave equation no longer is a problem. When claiming this, I guess they also set aside the ambition to process the ever-increasing quantity of data at our disposal. Despite the development of artificial intelligence algorithms and the increasing computational power, these data would benefit from innovative simulation methods in order to investigate them deeper and extract further information on geological environments and seismic sources. There are more and more seismic data because of i) new acquisition technologies such as distributed acoustic sensing and ii) new types of data such as rotational motion of particles, tremors from non-volcanic sources, and noise. The correlation of the latter between two seismic stations indeed gives information on the geological environment between the stations and therefore leads to a huge amount of data. Better using and understanding these data, with the help of better simulation methods and tools, is one of our missions.

d'importants dégâts. Ces méthodes interviennent également dans le design d'acquisitions sismiques en vue de surveiller ou de mieux imager un milieu géologique donné.

Quelle qu'en soit l'application, la plupart des méthodes de modélisation numérique nécessitent un maillage, c'est-à-dire une discrétisation de la géométrie du milieu. Or nous avons vu que les milieux géologiques étaient complexes par nature : ils sont hétérogènes, 3D, mais surtout discontinus et multi-échelles. Ce sont ces deux derniers traits qui posent les plus gros problèmes, car alors le maillage peut être très difficile à générer, et le coût de calcul de la propagation exorbitant (figure 2). Dans nombre d'applications, ces deux traits sont mis de côté. C'est le cas par exemple dans l'inversion de formes d'onde complètes où, pour demeurer au-dessus de la résolution sismique et ne pas aggraver la non-unicité du problème, on propose des modèles lisses, sans petites échelles. Pourtant, la géologie en contient, et la complexité de celle-ci me semble devoir rester un des phares de la recherche sismologique.

C'est peut-être parce qu'ils ont perdu de vue la complexité géologique que certains sismologues affirment que la résolution de l'équation des ondes sismiques est un problème résolu. Faut-il aussi qu'ils n'ambitionnent pas de traiter à fond l'énorme quantité de données dont nous disposons et qui recèle encore, faute de méthodes de simulation suffisamment performantes, bien des informations sur les milieux géologiques et les sources sismiques. Les données sismiques sont en effet très nombreuses, en raison notamment du développement de nouvelles technologies d'acquisition comme les systèmes de détection acoustique distribuée, mais aussi très variées. Elles ne concernent plus seulement les mouvements de translation des particules, mais aussi leur rotation. De plus, elle ne mettent plus seulement en jeu les séismes provoqués par une rupture soudaine de la croûte terrestre, mais aussi les tremors de tout type ainsi que le bruit sismique. La corrélation du bruit entre deux stations fournit en effet des informations sur le milieu géologique entre les stations. Ainsi s'ouvre un immense jeu de données. Mieux le comprendre et l'utiliser, en s'aidant notamment d'une meilleure simulation des signaux, est l'une de nos missions.

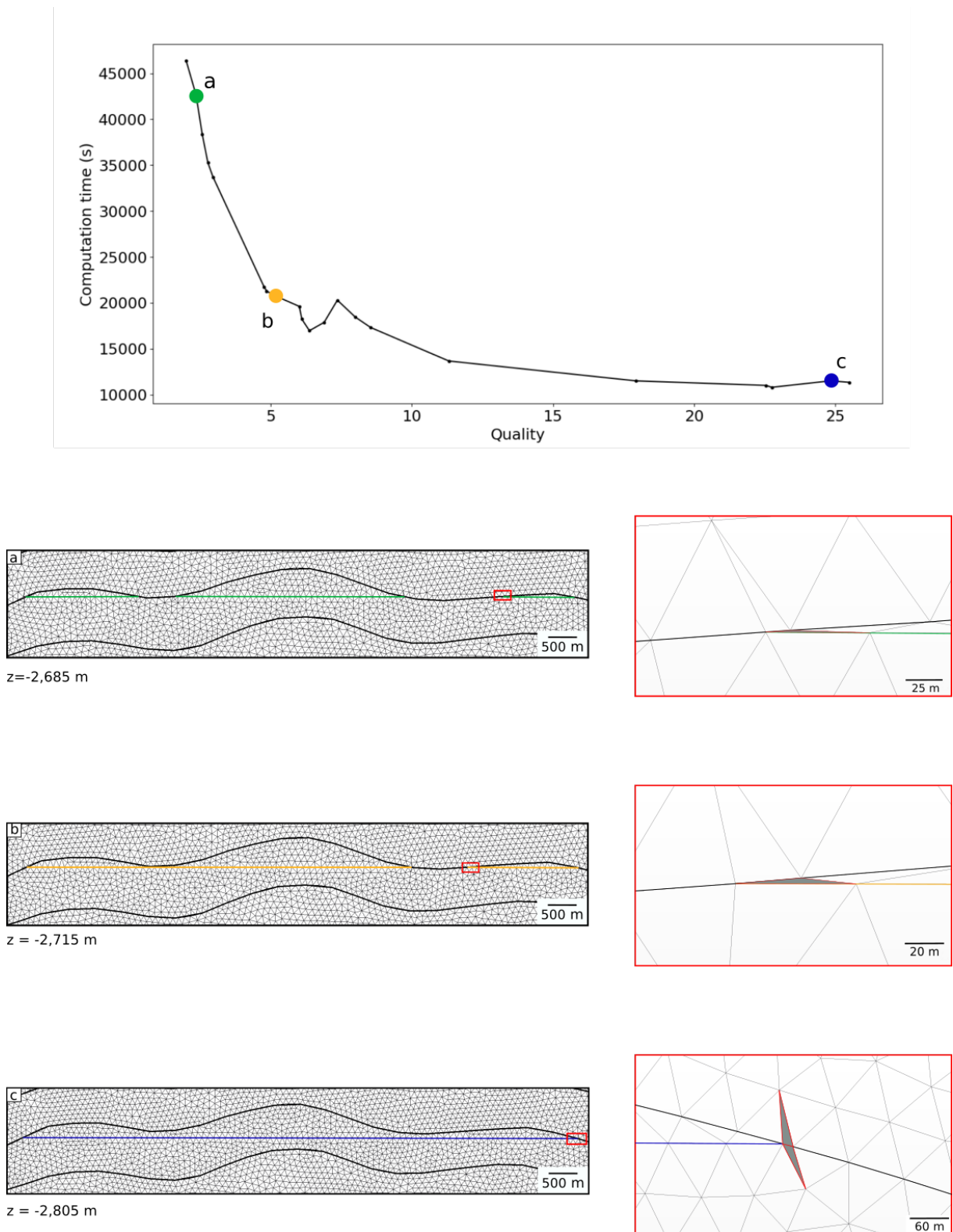


Figure 2: Mesh quality (smallest triangle height) versus computation time (using a discontinuous Galerkin method). In this example, a gas-water contact is inserted in a 2D subsurface model. 20 depths are considered, leading to 20 meshes. In some cases, the insertion of the contact leads to small and elongated elements which make the wave simulation very long (a). In some other cases, the quality of the elements is good enough so that the computation time is not so large (c). From Legentil et al., 2022 (Testing scenarios on geological models: Local interface insertion in a 2d mesh and its impact on seismic wave simulation. *Computers & Geosciences* 159, 105013).

Chapter 1

Research activities

Introduction

I started my PhD on January 3, 2005. My advisor, Yann Capdeville, suggested that I should work on the implementation of the spectral element method in 3D for setting up an adjoint-based full waveform inversion. "What the hell is he talking about?" I thought. Of course I had studied seismology before; I had learned wave physics and followed courses on inverse problem; I had completed a Master project on receiver functions which resulted in a paper (Ritsema et al., 2009); but "What is this adjoint stuff and the spectral elements? That sounds scary!". To defeat my ignorance and calm my anxiety down, I read Komatitsch (1997); Komatitsch and Vilotte (1998); Komatitsch and Tromp (1999); Chaljub (2000) and progressively understood what the spectral element method (SEM) was and why it enabled the efficient computation of seismic wave propagation in heterogeneous media (subsection 1.1.1). First, the SEM benefits from the ability of the finite element like methods to naturally handle complex interface geometry. Second, its convergence is spectral, which makes it very accurate. Third, it involves a diagonal mass matrix, which makes it computationally tractable. As a result, the SEM has been proved to overcome several limitations encountered in other numerical methods. A 2D code and a preliminary 3D code written by Gaetano Festa and Elise Delavaud also helped me in understanding it. I enriched the 3D version with i) a spherical deformation of the computational domain, ii) an automatic meshing of the solid earth surface and Moho topography along with PREM (Dziewonski and Anderson, 1981) or other spherical interfaces, iii) anisotropy and attenuation, iv) the ability of running on parallel architectures (subsection 1.1.2). That is how the structured version of RegSEM (Cupillard et al., 2012) was born.

As for the adjoint, 2005 was the year that Jeroen Tromp et al. (2005) and Yann Capdeville (2005) brought back to light the idea introduced by Lailly (1983) and Tarantola (1984, 1988) twenty years earlier. Reading these papers, I understood that the adjoint technique was 'just' an efficient way of computing the gradient of a misfit functional (section 1.3). Exploration seismologists were used to implement it for inverting P-waves under the 2D acoustic approximation, relying on finite difference algorithms to solve the forward problem (e.g. Bunks et al., 1995; Pratt, 1999; Dessa et al., 2004; Ravaut et al., 2004). At the global scale, the normal mode summation technique (e.g. Gilbert, 1971) along with a coupling theory (Woodhouse and Dziewonski, 1984; Li and Romanowicz, 1995) were able to provide synthetic seismograms in elastic models of the 3D earth (e.g. Li and Romanowicz, 1996), but only weak lateral heterogeneities and 2D sensitivity kernels could be considered at best in this framework. Thanks to the SEM, the accurate modeling of full elastic wavefields in 3D models with strong heterogeneities of any shape became feasible, so it was time to consider it within an adjoint-based inverse problem to image the earth interior from earthquake data. The first inversion of traveltimes has been applied to the southern California crust (Tape et al., 2009); the first waveform inversion has been applied to the mantle below the Australian continent (Fichtner et al., 2009).

Although I implemented the computation of adjoint kernels in RegSEM during my post-doc at UC Berkeley (2008-2010), I have never carried out an inversion myself. One of the reason is that the 2000s offered a new, very exciting type of seismological data in which I got interested: ambient noise corre-

lations (Shapiro and Campillo, 2004). Such data opened the path to seismic tomography with no earthquake (e.g. Shapiro et al., 2005; Yao et al., 2006; Lin et al., 2007; Yang et al., 2007). In the context of full waveform modeling and inversion mentioned in the previous paragraphs, my question was "Can noise correlation waveforms be modeled and inverted?". This question actually asked for what the amplitude of noise correlations contain. Whereas the phase was proved, under certain conditions, to correspond to the one of the Green's function, the behavior of the amplitude was much less clear, so I decided to investigate it (subsection 1.4.2). This led to my first significant contributions to seismology (Cupillard and Capdeville, 2010; Cupillard et al., 2011). Along these works, I proposed a method to model the noise correlations waveforms between a given central station and some other stations of a network (Cupillard, 2008). The method is based on building a virtual source at the central station to generate an anisotropic radiation which corresponds to the noise flux (subsection 1.4.3). Since 2008, Curtis and Halliday (2010); Tromp et al. (2010); Fichtner (2014, 2015); Fichtner et al. (2016); Sager et al. (2018) have made good progress in better understanding and accounting for the influence of noise source distribution on noise correlation amplitudes, which asks for an update of the method I proposed fourteen years ago.

At the present day, the propagation of seismic waves can be accurately modeled in 3D heterogeneous media, using either the SEM or another numerical technique (see Virieux et al., 2011; Igel, 2017, for a review of the various numerical methods available to model full waveforms). A decade ago, an important step toward this goal was achieved thanks to the development of the non-periodic homogenization method (Capdeville et al., 2010a,b; Guillot et al., 2010). Such a method enables to compute proper effective properties for the seismic wave equation, allowing the correct account for the effect of small heterogeneities on wave propagation (subsection 1.2.1). By small heterogeneities I mean structures which are smaller than the minimum seismic wavelength propagating in the medium. Handling them explicitly in numerical simulations can be extremely challenging because they control the spatial sampling of the medium and, consequently, the time sampling too (Courant et al., 1928). Small heterogeneities therefore induce massive, possibly prohibitive, computation costs, so working with effective properties is much preferable. Getting back to France after my post-doc at UC Berkeley, I worked with Yann Capdeville again to develop the 3D finite element version of the non-periodic homogenization (subsection 1.2.2). This led to a research paper (Cupillard and Capdeville, 2018) and multiple extended abstracts (Cupillard et al., 2015; Cupillard and Capdeville, 2017; Cupillard et al., 2020; Ibourichène et al., 2021) as well as an on-going collaboration with Yann on the homogenization method (Capdeville et al., 2015, 2020; Renat et al., 2022). Yann's works actually have a major influence on my current research, because I believe that homogenization is an amazing tool to describe the relationship between the waves and the various scales in which they stand or propagate. Therefore, I am pleased to put efforts in investigating how it can help in easing the forward modeling (subsections 1.2.3 and 1.2.4) as well as in downscaling full waveform inversion results (subsection 1.3.3) or source images (subsection 1.3.4). Contributing to these efforts, my current research team (RING, GeoRessources Lab, Université de Lorraine/CNRS) has been developing meshing tools (Merland et al., 2014; Pellerin et al., 2014; Botella et al., 2016; Pellerin et al., 2017; Anquez et al., 2019; Legentil et al., 2022) and structural modeling approaches (e.g. Caumon et al., 2009; Cherpeau et al., 2010; Bonneau et al., 2013; Julio et al., 2015; Ruiu et al., 2016; Godefroy et al., 2019; Irakarama et al., 2021) which I greatly benefit from.

In this chapter, I give some more details on the four topics I just introduced (i.e. SEM, non-periodic homogenization, waveform-based imaging and noise correlation) along with a summation of my contributions to each of them. Most of all, I draw some perspectives based on my ongoing research. These perspectives will guide my work as well as the students I will supervise in the next decades.

1.1 The spectral element method

Numerous techniques, like ray tracing or normal mode summation, have been developed for decades to compute synthetic seismograms. For many purposes, ray tracing is very convenient, but it relies on a high frequency approximation (e.g. Cervený, 2001), which means that it is only valid when seismic wavelength is much smaller than the scale of heterogeneity. The normal mode summation technique

is able to model low-frequency waves in 3D earth models (e.g. Li and Romanowicz, 1996), but the computation cost of such a technique is quickly prohibitive as the number of modes to couple increases with frequency. Moreover, it is limited to weak lateral heterogeneities.

To overcome these limitations, numerical solutions have been investigated. Finite difference schemes have been developed (e.g. Alterman and Karal, 1968; Boore, 1972; Kelly et al., 1976; Virieux, 1984, 1986; Igel et al., 1995; Robertsson et al., 1994; Moczo et al., 2007) but they present intrinsic problems in dealing with strong and deformed interfaces like basin edges, the Moho, the free surface and solid-fluid discontinuities. Such limitations do not exist in finite element methods, but the low polynomial order classically used in this kind of approaches make them inaccurate and dispersive when applied to elastodynamic problems (Lysmer and Drake, 1972; Backer, 1976; Marfurt, 1984; Toshinawa and Ohmachi, 1992). Since the 1990s, efforts have been focused on developing higher-order numerical modeling of seismic wave propagation. An important result from these efforts is the discontinuous Galerkin method (e.g. Dumbser and Käser, 2006). Another major result is the spectral element method (SEM). Initially introduced in fluid mechanics (Patera, 1984; Maday and Patera, 1989), this method has been successfully applied to elastodynamics with the increasing concern of developing numerical techniques ensuring both a great precision and a reasonable numerical cost (Seriani and Priolo, 1994; Faccioli et al., 1997; Komatitsch and Vilotte, 1998; Seriani, 1998; Komatitsch and Tromp, 1999).

1.1.1 Basics

We here outline the application of the SEM to elastodynamics. In particular, we point out the three features that makes it accurate and computationally tractable.

Strong formulation of the elastodynamic problem

In his strong form, the elastodynamic problem is expressed using Newton's second law of motion

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{f} = \rho \ddot{\mathbf{u}} \quad (1.1)$$

and Hooke's law

$$\boldsymbol{\sigma} = \mathbf{C} : \nabla \mathbf{u}. \quad (1.2)$$

In these equations, \mathbf{u} is the displacement field, $\boldsymbol{\sigma}$ is the stress field, \mathbf{f} is the body force, ρ is the density and \mathbf{C} is the elasticity tensor. Applying these equations to a finite volume Ω surrounded by a boundary Γ , given a Dirichlet condition $\mathbf{u} = \mathbf{g}$ on Γ_D and a Neumann condition $\boldsymbol{\sigma}_i n_j = h_i$ on Γ_N such that $\Gamma = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$, the solution \mathbf{u} of the problem exists and it is unique (up to a rigid motion).

Weak formulation of the elastodynamic problem

The SEM is a finite-element method (FEM), so it relies on the weak form of the problem to be solved. To write this form properly in the elastodynamic case, let us first define the functional spaces

$$V = \{ \mathbf{u} : (\Omega, T) \rightarrow \mathbb{R}^3 \mid \forall t \in T, \mathbf{u}_i(\cdot, t) \in H^1(\Omega), \mathbf{u}_i(\cdot, t) = \mathbf{g}_i(\cdot, t) \text{ on } \Gamma_D \} \quad (1.3)$$

and

$$V_0 = \{ \mathbf{u} : (\Omega, T) \rightarrow \mathbb{R}^3 \mid \forall t \in T, \mathbf{u}_i(\cdot, t) \in H^1(\Omega), \mathbf{u}_i(\cdot, t) = 0 \text{ on } \Gamma_D \}, \quad (1.4)$$

where T is the time range and $H^1(\Omega)$ is the order-1 Sobolev space on Ω , i.e. the set of functions for which the squares of both the function itself and its partial derivatives are Lebesgue integrable on Ω . Using such definitions, the weak-form of the elastodynamic problem can be written as

$$\text{Find } \mathbf{u} \in V \text{ such that} \quad (1.5)$$

$$\int_{\Omega} \mathbf{v} \cdot [\nabla \cdot (\mathbf{C} : \nabla \mathbf{u})] \, d\mathbf{x} + \int_{\Omega} \mathbf{v} \cdot \mathbf{f} \, d\mathbf{x} = \int_{\Omega} \mathbf{v} \cdot \rho \ddot{\mathbf{u}} \, d\mathbf{x}, \quad \forall \mathbf{v} \in V_0, \quad \forall t \in T.$$

Applying the divergence theorem to the first integral, the Neumann condition appears explicitly and our weak-form problem becomes

$$\text{Find } \mathbf{u} \in V \text{ such that} \\ \underbrace{\int_{\Omega} \nabla \mathbf{v} : \mathbf{C} : \nabla \mathbf{u} \, d\mathbf{x}}_{a(\mathbf{v}, \mathbf{u})} + \underbrace{\int_{\Omega} \mathbf{v} \cdot \rho \ddot{\mathbf{u}} \, d\mathbf{x}}_{(\mathbf{v}, \rho \ddot{\mathbf{u}})} = \underbrace{\int_{\Omega} \mathbf{v} \cdot \mathbf{f} \, d\mathbf{x}}_{(\mathbf{v}, \mathbf{f})} + \underbrace{\int_{\Gamma_N} \mathbf{v} \cdot \mathbf{h} \, d\Gamma}_{(\mathbf{v}, \mathbf{h})_{\Gamma}}, \quad \forall \mathbf{v} \in V_0, \forall t \in T. \quad (1.7)$$

Galerkin approximation

V is an infinite-dimensional space. To solve equation (1.7) numerically, we discretize this space, i.e. we look for an approximated solution \mathbf{u}^d in a subspace of V , called V^d , which is of finite dimension. We therefore rewrite equation (1.7) as

$$\text{Find } \mathbf{u}^d \in V^d \text{ such that} \\ a(\mathbf{v}^d, \mathbf{u}^d) + (\mathbf{v}^d, \rho \ddot{\mathbf{u}}^d) = (\mathbf{v}^d, \mathbf{f}) + (\mathbf{v}^d, \mathbf{h})_{\Gamma}, \quad \forall \mathbf{v}^d \in V_0^d, \forall t \in T. \quad (1.9)$$

In this last equation, we use the notations introduced underneath the underbraces in equation (1.7).

The question then arises to find a relevant set of basis functions for V^d . Within all the FEMs, piecewise polynomial functions are used, the pieces being the elements of the geometric discretization of the domain Ω .

SEM features

In the SEM, the geometric discretization relies on hexahedral elements in order to benefit from advantageous properties of tensorization. Although hexahedra are less favorable than tetrahedra for meshing geometrically complex structures, a certain flexibility is ensured by a local geometrical transformation from a reference element (i.e. the unit cube $[-1; 1]^3$) to any deformed element. Based on such a geometric discretization, the three SEM features are the following:

- The basis functions are built from the tensorization of degree- n Lagrange polynomials associated to $n+1$ interpolation nodes in $[-1; 1]$.
- The interpolation nodes are chosen to be the Gauss-Lobatto-Legendre (GLL) points.
- The integrals involved in our problem are evaluated using a GLL quadrature. As a consequence, the interpolations nodes and the quadrature nodes are the same.

Inserting the polynomial interpolation and quadrature rules into the weak form (1.9) leads to an algebraic system of equations which governs the displacement and its second-order time derivative at the nodal positions for any instant t . This system can be written as

$$\mathbb{K} \mathbf{U}(t) + \mathbb{M} \ddot{\mathbf{U}}(t) = \mathbf{F}(t), \quad (1.10)$$

where \mathbf{U} is a vector which contains the components of the displacement at the nodes, \mathbb{K} is the stiffness matrix, \mathbb{M} is the mass matrix, and \mathbf{F} is a vector which contains the contributions from both the body force and the boundary conditions.

SEM properties

The choice of the basis functions as tensorized Lagrange polynomials and quadrature nodes collocated with the interpolation nodes leads to the first interesting property of the SEM: the mass matrix is diagonal. Such a property makes the computation of \mathbb{M}^{-1} straightforward, which enables to discretize the time evolution of our problem using an explicit time-stepping. In most of the SEM implementations, this is ensured by a second-order finite-difference scheme.

The choice of a Lagrangian interpolation associated with GLL points provides the second interesting property of the SEM: the convergence is exponential (i.e. an increase of the polynomial degree leads to an exponential decrease of the aliasing error). This property, called spectral precision, gives its name to the method. It enables to use high-degree polynomials and to get a better accuracy as compared to more classical FEMs.

1.1.2 Contributions

Regional scale implementation

The first applications of the SEM have dealt with the global scale on the one hand (Chaljub, 2000; Komatitsch and Tromp, 2002a,b; Komatitsch et al., 2002; Capdeville et al., 2003; Chaljub et al., 2003; Chaljub and Valette, 2004) and ground motion simulations on the other hand (Komatitsch et al., 2004; Delavaud et al., 2006; Lee et al., 2008; Stupazzini et al., 2009; Chaljub et al., 2010). One of my contributions has been to implement the method at the regional scale, which has given rise to RegSEM (Cupillard et al., 2012). This code actually has two versions: a continental-scale version which is able to generate regular meshes of crustal and mantle structures in spherical geometry, and a local-scale version which uses an external mesh generator, CUBIT (<http://cubit.sandia.gov>), to produce 3D unstructured meshes essentially designed to study the seismic response of sedimentary basins. My effort focused on the continental-scale version (Cupillard, 2008); Elise Delavaud (2007) was in charge of the local-scale version.

The continental-scale version of RegSEM can provide a regular mesh of any chunk of the earth whose lateral size is smaller than 90° . To do so, the code uses the so-called cubed sphere mapping (Ronchi et al., 1996). For each element, this mapping allows defining the Cartesian coordinates of 27 control points. Using the Lagrange polynomials of degree 2 associated with these control points, the unit cube can be deformed and the shape and position of each element in the chunk can be defined. Moreover, spherical discontinuities within the velocity model can be honored in a versatile way: one just needs to introduce the radius of each discontinuity, and then the code fills the seismic layers with the appropriate number of elements. Of course, this number depends on the vertical size of the elements. This size is first equal to the horizontal size introduced by the user, and then it is adjusted in each seismic layer to fit the thickness of the layer. In the case of PREM (Dziewonski and Anderson, 1981), there is one more level of sophistication because not only the thickness of the layers but also the seismic velocities in the layers are used by the code to constrain and optimize the vertical size of the elements.

RegSEM can also mesh any surface and Moho topography. This is an important feature because the crust has significant effects on surface waves (Montagner and Tanimoto, 1991; Curtis et al., 1998; Shapiro and Ritzwoller, 2002; Marone and Romanowicz, 2007; Capdeville and Marigo, 2008; Ferreira et al., 2010), even at relatively long period (up to about 60 s). The capability to consider any model with a realistic crust is therefore a major benefit. To do so, the code uses only one layer of elements in the crust. This means that discontinuities within the crust, such as the sediment-rock interface and the upper-lower crust interface, cannot be taken into account. Moreover, the fact that only one layer of elements is used to mesh the crust limits the frequency content that the simulation can handle. For example, when performing a simulation in Tibet (where the crust is more than 70 km thick) with degree-8 polynomials, the highest frequency to be propagated will be approximately 0.1 Hz. A simulation in a 3D crustal model of Europe is shown in figure 1.1.

To avoid artificial reflections at the border of the chunk, absorbing boundary conditions have to be implemented. RegSEM uses the velocity-stress formulation of the so-called Perfectly Matched Layers (PML, Festa and Vilotte, 2005). This formulation requires an unphysical splitting of the field variables along the directions of normal and parallel derivatives with respect to the interface between the PML and the 'normal' elements. This means that in practice, the splitting directions have to be known at every GLL node belonging to the PML, which is not obvious when working with Cartesian coordinates in a deformed layer. Therefore, we here make an assumption: for all the GLL nodes of a given element, the splitting directions defined at the center of the element are used. The effect of such an assumption on the stability

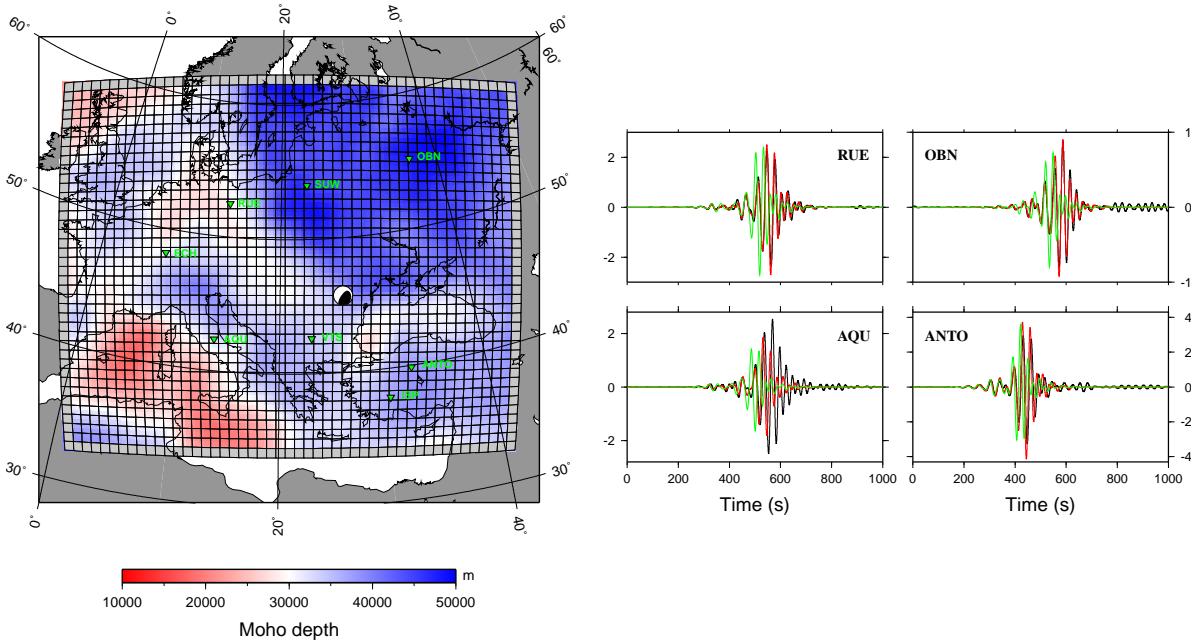


Figure 1.1: Left: Top view of a regular hexahedral mesh of the European region. PML elements are colored in light grey. The other elements show the depth of the Moho which is honored by the mesh. To model the propagation of seismic waves, an earthquake is introduced in Romania. Stations are marked in green. Right: Rayleigh waves obtained in the 0.02 - 0.05 Hz frequency range. Real data are shown in black; results of a numerical simulation in a 3D model (CUB, Shapiro and Ritzwoller, 2002) are in red; results in a 1D model (PREM, Dziewonski and Anderson, 1981) are in green.

and accuracy of the PML is not clear. Furthermore, our PML are isotropic, so spurious reflections can be created when considering anisotropic media. To mitigate them, RegSEM takes advantage of the SEM ability to use a different polynomial degree in each direction: in the PML, a higher polynomial degree is introduced in the damping directions.

Solving the forward problem in waveform inversion

RegSEM has been used for various applications. Among those I have been directly involved in, two aim at better understanding noise correlations (Stehly et al., 2011; Saade et al., 2015), one deals with the numerical implementation of time reversal mirrors (Masson et al., 2014), and two aim at inverting seismic waveforms for imaging the lithosphere and the asthenosphere at a continental scale (Yuan et al., 2014; Zunino et al., 2016). I here provide the abstracts of these two latter studies.

Yuan et al. (2014), Lithospheric expression of geological units in central and eastern North America from full waveform tomography, in *Earth and Planetary Science Letters* 402.SI, pp. 176–186: *The Earth-Scope TA deployment has provided dense array coverage throughout the continental US and with it, the opportunity for high resolution 3D seismic velocity imaging of both lithosphere and asthenosphere in the continent. Building upon our previous long-period waveform tomographic modeling in North America, we present a higher resolution 3D isotropic and radially anisotropic shear wave velocity model of the North American lithospheric mantle, constructed tomographically using the spectral element method for wavefield computations and waveform data down to 40 s period. The new model exhibits pronounced spatial correlation between lateral variations in seismic velocity and anisotropy and major tectonic units as defined from surface geology. In the center of the continent, the North American craton exhibits uniformly thick lithosphere down to 200–250 km, while major tectonic sutures of Proterozoic age visible in the surface geology extend down to 100–150 km as relatively narrow zones of distinct radial anisotropy, with $V_{SV} > V_{SH}$. Notably, the upper mantle low velocity zone is present everywhere under the craton between 200 and 300 km depth. East of the continental rift margin, the lithosphere is broken up into*

a series of large, somewhat thinner (150 km) high velocity blocks, which extend laterally 200–300 km offshore into the Atlantic Ocean. Between the craton and these deep-rooted blocks, we find a prominent narrow band of low velocities that roughly follows the southern and eastern Laurentia rift margin and extends into New England. We suggest that the lithosphere along this band of low velocities may be thinned due to the combined effects of repeated rifting processes and northward extension of the hotspot related Bermuda low-velocity channel across the New England region. We propose that the deep rooted high velocity blocks east of the Laurentia margin represent the Proterozoic Gondwanian terranes of pan-African affinity, which were captured during the Rodinia formation but left behind after the opening of the Atlantic Ocean. Our results suggest that recurring episodes of tectonic events that are well exposed at the surface also leave persistent scars in the continental lithosphere mantle, marked by isotropic and radially anisotropic velocity anomalies that reach as deep as 100–150 km.

Zunino et al. (2016), Constitution and Structure of Earth's Mantle: Insights from Mineral Physics and Seismology, in Integrated Imaging of the Earth: Theory and Applications, AGU Monograph Series, pp. 219–243: *This chapter describes a quantitative approach that integrates data and results from mineral physics, petrological analyses, and geophysical inverse calculations to map geophysical data directly for mantle composition and thermal state. Seismic tomography has proved an important tool to image the inaccessible parts of the Earth. Computation of physical properties using thermodynamic models is described and discussed, and an application of the joint inverse methodology is illustrated in a case study where mantle composition and thermal state beneath continental Australia is determined directly from seismic data. There is a growing consensus that the cause of the imaged wavespeed anomalies not only relates to variations in temperature, but also bears a strong compositional component. However, separation of thermal and chemical effects from seismic wave speeds alone is difficult and is further complicated by the general insensitivity of seismic wave speeds to the density contrasts that are responsible for driving mantle convection.*

1.2 Non-periodic homogenization of the elastic wave equation

The homogenization method emerged in the seventies from research in micromechanics for predicting the macroscopic response of composite and random materials to either static or dynamic excitations (Bensoussan et al., 1978; Papanicolaou and Varadhan, 1981; Sanchez-Palencia, 1980). Since then, the method has been successfully applied to many physical processes, such as heat transfer (e.g. Allaire and Habibi, 2013), Stokes flow (e.g. Hornung, 1997), neutronic diffusion (e.g. Allaire and Capdeboscq, 2000), magnetization (e.g. Santugini-Repiquet, 2007) and elastic wave propagation (e.g. Boutin and Auriault, 1993; Fish and Chen, 2001, 2004; Parnell and Abrahams, 2008; Bacigalupo and Gambarotta, 2014). In this latter field, it has been adapted to non-periodic media by Capdeville and Marigo (2007, 2008); Capdeville et al. (2010a,b); Guillot et al. (2010), which has opened the path to the upscaling of general elastic media, with no constraint on the shape and size of the heterogeneities.

1.2.1 Basics

Periodic homogenization

The homogenization theory relies on an ansatz for the solution of the physical problem in consideration. The solution is indeed postulated to be a two-scale asymptotic expansion. For instance, the solution $u(x)$ of the 1D elastodynamic problem is written

$$u(x) = \sum_{i=0}^{+\infty} \varepsilon^i u_i \left(x, \frac{x}{\varepsilon} \right), \quad (1.11)$$

where $\varepsilon = l/\lambda_{min}$ is the ratio of the periodic cell which constitutes the 1D medium to the minimum wavelength propagating in this medium. By definition, l is microscopic and λ_{min} is macroscopic. ε , which is called scaling parameter, is therefore much smaller than 1. It enables to explicitly separate

the scales within the coefficients u_i of series (1.11), x being the macroscopic scale and $y = \frac{x}{\varepsilon}$ being the microscopic scale. Plugging series (1.11) into the elastodynamic problem (i.e. the wave equation and Hooke's law) yields a cascade of equations which can be solved for each i using the average over the periodic cell $\langle f \rangle(x) = \frac{1}{\lambda_{min}} \int_0^{\lambda_{min}} f(x, y) dy, \forall f: \mathbb{R}^2 \rightarrow \mathbb{R}$, and the periodicity in y of the problem. It then turns out that the zeroth-order term u_0 does not depend on the microscopic scale y . One shows that u_0 actually is the solution of the so-called homogenized problem, which is a classical elastodynamic problem with homogeneous effective properties

$$E^* = \left\langle \frac{1}{E(y)} \right\rangle^{-1} \quad \text{and} \quad \rho^* = \langle \rho(y) \rangle, \quad (1.12)$$

E being the Young modulus and ρ the density. While the effective density ρ^* is the average of the initial density ρ , we note that the quantity to average for obtaining the effective Young modulus E^* is the inverse of the initial Young modulus E .

A crucial point in the homogenization theory is that the asymptotic convergence can be proved mathematically. One actually does not directly show the convergence of (1.11) towards the true solution. Rather than studying a single problem for the physically relevant value of ε , one considers a sequence of problems indexed by ε which is now regarded as a small parameter going to zero. The demonstrated convergence then is $\lim_{\varepsilon \rightarrow 0} u_\varepsilon = u_0$ (e.g. Allaire, 1992). Such a demonstration provides rigorous mathematical foundations for the homogenization theory. Of course, in practice, ε does not tend to zero; it is what it is, and computing higher order terms may be valuable for reaching a better precision (e.g. Fish and Chen, 2001).

Non-periodic homogenization

The non-periodic homogenization cannot rely on the scaling parameter ε defined in the periodic case. It is necessary to introduce another parameter $\varepsilon_0 = \lambda_0/\lambda_{min}$, where $\lambda_0 \ll \lambda_{min}$ is the limit between the small and the large length scales (λ_{min} still being the minimum wavelength propagating in the medium). ε_0 therefore is a user-defined parameter which tells what is considered as small and what is considered as large. Among other things, this parameter enables to build a low-pass filter $\mathcal{F}^{\varepsilon_0}$. When applied to a given quantity defined in the domain, this filter simply removes all the scales smaller than λ_0 from this quantity.

As in the periodic case, the zeroth-order term of the non-periodic homogenization is the solution of a classical elastodynamic problem with effective properties. To compute these upscaled properties, an elastostatic equation has to be solved with periodic boundary conditions. Denoting by \mathbf{C} the 4th-order stiffness tensor of the initial medium, we can write this equation as

$$\nabla \cdot \left\{ \mathbf{C} : \left[\mathbf{I} + \frac{1}{2} (\nabla \chi + {}^t \nabla \chi) \right] \right\} = \mathbf{0}, \quad (1.13)$$

where \mathbf{I} is the 4th-order identity tensor and the solution χ is a 3rd-order tensor known as the corrector. This solution is used to build two intermediate 4th-order tensors:

$$\mathbf{G} = \frac{1}{2} (\nabla \chi + {}^t \nabla \chi) + \mathbf{I} \quad \text{and} \quad \mathbf{H} = \mathbf{C} : \mathbf{G}. \quad (1.14)$$

\mathbf{G} and \mathbf{H} are called strain and stress concentrators, respectively. After filtering the components of these two tensors, the effective stiffness tensor \mathbf{C}^* is finally obtained with the double contraction

$$\mathbf{C}^* = \mathcal{F}^{\varepsilon_0}(\mathbf{H}) : \mathcal{F}^{\varepsilon_0}(\mathbf{G})^{-1}. \quad (1.15)$$

As for the effective density ρ^* , it is obtained by simply filtering the initial density ρ :

$$\rho^* = \mathcal{F}^{\varepsilon_0}(\rho). \quad (1.16)$$

1.2.2 Contributions

In the last decade, my main contributions to the non-periodic homogenization were the implementation of the method in 3D using a finite element analysis (Cupillard and Capdeville, 2017, 2018) and the application of it to complex geological media (Cupillard et al., 2015; Cupillard and Capdeville, 2018; Capdeville et al., 2020). This effort allowed the benefit of working with homogenized media to be shown (Cupillard and Capdeville, 2018; Capdeville et al., 2020) and apparent anisotropy to be investigated (Cupillard et al., 2020).

Finite element resolution in 3D

Differential equation (1.13) can be seen as a classic elastostatic problem with a specific load consisting in the divergence of the elastic tensor \mathbf{C} . Such a divergence yields a 3rd-order tensor $\partial_i C_{ijkl}$. Thanks to the symmetry of \mathbf{C} ($C_{ijkl} = C_{ijlk}$), this tensor reduces to six force vectors. To determine the six corresponding displacements, i.e. to solve equation (1.13), I implemented a tetrahedral finite element method (Cupillard and Capdeville, 2017, 2018). The resulting code offers polynomial interpolations of degree 1, 2 or 3 and various quadrature rules which make it possible to investigate the behavior of the solution with respect to discretization. Moreover, both linear and quadratic tetrahedra are enabled, leading to either iso-, super- or sub-parametric elements. To solve the finite element linear system, PARDISO (Schenk and Gärtner, 2006) is chosen. Finally, the low-pass filter $\mathcal{F}^{\varepsilon_0}$ is applied in the space domain to obtain ρ^* and \mathbf{C}^* .

In theory, periodic boundary conditions are imposed to equation (1.13), meaning that the medium is supposed to repeat itself periodically in the three dimensions. When dealing with geological media, this condition is obviously not fulfilled. We therefore replace it by homogeneous Dirichlet conditions $\chi = \mathbf{0}$ at the border. The effect in the volume of such an artificial condition decays exponentially (Dumontet, 1990), so our numerical solution is meaningless in a thin layer from the border of the domain. Such meaningless values of χ do not matter, because the filtering process cannot be performed near the border anyway. Some elastic material to be convolved with the filter wavelet is actually missing there, so we are not able to compute the effective properties ρ^* and \mathbf{C}^* using equations (1.15) and (1.16). The thickness of the layer in which the filter cannot be applied is equal to half of the wavelet support. Solutions to these boundary issues in the layered case are proposed by Capdeville and Marigo (2007, 2008) and discussed in subsection 1.2.4.

When handling large models, the memory needs for achieving the computation of the effective properties can be very large. This is mainly because solving large linear systems, even sparse and symmetric, is memory-demanding. This is also because high-order tensors are involved in the homogenization process. For these two reasons, a distributed-memory computation is necessary. To speed up our code, we also perform multithreaded computations whenever possible. Moreover, efficient algorithms based on k-d trees and stack data structures are used to search for elements or points across the finite element mesh.

To test the performance of our implementation, we stress our code with a large and highly diffractive medium made of 1000000 cubes (100 in each direction). Each cube is 1 km^3 large and has isotropic properties randomly taken in the range $2000 - 4000\text{ kg.m}^{-3}$ for density, $2.5 - 5\text{ km.s}^{-1}$ for S-wave velocity, and $4 - 8\text{ km.s}^{-1}$ for P-wave velocity. Because we are not able to compute effective properties near the border of the domain, we embed the random cubes with a 13 km thick homogeneous medium (figure 1.2, left-hand side). To mesh the whole medium, 12002256 tetrahedra are used. Because we set λ_0 to 1.6 km (i.e. $\lambda_{min} = 8\text{ km}$ and $\varepsilon_0 = 0.2$), cubic interpolation functions are necessary to capture a sufficient amount of details in the solution χ of our finite element analysis. Such a high degree leads to a stiffness matrix which is 160747899^2 large (i.e. 3 components at 53582633 interpolation points). To achieve the computation, the domain is split into 100 overlapping subdomains, each of them being treated independently from the others. Calculating the effective properties in a single subdomain then requires about 116 GB and four computing hours on a Intel Xeon X5680 processor (6 cores, 3.33 GHz , 12 MB Cache). The obtained homogenized model is shown in figure 1.2, right-hand side.

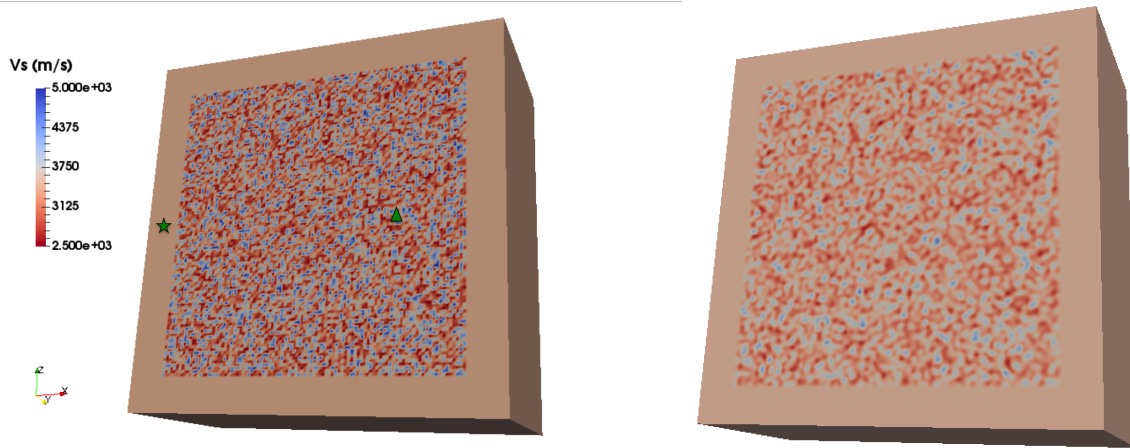


Figure 1.2: A cut in the initial (left-hand side) and effective (right-hand side) random cube models used to challenge our 3D non-periodic homogenization code. The green star represents the location of the seismic source which produces \mathbf{u} in the initial medium and \mathbf{u}^* in the effective medium. The green triangle shows the location at which the waveforms in figure 1.3 are computed. From Cupillard and Capdeville (2017).

To assess the accuracy of the computation, we compare a wavefield \mathbf{u}^* in the effective medium with a wavefield \mathbf{u} in the initial model. Both wavefields are generated by a point-source positioned in the homogeneous layer (figure 1.2) where a force-vector acts in the z direction. The obtained displacements are computed at 100 randomly chosen receiver locations using the SEM (Cupillard et al., 2012) with absorbing boundary conditions.

We first compare the wavefields \mathbf{u} and \mathbf{u}^* at a single receiver (figure 1.3). The location of the receiver is shown in figure 1.2. At this location, \mathbf{u} and \mathbf{u}^* are very similar. The residual $\mathbf{u} - \mathbf{u}^*$ is 8% at most in the ballistic wave. In the more complex, smaller-amplitude, diffracted wavefield, it reaches 12%. When computing waveforms in a model obtained by just filtering the component of the initial elastic tensor, the residual is way worse, reaching 150% in both the ballistic and the diffracted wavefields.

The homogenization theory states that the error of the wavefield reconstructed in the effective medium is a function of ε_0 : $\mathbf{u} - \mathbf{u}^* = O(\varepsilon_0)$. By computing this error over the 100 receivers for different values of ε_0 , we can check if we retrieve this convergence numerically. Figure 1.4 shows that a decay in $\varepsilon_0^{3/2}$ is actually retrieved. This is probably because all our receivers are positioned in the middle of a cube, ‘far’ from the discontinuities, so the contribution of the higher-order terms of the homogenized solution are particularly small. When filtering each component of the initial elastic tensor instead of computing the effective medium with the homogenization technique, the convergence is way worse.

Application to complex subsurface models

Our 3D homogenization code has been applied to various geological models: Furfooz (Dewaide et al., 2014), Ribaute (Caumon et al., 2009), Groningen (Romijn, 2017) and the SEG-EAGE overthrust (Aminzadeh et al., 1997). This latter application is presented below along with wave simulations in both the original medium and the homogenized medium, which emphasizes the benefit of the homogenization.

The SEG-EAGE overthrust model is $20\text{km} \times 20\text{km} \times 4\text{km}$ large. It is made of twelve faulted and folded layers. All of them are isotropic. The P-wave velocity ranges from $2500\text{m}\cdot\text{s}^{-1}$ to $6000\text{m}\cdot\text{s}^{-1}$; the S-wave velocity ranges from $1600\text{m}\cdot\text{s}^{-1}$ to $3500\text{m}\cdot\text{s}^{-1}$ (figure 1.5a). To perform wave propagation simulations in such a medium, the finite-difference method can be considered at the price of using an extremely small space-step to capture the effect of the discontinuities. Finite-element methods are more suitable for taking discontinuities into account. This kind of methods relies on meshing the medium with simple shapes like tetrahedra, hexahedra, pyramids, prisms, or a mix of those. Because tetrahedra enable the maximum flexibility, we use them to mesh the overthrust (figure 1.5a’) and we rely on a mass-lumped

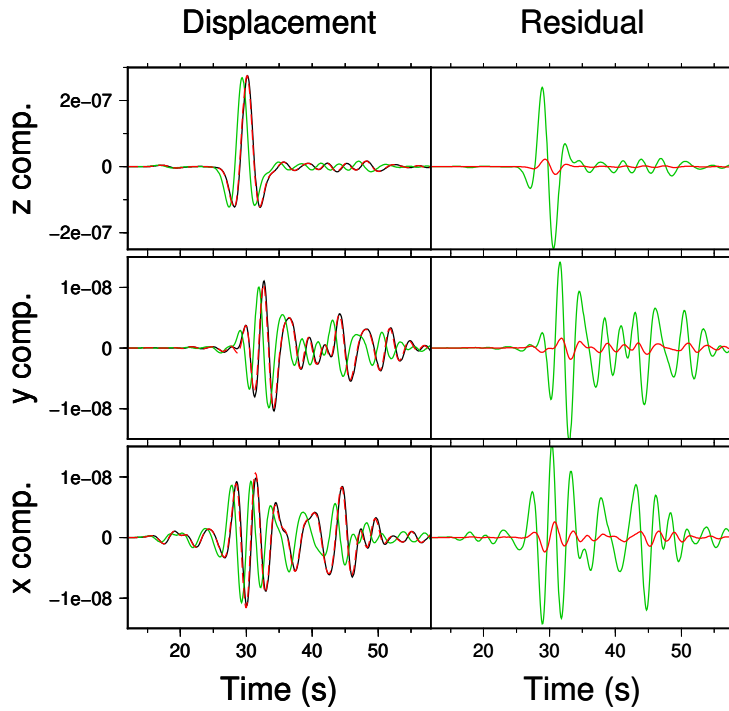


Figure 1.3: Comparison of the wavefield \mathbf{u} in the initial random cube model (black line) with the wavefield \mathbf{u}^* in the corresponding effective medium (red dashed line) computed at the receiver shown in figure 1.2. Both the ballistic wave (visible in the z component) and the diffracted wavefield (visible in the three components) are well-recovered in the effective medium. When filtering each component of the initial elastic tensor instead of computing \mathbf{C}^* with the homogenization technique, the green waveforms are obtained, which do not fit \mathbf{u} at all. From Cupillard and Capdeville (2017).

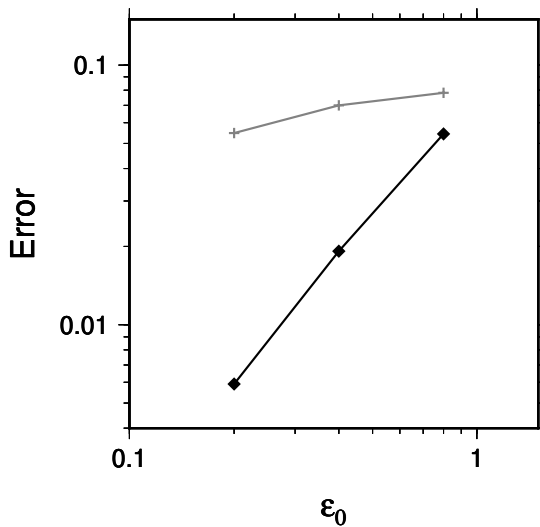


Figure 1.4: Black diamonds show the error $\frac{\sum_{i=1}^{100} \sqrt{\int (\mathbf{u}_i - \mathbf{u}_i^*)^2 dt}}{\sqrt{\int \mathbf{u}_i^2 dt}}$ (i being the index of the receivers) for three different values of ϵ_0 (0.2, 0.4 and 0.8). Gray crosses show the same error for a wavefield computed in a model obtained by just filtering the component of the initial elastic tensor. From Cupillard and Capdeville (2017).

finite-element method (Geevers et al., 2019) to compute seismic wave propagation. After an important work for building the best possible mesh, using various surface and volume meshing algorithms in a row (Pellerin et al., 2014; Si, 2015; Kononov et al., 2012) to benefit from the relevant features of each of them, we obtain 75.5 cm for the inner-sphere radius of the smallest element. Such a small value leads to a small time-step within the wave propagation simulation, for numerical stability reasons. Computing a 12 s long wavefield (hereafter denoted by \mathbf{u}) for a single $f_0 = 3.125\text{ Hz}$ Ricker source therefore requires 12.6 days on 40 Skylake cores.

Let us now compute the effective overthrust medium. To do so, we use the tetrahedral finite-element code described in the previous paragraph. This code first solves equation (1.13), which is a static problem so there is no time-step involved and small elements in the mesh does not lead to computational limitations. The main difficulty when solving this problem is to handle the large memory requirement. We therefore cut the model into 200 overlapping subdomains, each of them being treated independently from the others. λ_{min} and ε_0 are set to 200 m and 0.75 , respectively, so $\lambda_0 = 150\text{ m}$. In this configuration, the whole computation (i.e. the resolution of the cell problem followed by the filtering of the stress and strain concentrators) requires 3 hours and 100 Gb on a single PowerEdge M610 for each subdomain. The result is shown in figure 1.5b and c.

By construction, the homogenized overthrust is smooth: it has no spatial variations smaller than λ_0 . Computing waveforms in it therefore is very light: the mesh no longer need to honor geological structures, and the size of the elements is constrained by λ_0 only. Figure 1.5b' shows a zoom in a regular hexahedral mesh of the homogenized overthrust. All the elements are 200^3 m^3 large. They are used in a degree-6 spectral element method (Cupillard et al., 2012) to compute the zeroth-order displacement \mathbf{u}^* corresponding to the wavefield \mathbf{u} generated in the original overthrust model. The computation cost of the spectral-element simulation is 4163 s on two Xeon Gold 6130 processors, which is 260 times less than the computation cost required for \mathbf{u} . A comparison between \mathbf{u}^* and \mathbf{u} at a given point in space is shown in figure 1.6. The error at this point is 8.73% . Such an error averaged over 200 randomly-positionned points is 7.57% .

Using the same regular hexahedral mesh, we perform a spectral-element simulation in the original overthrust model. In this case, the geological structures are not honored by the mesh. They are smoothed by the numerical method itself, which is not a physical smoothing. As a consequence, the obtained wavefield $\mathbf{u}^{\text{brutal}}$ does not match \mathbf{u} . The error between the two wavefields reaches 21.8% . Refining the regular mesh (using 150^3 m^3 , 120^3 m^3 , 100^3 m^3 , 80^3 m^3 and 60^3 m^3 large elements), the error decreases but the computation cost increases drastically, as shown in figure 1.7. This figure also shows that the refinement has no impact on the wavefield computed in the homogenized medium. This is because all the heterogeneities of the medium are properly captured by the coarsest mesh (i.e. 200^3 m^3 large elements).

1.2.3 Perspective №1: Homogenization of multiscale fractured media

Tectonic processes and the industrial exploitation of the subsurface induce brittle deformations in the earth crust, leading to fractures at all scales. These fractures are organized in networks which are basically characterized by their density, connectivity and distribution of aperture, length and orientation. Determining these parameters are essential for predicting the hydrogeological behavior of reservoirs or understanding the fatigue of soils and engineering structures. However, direct measurements of fracture parameters are rarely available. Apart from outcrops, cores and borehole images, fractured rocks are seen in an effective way through mechanical properties derived from mechanical tests or seismic wave data. It is therefore of practical interest to better understand how the waves upscale fractured media.

Geological observations have evidenced that a power law is appropriate to describe the density of a fracture set as a function of the fracture size (Allègre et al., 1982; Bour and Davy, 1997; Bonnet et al., 2001; Bour et al., 2002; Nieto-Samaniego et al., 2005; Neuman, 2008). Such a law suggests that a fractured medium accommodates a continuum of fracture lengths. Nevertheless, for either theoretical or computational reasons, studies on seismic wave propagation in fractured media have been restricted to a short range of fracture sizes so far.

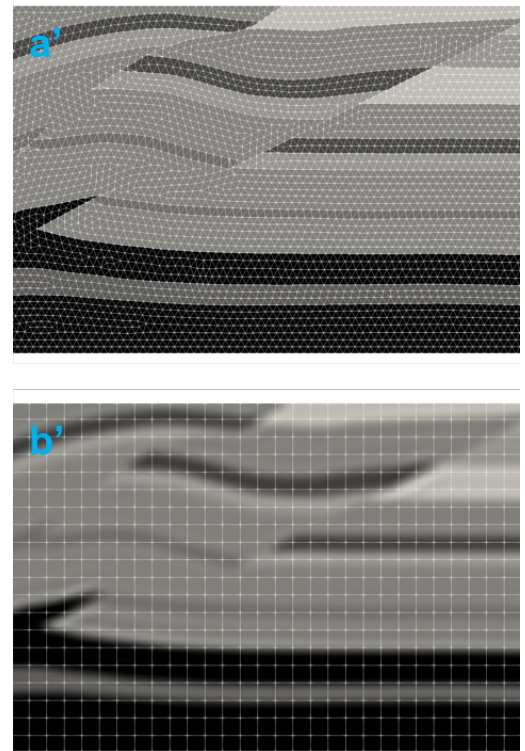
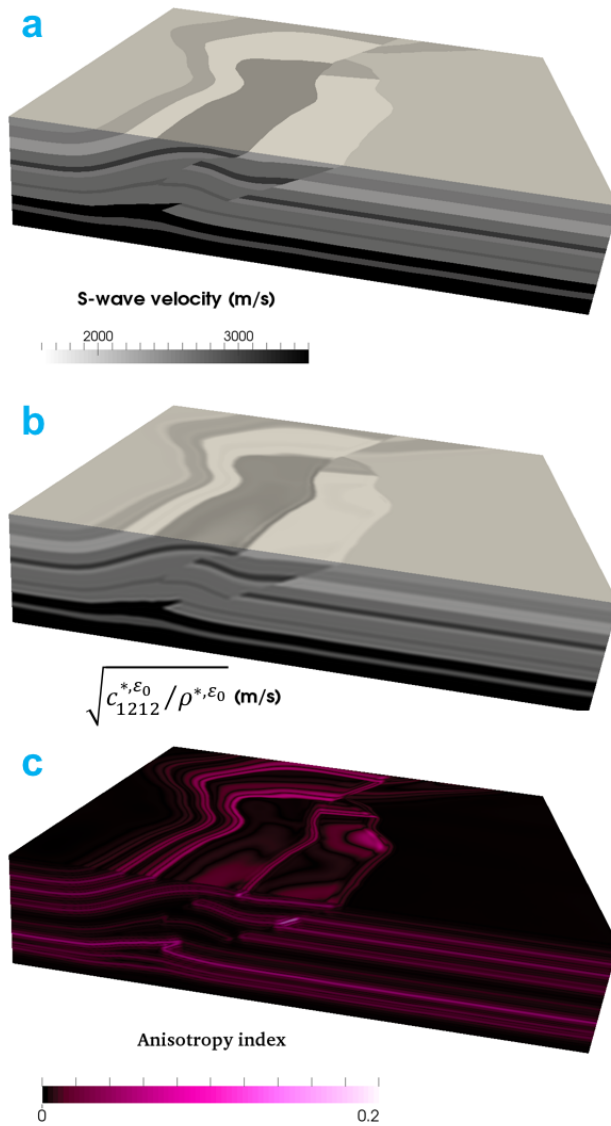


Figure 1.5: a) S-wave velocity structure of the overthrust model. a') Zoom in a lateral border of the model meshed with tetrahedra. b) The homogenized version of the model; the plotted quantity is the Sh-wave velocity. b') Zoom in a lateral border of the homogenized model meshed with hexahedra. c) Total anisotropy of the homogenized overthrust. From Capdeville et al. (2020).

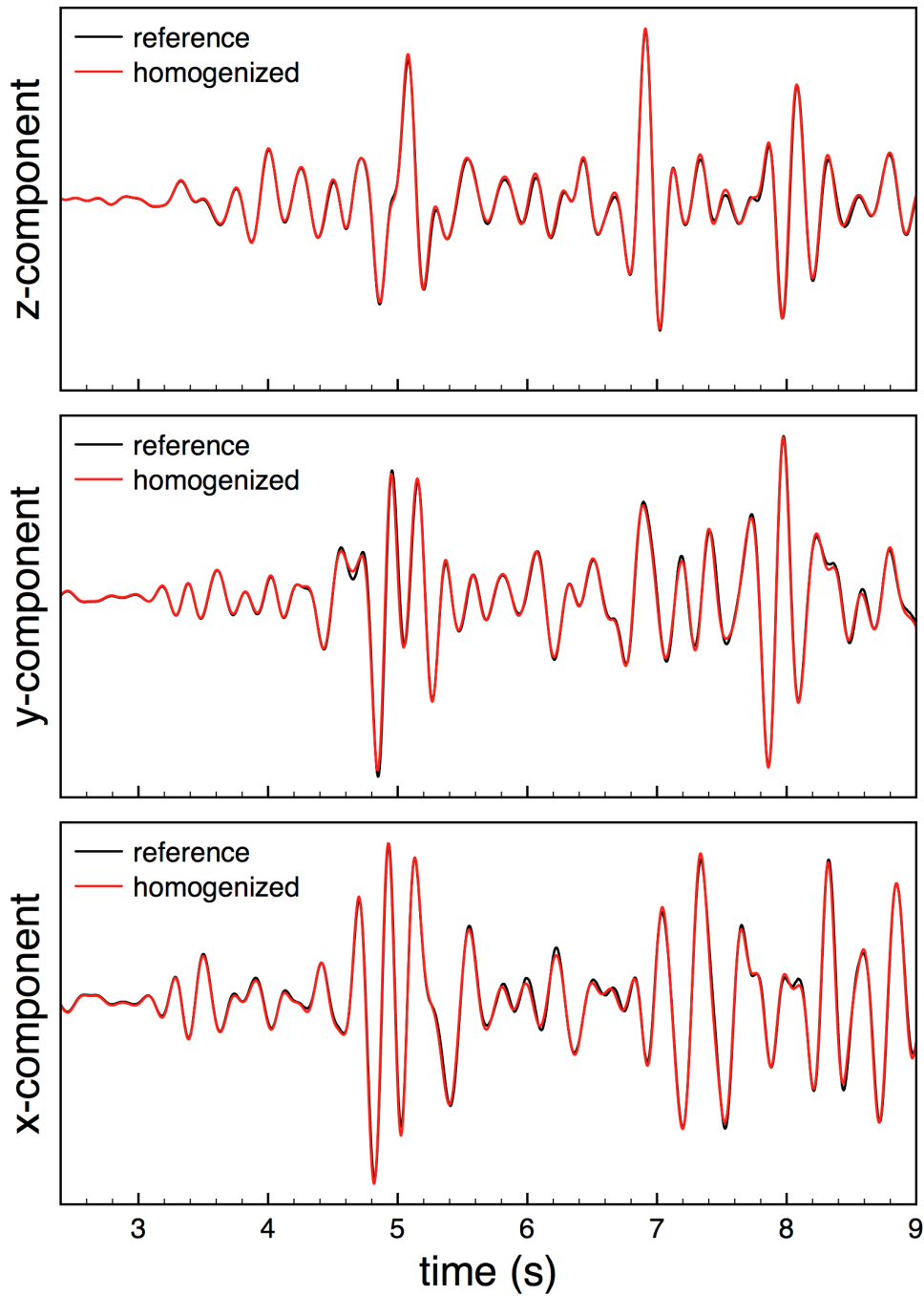


Figure 1.6: Comparison between \mathbf{u} (black) and \mathbf{u}^* (red) at a randomly-chosen point in space. Many wiggles are observed because there is no absorbing boundaries in the two simulations. From Capdeville et al. (2020).

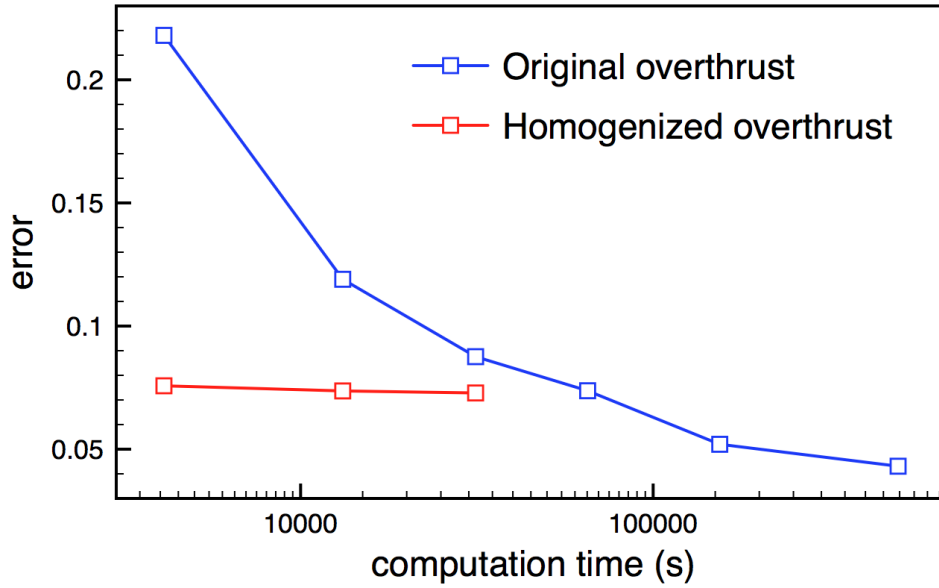


Figure 1.7: Error of $\mathbf{u}^{\text{brutal}}$ (blue) and \mathbf{u}^* (red) with respect to \mathbf{u} . Both $\mathbf{u}^{\text{brutal}}$ and \mathbf{u}^* are computed using a regular hexahedral mesh within a spectral-element method. The error of these two wavefields is plotted as a function of the computation time associated with the size of the hexahedra ($200^3 m^3$, $150^3 m^3$, $120^3 m^3$, $100^3 m^3$, $80^3 m^3$ and $60^3 m^3$). From Capdeville et al. (2020).

Two main approaches have been developed to study the seismic wave propagation in fractured media. The first one relies on numerical simulations of waves in explicit representations of fractures as planar displacement jumps ensured by the so-called linear slip model (e.g. Vasilyeva et al., 2019; Cho et al., 2019). For computational reasons, this approach meets considerable difficulties in handling fractures smaller than the wavelength and is often limited to 2D applications. The second approach is based on effective medium theories which rely on an Elementary Representative Volume (ERV) (e.g. Schoenberg, 1980; Hudson, 1981). Assuming the ERV small with respect to the wavelength, then it is in a quasi-static regime of stress so that techniques from micromechanics can be used to compute an equivalent medium. In the case of penny-shaped cracks, these techniques can provide analytical results whereas a numerical framing is necessary in more complex cases (Grechka and Kachanov, 2006a,b).

In the frame of a 18-month post-doc, Anaïs Ibourichène started working on the application of the non-periodic homogenization method to fractured media to go beyond the ERV setting and explore the effect of a wide distribution of crack sizes on seismic wave propagation. The preliminary results obtained by Anaïs in 2D (Ibourichène et al., 2021) are presented below. First, the effective properties computed using the homogenization in the case of Eshelby problems (Eshelby, 1957) are compared to analytical solutions provided by various effective medium theories. Then, the homogenization is applied to a synthetic mass rock which contains different fracture sets, each set being characterized by a length l (up to $\lambda_{\min}/4$) and a density $d \sim l^{-3}$.

Comparison with effective medium theories in the Eshelby frame

As a preliminary investigation, the analytical solutions provided by various effective medium theories are compared with the results of the non-periodic homogenization in the case of elliptical inclusions in a host matrix (Eshelby, 1957). We represent the fractures by ellipses and we embed them in a host matrix with properties similar to crustal rocks: $V_P = 4000 m.s^{-1}$, $V_S = 2500 m.s^{-1}$, $\rho = 2400 kg.m^{-3}$. To mimick empty cracks, the fractures are filled with a thousandth of these values: $V_P = 4 m.s^{-1}$, $V_S = 2.5 m.s^{-1}$, $\rho = 2.4 kg.m^{-3}$. For all the computations, we set the density of fracture centers to $(8/\lambda_{\min})^2$ and we align the major axis of all the ellipses along the x -direction. Moreover, interactions between fractures are not allowed to ensure no departure from the length we set in the different computations. The Gmsh

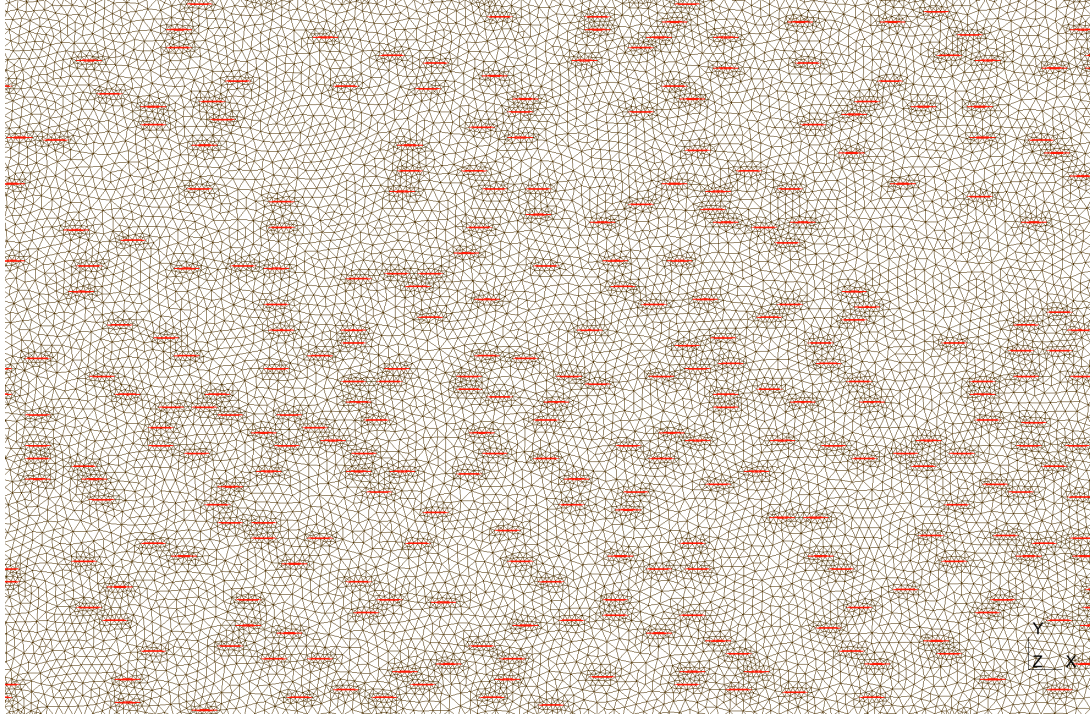


Figure 1.8: A triangular mesh of a fractured medium. In this example, the length of the fractures is equal to $\lambda_{min}/16$ and the aspect ratio is equal to 0.05.

software (Geuzaine and Remacle, 2007) is used to build the 2D mesh in which the homogenization will be performed. As shown in figure 1.8, the mesh is refined in the vicinity of the fractures to ensure that the heterogeneities will be correctly handled by our homogenization code.

As a first experiment, the aspect ratio of the elliptical fractures varies from $1/50$ to $2/5$ and we set the length of the fractures to $\lambda_{min}/16$. This length is much smaller than the minimum wavelength, so the effective medium theories against which we want to compare the homogenization are valid. Posing $1 = xx$, $2 = yy$ and $3 = xy$, we plot the effective C_{11} , C_{22} , C_{33} and C_{12} components obtained by the various theories (figure 1.9). Among the six independent components of the effective stiffness matrix, these four are the significant ones because the effective medium of a rock mass affected by horizontally aligned cracks is vertical transversely isotropic (VTI). In other words, C_{13} and C_{23} vanish. The results show that the homogenized stiffness matrix fits the solutions derived from the self-consistent (O’Connell and Budiansky, 1974; Budiansky and O’Connell, 1976; Henyey and Pomphrey, 1982) and the differential (Vavakin and Salganik, 1975; Hashin, 1988; Saenger and Shapiro, 2002; Orlovsky et al., 2003; Sævik et al., 2013) schemes when the aspect ratio is lower than 0.1. For larger aspect ratios, the homogenization result gets closer to the Mori-Tanaka (Mori and Tanaka, 1973) and the dual dilute (Sevostianov and Kachanov, 2012) solutions. These two schemes are both based on the non-interaction approximation so they are expected to give similar results.

In a second experiment, the aspect ratio of the fractures is set to 0.05 and we vary the length of the ellipses from 0.3 to $0.8 \times \lambda_{min}/16$. Figure 1.10 shows that the effective C_{11} , C_{22} and C_{12} components from the homogenization are close to the solution obtained using the differential scheme. Because the aspect ratio is small, the homogenized stiffness matrix is close to the self-consistent solution as well. As for the Mori-Tanaka and the dual dilute schemes, they both provide C_{ij} values higher than the homogenized ones. Two other analytical solutions are shown: the Castañeda and Willis (1995) one, which provides C_{ij} values smaller than the homogenized ones, and the dilute one (e.g. Sevostianov and Kachanov, 2012), which doesn’t perform well in this case.

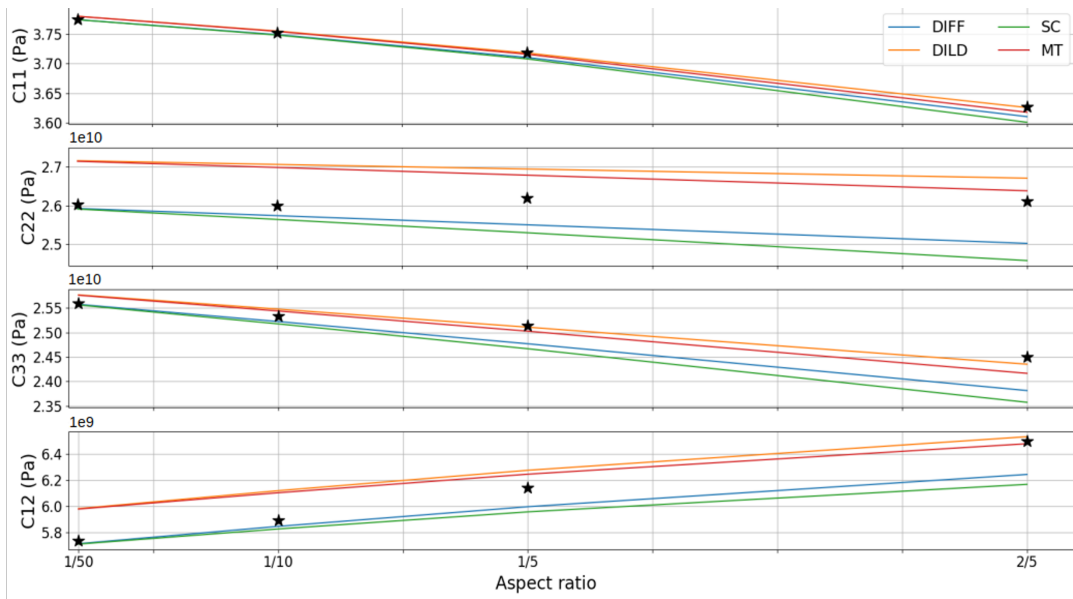


Figure 1.9: C_{11} , C_{22} , C_{33} and C_{12} components of the effective stiffness matrix as a function of the aspect ratio of the elliptical fractures. The lines refer to the solutions from the different effective medium theories (MT: Mori-Tanaka; DIFF: differential; DILD: dual dilute; SC: self-consistent) and the black stars to the solution estimated by the non-periodic homogenization.

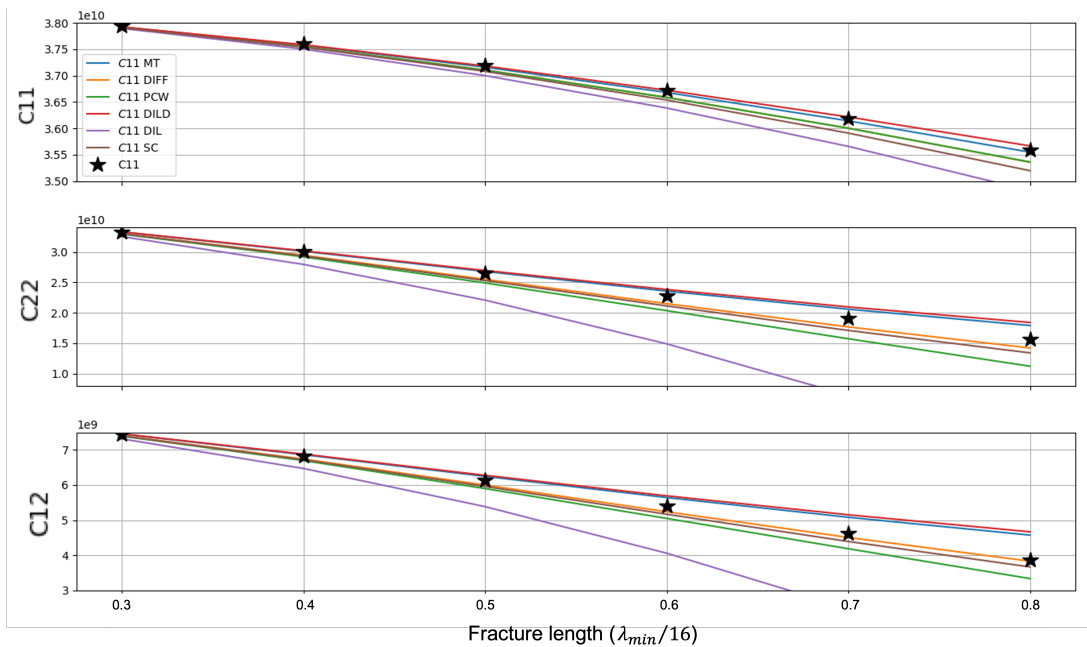


Figure 1.10: C_{11} , C_{22} , and C_{12} components of the effective stiffness matrix as a function of the fracture length. The lines refer to the solutions provided by the different effective medium schemes (MT: Mori-Tanaka; DIFF: differential; DILD: dual dilute; SC: self-consistent; PCW: Castaneda-Willis; DIL: dilute) and the black stars correspond to the solutions calculated by the non-periodic homogenization.

Homogenization of fractures with a density-length distribution following a power law

As a first application of the homogenization to a medium which contains fractures of different sizes, we consider an isotropic host matrix in which we progressively add larger and larger fractures, following a power law $d \sim l^{-\alpha}$ for the density of fractures d as a function of the fracture length l . Geological observations suggest that $2 \leq \alpha \leq 3$ (Allègre et al., 1982; Bour and Davy, 1997; Bonnet et al., 2001; Bour et al., 2002; Nieto-Samaniego et al., 2005; Neuman, 2008). In our application, we set α equal to 3.

The fractures are generated using a Discrete Fracture Network (DFN) algorithm which creates horizontally aligned cracks with the desired density. The fractures are elliptical with an aspect ratio equal to 0.05. The first set of fractures that we introduce in the host matrix has a density equal to $d_{max} = (8/\lambda_{min})^2$ and a fracture length equal to $l_{min} = \lambda_{min}/16$. These parameters lead to medium number 1 in figure 1.11. On the top of this medium, a second set of fractures is introduced using a length equal to $2l_{min}$. Following the power law, the density of this set is $d_{max}/8$. This leads to medium 2 in figure 1.11. Finally, a third medium is generated by introducing a third set of fractures with $l = 4l_{min}$ and $d = d_{max}/64$ (figure 1.11). The fracture length of this set is no longer small with respect to the minimum wavelength, so the effective medium theories mentioned in the previous paragraph could not be applied to medium 3.

We homogenize the three fractured media. The obtained density and both the vertical and the horizontal P-wave velocity are plotted in figure 1.11. Because the homogenized media are heterogeneous, each effective quantity is shown as a dispersion bar around a mean value. As expected, when adding more and more fractures, the density and the P-wave velocities decrease, and the amount of heterogeneities increases. Interestingly, the horizontal P-wave velocity starts changing significantly when adding a few long fractures (medium 3) whereas the variations of the density and the vertical P-wave velocity become a bit smaller. This is one of the effects of having multiple fracture lengths with a realistic density-length distribution in rocks.

In this preliminary experiment, the range of fracture lengths is quite narrow. To better understand how the seismic waves behave as a function of the power law $d \sim l^{-\alpha}$, we have to push this work further by considering a larger range of lengths and different values of α . Moreover, the comparison between the non-periodic homogenization and the effective medium theories developed in mechanics asks for a deeper analysis, looking at the details of each theory to understand the similarities and the differences of the results.

1.2.4 Perspective №2: Effective media for ground motion modeling

Seismic hazard assessment and in particular the estimation of ground motion at the local scale are essential for mitigation policies as well as compensation for disasters supporting the resilience of the affected territories. Indeed, seismic waves are modified, often amplified, by local geological and geotechnical conditions. These variations are known as site effects and are of two different types: lithological site effects in the case of resonance of seismic waves in shallow geological layers and topographical site effects in presence of irregular topographic features (Bard, 1997). Such a resonance amplifies some frequencies of the waves at the surface and extends the ground motion duration, which potentially increases the seismic impact on buildings and people.

To get a quick estimation of the ground motion generated by an event that just occurred or in probabilistic studies such as Probabilistic Seismic Hazard Analysis (PSHA), Ground Motion Prediction Equations (GMPEs) are available. They are simple analytical expressions which allow translating the basic features of an earthquake (magnitude, distance, tectonic regime) in a Peak Ground Acceleration (PGA) to which the stakes located on the surface are submitted. However, these GMPEs do not explicitly depend on the nature and geometry of the geological objects in which the seismic waves propagate. Moreover, most of GMPEs are established using real seismicity catalogs in specific regional contexts, so they are hardly transposable to other regions. The large uncertainty on the PGA estimated from GMPEs (e.g. Douglas, 2003) reflects this disregard of the propagation medium.

To properly take into account the effects of shallow geological layers on seismic waves arriving at the earth surface, various numerical methods are available. We can cite the discrete wavenumber method

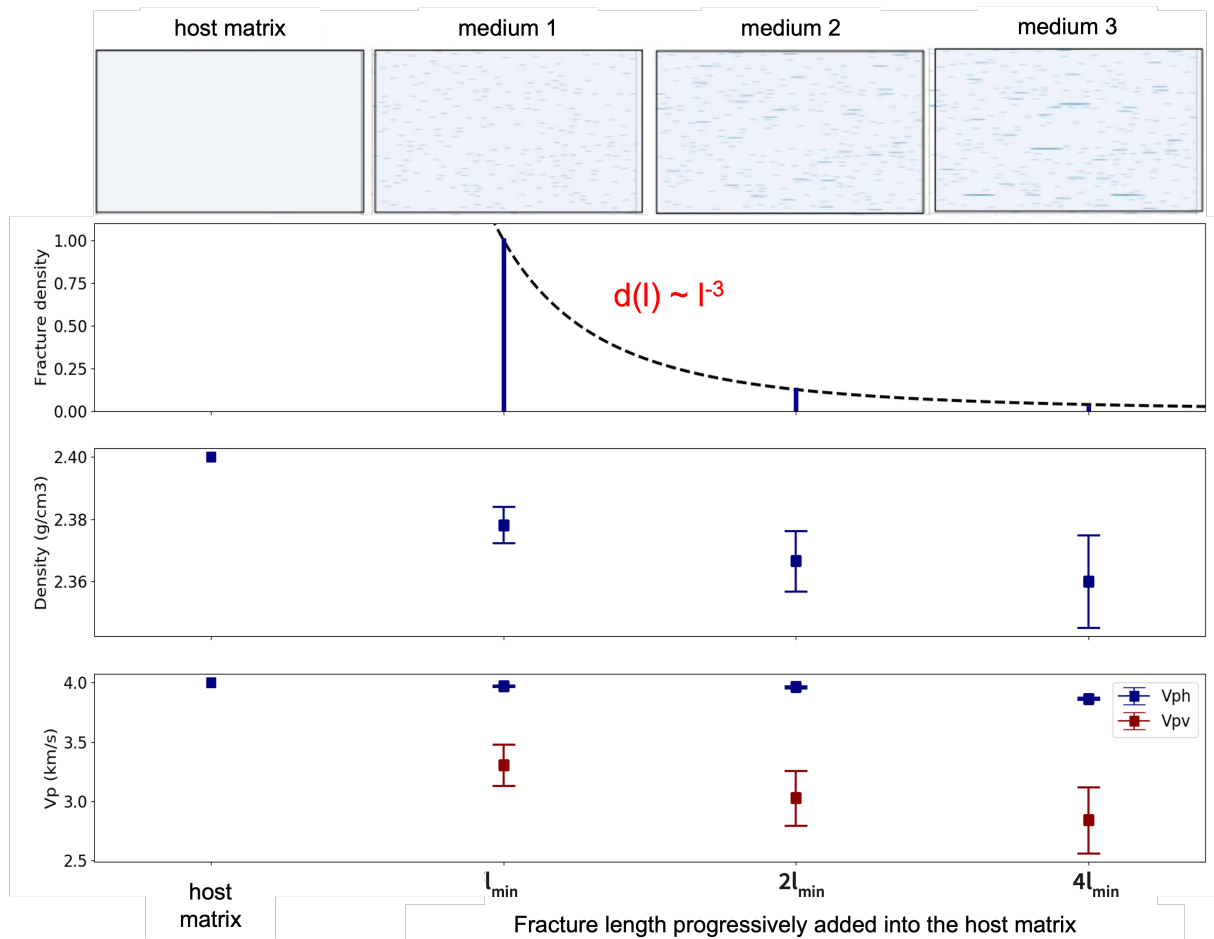


Figure 1.11: The effective density (third row) and P-wave velocities (bottom row) obtained when homogenizing the fractured media shown in the top row. These media are built by progressively adding fracture sets in a host matrix. Each fracture set is characterized by a length (from $l_{min} = \lambda_{min}/16$ to $4l_{min}$) and its associated density (second row). Modified from Ibourichène et al. (2021).

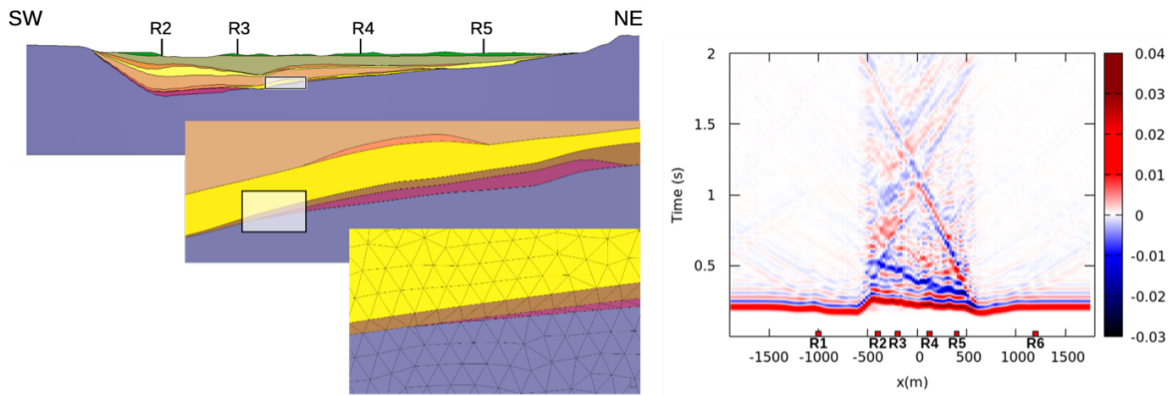


Figure 1.12: Left: Geological cross-section of the lower Var valley near the Nice airport, France. Two close-ups show several tangential contacts which imply small and elongated elements. Right: Ground motion velocity induced by a vertical incident plane wave. The maximum frequency is 25 Hz. Because of the small elements, the DGM simulation from which this result comes required 6 days on an Intel Core I7 2.70 GHz processor. Modified from Anquez et al. (2021).

(DWM, e.g. Campillo and Bouchon, 1985), the boundary element method (BEM, e.g. Chaillat et al., 2009), the finite difference method (FDM, e.g. Moczo et al., 2014), the finite element method (FEM, e.g. Bielak et al., 2005), the spectral element method (SEM, e.g. Komatitsch and Vilotte, 1998; Stupazzini et al., 2009) and the discontinuous Galerkin method (DGM, e.g. Mercerat and Glinsky, 2015). All these techniques require a proper spatial discretization of the geological interfaces (i.e. a mesh) in order to accurately account for their effects on wave propagation. In the common case where uncertainties exist about the geometry of geological structures, this discretization should be computed on several possible geological models for the rigorous assessment of the non-linear effects of geological uncertainties on the simulated wavefield (Wellmann and Caumon, 2018). However, when the geometry of the interfaces is complex, obtaining even a single deterministic mesh can be a difficult task and the computation of wave propagation can be extremely costly. In particular, when the interfaces are close one to another, involving tangential contacts and thin geological layers, the mesh required in the finite-element type methods (FEM, SEM or DGM) can be challenging to generate. Most of all, this mesh may contain a large number of elements and a few elongated elements, which results in gigantic wave simulation costs, as exemplified above (subsection 1.2.2) as well as in figure 1.12.

To deal with this difficulty, the DGM enables the so-called p -adaptivity, which consists in using a non-uniform polynomial degree distribution to relax constraints on the time-step in small elements (Dumbser et al., 2007; Etienne et al., 2010). This method is all the more efficient when associated to non-conforming meshes, as explored for finite-element type methods (Chaljub et al., 2003; Diaz and Joly, 2005), but the coupling between the coarse and fine non-conforming elements in this context can prove to be very complex, in particular for 3D meshes based on tetrahedra. Another solution for dealing with complex geometry consists in implementing hybrid methods to benefit from the advantages of one method in a given part of the space domain (e.g. the flexibility of the DGM where a complex geometry is involved) and of another method in the rest of the domain (e.g. the efficiency of the SEM where hexahedra can be easily handled). Hybrid methods include BEM-DWM (e.g. Bouchon et al., 1989), FEM-BEM (e.g. Bielak et al., 1991), FEM-FDM (e.g. Moczo et al., 1997), tetrahedral-hexahedral FEM (e.g. Ichimura et al., 2009), FEM-SEM (Casadei et al., 2002) and DGM-SEM (Terrana et al., 2017). Most of these methods call for hex-dominant meshes, which can be challenging to generate (Yamakawa and Shimada, 2003; Botella et al., 2016; Baudouin et al., 2014).

Instead of adapting the mesh to the geological medium, the non-periodic homogenization proposes to adapt the medium to the physics of wave propagation. As explained and illustrated above, this method allows smoothing a given geological medium (which contains all the scales, as shown in figure 1.12) up

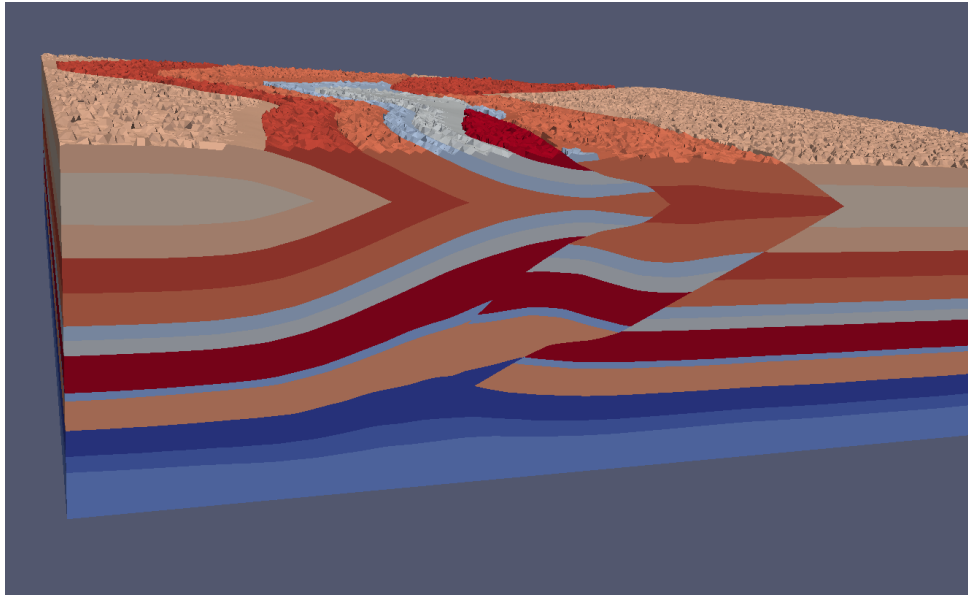


Figure 1.13: Example of a symmetric extension. The top surface of the overthrust model shown in figure 1.5a acts as a mirror, so tetrahedra fill the space above the surface.

to the minimum wavelength of the wavefield. The upscaled medium no longer contains discontinuities so it can be discretized using a coarse mesh, the size of the elements (and, consequently, the time-step of any wave simulation) being driven by the wavefield frequency instead of the geological structures.

The main difficulty when applying the homogenization to ground motion modeling is that a free surface is involved, so the boundary issues mentioned in subsection 1.2.2 have to be tackled to compute accurate effective properties at the border of the domain. In the layered case, Capdeville and Marigo (2007) showed that any extension of the medium above the free surface (e.g. extending the boundary values of density and elastic coefficients) leads to a zeroth-order solution, but such a solution is not accurate enough for surface or grazing incident waves. Nonetheless, the symmetric extension (i.e. when the boundary acts as a mirror, e.g. figure 1.13) naturally leads to a first-order solution. Moreover, deriving a second-order correction is relatively easy in this case, which enables to compute an accurate solution (Capdeville and Marigo, 2007). Such a result is still valid when smooth lateral variations are present. In alluvial basins and sedimentary valleys where amplifications occur, strong lateral variations can be involved, especially at the bedrock edges (e.g. figure 1.12). Moreover, the free surface can be non-flat. In these cases, a relevant extension has to be found, and improving the zeroth-order solution locally by an adaptive filtering can be considered.

1.3 Inferring the earth interior and seismic source features from full waveforms

1.3.1 Basics

Deterministic inversion consists in finding a model \mathbf{m} which explains some observed data \mathbf{d}^{obs} . To do so, an objective function $\chi(\mathbf{m})$ involving the misfit between the observed data and synthetic data $\mathbf{d}^{syn} = g(\mathbf{m})$, is minimized. g represents the forward operator, i.e. a tool which enables to evaluate the response of a model \mathbf{m} to the phenomenon at play. For instance, when inverting P-wave arrival-times for estimating the P-wave speed structure of a given medium, g can be a ray tracing software. Table 1.1 shows some examples of forward modeling approaches as a function of the data type to be modeled. It also provides some references to articles which implement these approaches to infer seismic velocities within the earth.

Data type and χ	g	Ray theory	Numerical simulations
Infinite frequency arrival-times $\frac{1}{2} \sum_i [\mathbb{T}_i^{syn}(\mathbf{m}) - \mathbb{T}_i^{obs}]^2$		e.g. Ritsema et al (1999)	e.g. Noble et al (2014) Solving the eikonal equation
Finite frequency arrival-times $\frac{1}{2} \sum_{i,\omega} [\mathbb{T}_i^{syn}(\mathbf{m}, \omega) - \mathbb{T}_i^{obs}(\omega)]^2$		e.g. Montelli et al (2004)	e.g. Tape et al (2010) e.g. Bozdag et al (2016) Solving the wave equation
Full waveforms $\frac{1}{2} \sum_{e,r,t} \ \mathbf{u}_{e,r}^{syn}(\mathbf{m}, t) - \mathbf{u}_{e,r}^{obs}(t)\ _2^2$			e.g. Virieux & Operto (2009) e.g. Fichtner et al (2009) Solving the wave equation

Table 1.1: References to some articles which involve a given forward modeling tool g as a function of the seismic data type the authors aim at inverting. The ultimate goal of all the mentioned articles is to image the seismic velocities within the earth. For each type of data, the classic L^2 -norm objective function is provided.

Deterministic is opposed to *Bayesian*. Ideally, all the inverse problems should be solved in a Bayesian way, i.e. estimating the likelihood of a large amount of models to derive the posterior probability of any sample \mathbf{m} of the model space. Unfortunately, when this later is a high-dimensional space or when the forward problem is computationally demanding, such an approach is not tractable. For instance, when considering full waveforms to infer seismic velocities in 3D, computing synthetic data on a single model requires a significant amount of computational resources, and the model vector \mathbf{m} can be as large as millions, so exploring the model space is unfeasible. In this case, a deterministic approach is required.

Whatever the minimization algorithm, the gradient of the objective function $\partial\chi/\partial\mathbf{m}$ has to be computed at each iteration. Thanks to the adjoint technique (Lailly, 1983; Tarantola, 1984, 1988; Tromp et al., 2005; Tape et al., 2007; Fichtner et al., 2006a,b; Liu and Gu, 2012), this can be done in an efficient way. For each seismic event $e = 1, \dots, N_e$, the adjoint technique yields *event kernels* $K_e(\mathbf{m}; \mathbf{x})$ with respect to any given parameter (such as $V_p, V_s \dots$) at any point \mathbf{x} in the medium via the interaction of a regular wavefield, propagating from the event to the receivers, and an adjoint wavefield, propagating from the receivers back to the event. When considering the classic L^2 -norm objective function within a full waveform inversion (table 1.1), the adjoint source associated with the event e is given by

$$\mathbf{f}_e(\mathbf{m}; \mathbf{x}, t) = \sum_{r=1}^{N_r} [\mathbf{u}_{e,r}^{syn}(\mathbf{m}; T-t) - \mathbf{u}_{e,r}^{obs}(T-t)] \delta(\mathbf{x} - \mathbf{x}_r), \quad (1.17)$$

where N_r is the number of receivers and T is the length of the seismic waveforms. Then, the gradient is obtained by

$$\frac{\partial\chi}{\partial\mathbf{m}} = \int K(\mathbf{m}; \mathbf{x}) \mathbf{B}(\mathbf{x}) d\mathbf{x}, \quad (1.18)$$

where $\mathbf{B}(\mathbf{x})$ is the vector of the basis functions in which to expand the model (e.g. splines, pixels) and $K(\mathbf{m}; \mathbf{x})$ is simply the sum over e of the event kernels $K_e(\mathbf{m}; \mathbf{x})$. Such a gradient computation does not depend on neither the number of parameters nor the number of receivers. It actually requires N_e forward simulations + N_e adjoint simulations = $2N_e$ simulations at each iteration of the minimization scheme, which makes it possible to use advanced wave equation solvers to infer seismic velocities (e.g. Brossier et al., 2009; Fichtner et al., 2009, 2013; Pageot et al., 2013; Brossier et al., 2015; Operto et al., 2015; Beller et al., 2017) and seismic source features (Liu et al., 2004; Kim et al., 2011; Somala et al., 2018; Willacy et al., 2019) from full waveforms.

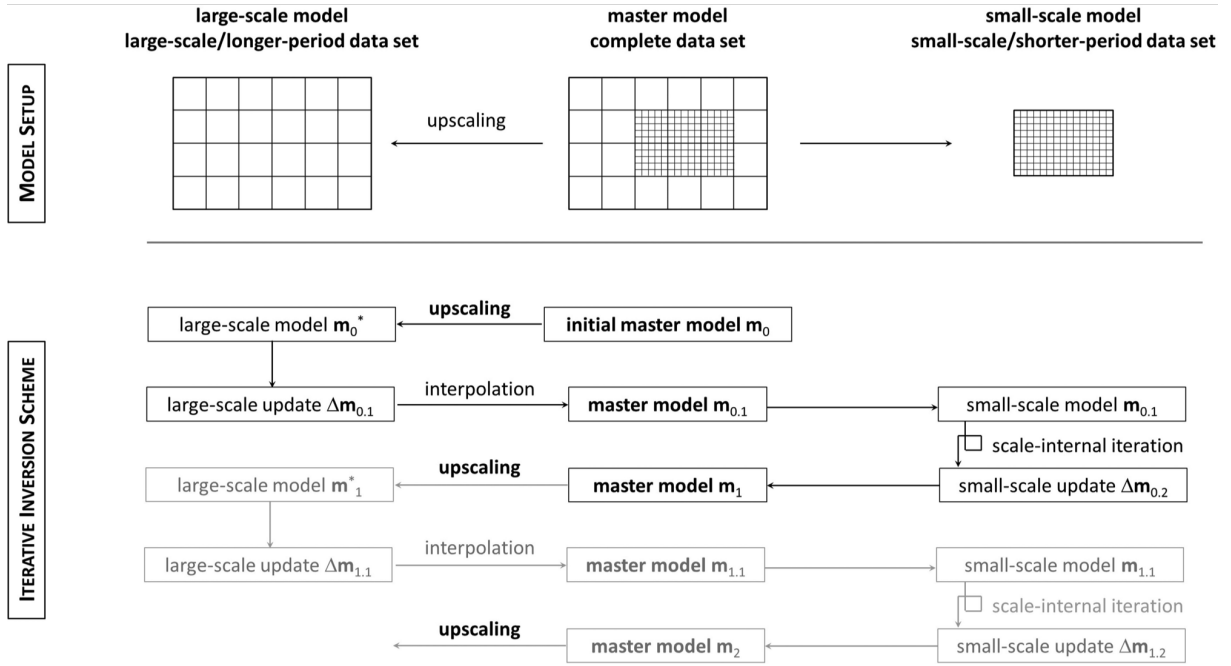


Figure 1.14: Multiscale FWI strategy proposed by Fichtner et al. (2013). The upscaling of master models to large-scale models is ensured by the homogenization method (section 1.2).

One of the main issues when performing a deterministic inversion is the risk of getting trapped in a local minimum of the objective function. To mitigate this risk, significant efforts have been made to go beyond the classic L^2 -norm (table 1.1) and to design more robust, convex objective functions (Fichtner et al., 2008; van Leeuwen and Mulder, 2010; Bozdağ et al., 2011; Ma and Hale, 2013; Métivier et al., 2016). Another issue with deterministic inversions is the lack of uncertainty estimation: only one single 'best' model is returned with no probabilistic view of the model space. In the context of full waveform inversion, remarkable papers recently came out to overcome this limitation and drew Bayesian ways of inverting seismic full waveforms based on the construction of an artificial Hamiltonian system (Fichtner and Simutè, 2018; Fichtner et al., 2019; Fichtner and Zunino, 2019; Gebraad et al., 2020) or a dipping layer parameterization combined with model stitching (Visser et al., 2019; Guo et al., 2020).

1.3.2 Contributions

In subsection 1.1.2, I mention two continental scale imaging studies (Yuan et al., 2014; Zunino et al., 2016) to which I have contributed by providing the forward modeling solution. In collaboration with Andreas Fichtner from ETH Zürich, I have also participated to the development of a full waveform inversion (FWI) method which allows assimilating seismic data over a broad range of space scales (Fichtner et al., 2013). The earth indeed is unevenly sampled by the data because both seismic stations and events are unevenly distributed. As a consequence, the maximum frequency of the data and the resolution of the tomographic models considerably varies from a region to another. The method Andreas developed aims at handling this heterogeneity. It enables to invert different dataset in a consistent way (figure 1.14). A key component of it is the homogenization technique (section 1.2) which allows including the knowledge of the small scales in the large-scale inversions.

Our multiscale FWI method was applied to Eurasia (Fichtner et al., 2013), with a focus on Anatolia where the coverage of seismic stations is particularly dense. At the continental scale, the structure is constrained by waveforms with periods between 30 and 200 s. Besides the well-known structural elements of the Eurasian mantle, our model reveals a wide variety of more subtle features, such as the Armorican Massif, the Rhine Graben and the Massif Central. Anatolia is crossed by waveforms with periods between 8 and 200 s. The model obtained shows the depth signature of the North Anatolian fault. Within

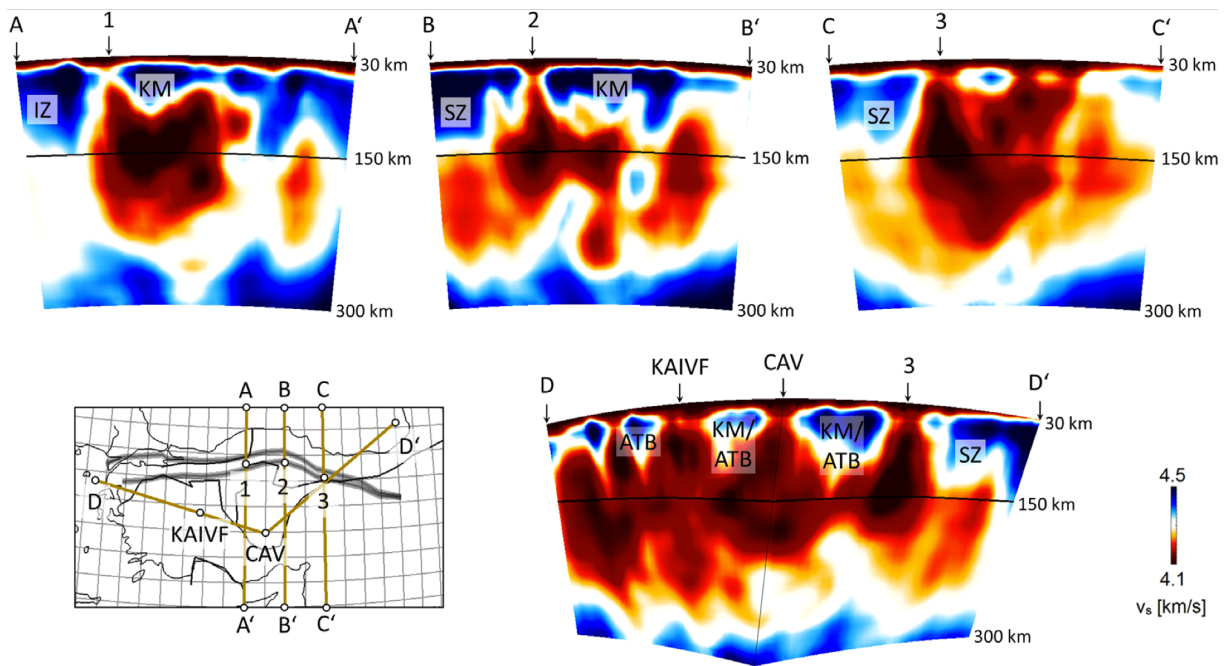


Figure 1.15: Vertical slices (stretched by a factor of 3) through the Anatolia model obtained by multiscale FWI. Key to marked features: ATB, Anatolide–Tauride Block; CAV, Central Anatolian Volcanics; IZ, İstanbul Zone; KAIVF, Kırka–Afyon–Isparta Volcanic Field; KM, Kırşehir Massif; SZ, Sakarya Zone. From Fichtner et al. (2013).

the crust, the fault zone is mostly bounded by several high-velocity blocks, suggesting that it developed along the edges of continental fragments with high rigidity. Below the crust, the surface expression of the eastern and central parts of the North Anatolian Fault Zone correlate with a pronounced low-velocity band that extends laterally over 600 km. Around 100 km depth, the low-velocity band merges into the shallow Anatolian asthenosphere, thereby providing a link to the Kırka–Afyon–Isparta Volcanic Field and the Central Anatolian Volcanics (figure 1.15). We interpret the low-velocity band beneath the North Anatolian Fault Zone as the upper-mantle expression of the Tethyan sutures that formed 60–15 Ma ago as a result of Africa–Eurasian convergence. The structurally weak suture facilitated the formation of the younger (less than 10 Ma) crustal fault zone. In this sense, the North Anatolian Fault Zone is not only a crustal feature, but a narrow zone of weakness that extends into the upper mantle.

A latter article (Fichtner et al., 2018) makes the multiscale FWI method evolutionary, allowing for posterior refinements of the model while keeping prior knowledge from the previous updates. Each refinement may come from a particular inversion technique and rely on its own parameterization (figure 1.16a). Possible conflicting updates (meaning that previously assimilated data may be worse explained by the updated model) are handled using a Gaussian approximation to find a deterministic solution that optimally agrees with the independent updates. The methodology is applied at the global scale, leading to the first-generation Collaborative Seismic Earth Model. This later aims at harnessing distributed human and computing power. It comprises twelve refinements from full seismic waveform inversion, ranging from regional crustal- to continental-scale models (figure 1.16b).

1.3.3 Perspective №1: Bayesian inversion of FWI images for reducing structural uncertainties

Owing to the lack and incompleteness of subsurface data, significant uncertainties exist on the position of structural surfaces (faults and horizons). Because structural modeling plays a key role in reservoir planning and development, as well as in various other fields in geosciences such as nuclear waste storage and geological hazard prediction, it is important to quantify and reduce these uncertainties.

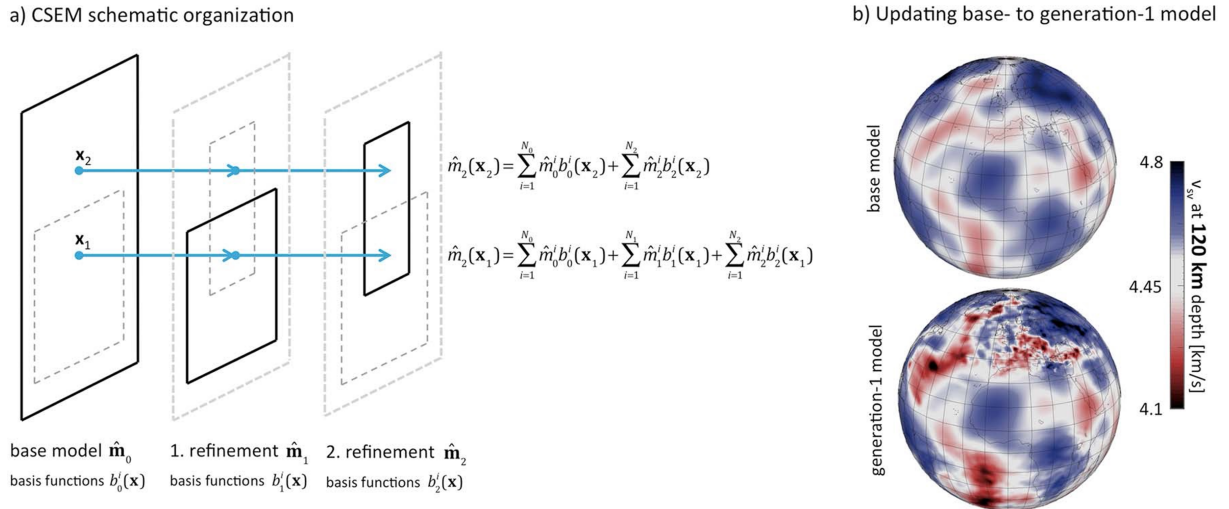


Figure 1.16: a) Schematic illustration of successive refinements $\hat{\mathbf{m}}_1$, $\hat{\mathbf{m}}_2$, ... being added to the smooth initial model $\hat{\mathbf{m}}_0$. Different updates may be parametrized differently, that is, have different types of basis functions. b) Concrete example of V_{SV} at 120-km depth in the initial model (top) and in the first-generation Collaborative Seismic Earth Model (bottom). Refinement regions visible in this view include Europe, Turkey, the Sea of Marmara region, the North Atlantic, the South Atlantic and North America. From Fichtner et al. (2018).

A common way of quantifying structural uncertainties is to sample the ensemble of acceptable models by generating a relevant number of them stochastically (e.g. Cherpeau et al., 2010). Then, dynamic reservoir data can be integrated to appraise the models and reduce the uncertainties (e.g. Suzuki et al., 2008; Cherpeau et al., 2012). In collaboration with Guillaume Caumon, I propose to integrate ancillary seismic data for reducing uncertainties on fault and horizon geometry. We see this as an essential pathway to reduce uncertainties because seismic data and flow data are sensitive to different structural features. Flow data are primarily sensitive to permeability while seismic data are sensitive to elastic properties. Moreover, both flow and seismic data are sensitive to fault throw and layer thickness, but at different scales. As a consequence, combining the two sources of information during the appraisal of models should result in reducing structural uncertainties.

Stochastic structural models being created from reflection seismic data, we plan to focus our work on the usage of borehole seismic data in order to benefit from a different illumination of the structures. As in the work of Irakarama et al. (2019), we will work on raw data to avoid biases due to stacking, migration, interpretation, time-to-depth conversion, etc. The uncertainty related to these processes will be managed by the prior sampling of uncertainties using stochastic structural modeling (Thore et al., 2002; Cherpeau and Caumon, 2015).

Raw seismic waveforms are explained by velocity models derived from FWI. As a consequence, the data could be such velocity models instead of the waveforms themselves. In this context, the appropriate forward modeling tool for mapping the model space (i.e. the geological structures) into the data space (i.e. a FWI result) is the non-periodic homogenization operator presented in section 1.2. This operator enables to compute the effective elastic medium of any structural model filled with elastic properties. The obtained smooth medium is what the waves 'see', i.e. the result of a FWI in case of sources and receivers all around the medium. Weighting homogenization results by the actual source-receiver configuration therefore leads to synthetic FWI images to be compared to the data.

The ingredients of the method to be implemented are (figure 1.17):

1. Priors on the structural parameters to be estimated along with a way of sampling them,
2. The homogenization operator for computing synthetic FWI images from structural models,
3. An appropriate function to estimate the local (i.e. spatialized) likelihood between the synthetic images and the real FWI result,

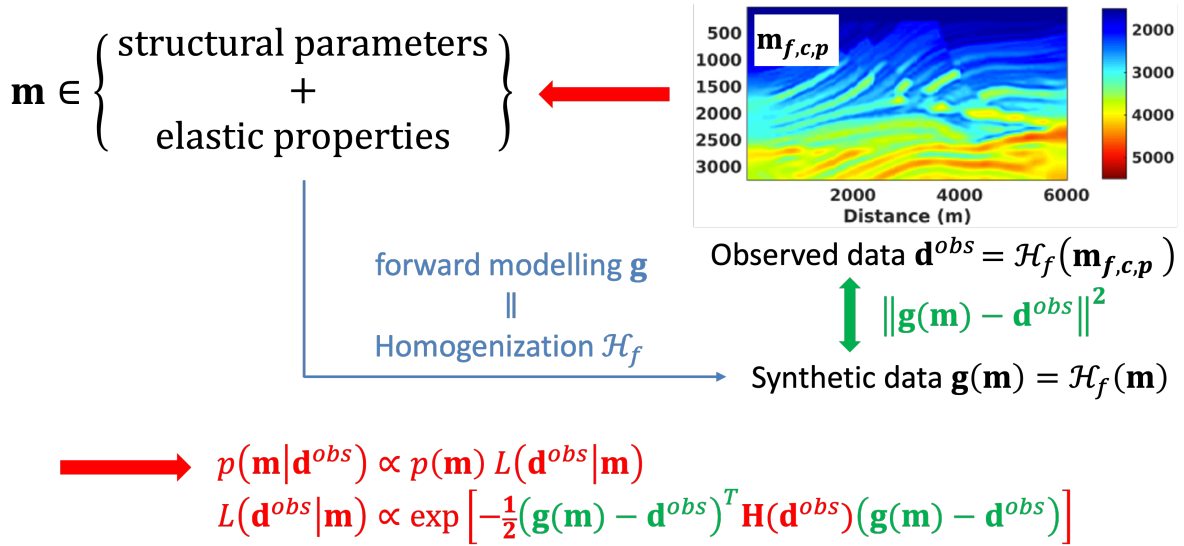


Figure 1.17: Principle of the inverse homogenization of a FWI image $\mathbf{m}_{f,c,p}$ to estimate structural parameters. Indices f,c,p point out the fact that the FWI image depends on the frequency content of the seismic data, on the source-receiver setting and on the parameterization of the inversion, respectively.

4. Bayes' theorem to get the posterior distribution of the structural parameters.

This method was initiated by Bodin et al. (2015) who applied it to the layered case. In the frame of the HIWAI project (ANR-16-CE31-0022-01), more complex synthetic cases were investigated (Capdeville and Métivier, 2018; Hedjazian et al., 2021) and applications to seismic anisotropy in the earth mantle were proposed (Alder et al., 2017; Magali et al., 2021). We intend to test the method at the exploration scale in order to reduce structural uncertainties derived from reflection seismic images. In a preliminary work (Irakarama et al., 2019), we showed how to rank structural interpretations using additional, VSP data. Implementing the inverse homogenization method presented here would allow to go further. To do so, we will benefit from stochastic structural modeling tools (e.g. Cherpeau and Caumon, 2015; Wellmann and Caumon, 2018) and the recent advance from Legentil et al. (2022) in mesh local updating (figure 1.18) to efficiently represent structural priors.

1.3.4 Perspective №2: Inverse homogenization of time-reversal focal spot

As mentioned in subsection 1.3.1, FWI can be set up to infer seismic source features such as location and moment tensor (Liu et al., 2004; Kim et al., 2011; Somala et al., 2018; Willacy et al., 2019). Other waveform-based methods are available for such a goal, including source scanning (e.g. Kao and Shan, 2004), match field processing (e.g. Cros et al., 2011), back-projection (e.g. Ishii et al., 2007) and time reversal (e.g. Li and van der Baan, 2016). They are often called source imaging methods because they all somehow focus waves to create an 'image' of the source. As compared to FWI, the main advantage of these methods is the absence of prior source model and the direct (i.e. non-iterative) calculation of the solution. However, they may require data processing to clean the input signals, with the notable exception of time reversal, which is weakly sensible to noise (e.g. Gajewski and Tessmer, 2005; Cao et al., 2012).

In collaboration with Zoé Renat and Yann Capdeville, I recently investigated the link between time reversal focal spot (which is smooth in space) and actual seismic source (which holds a discontinuity) in the light of point-source homogenization (Capdeville, 2021). This led to a paper submitted to *Geophysical Journal International* (Renat et al., 2022) whose content is given in chapter 3. Assuming a point-source model, we show that the wavefield propagated backward in time from a closed surface of receivers is equivalent to the sum of two low-wavenumber fields resulting from the homogenization of the original point-source (figure 1.19). In other words, the homogenized point-source is the equivalent

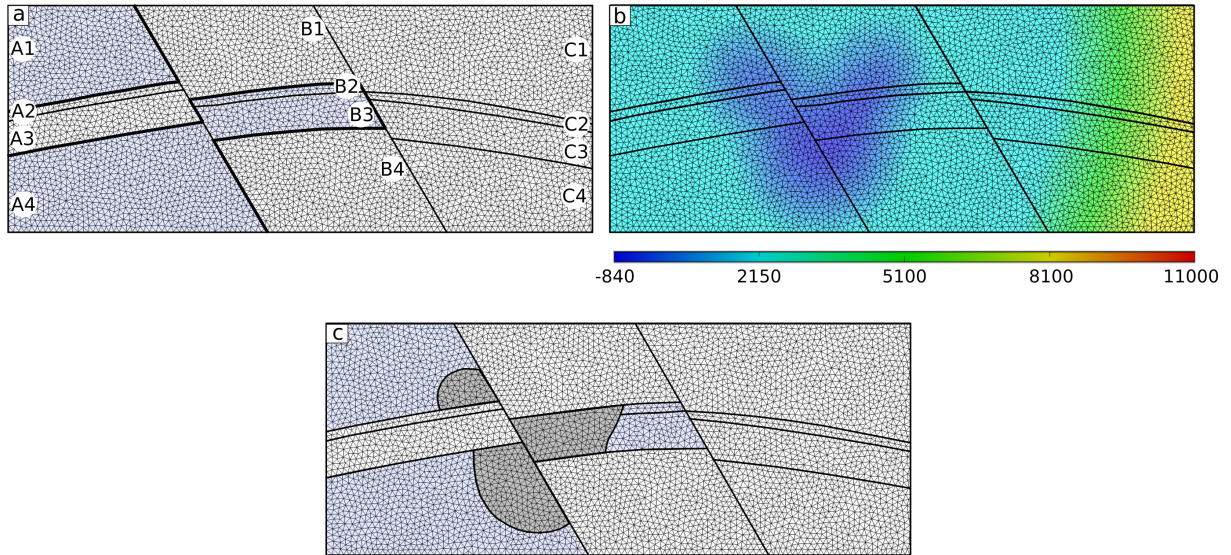


Figure 1.18: Insertion of a shape defined by an iso-value of a scalar field (b) in four regions of a geological model (a). (c) shows the resulting model for the scalar value 1000; only the four light blue regions (A1, B2, B3, A4) are remeshed. From Legentil et al. (2022).

force for producing the focal spot. This result opens the path to the downscaling of time reversal focal spots, following the inverse homogenization method described in the previous subsection (1.3.3).

1.4 Ambient noise correlations

By showing that a coherent signal can emerge from the correlation of seismic coda waves recorded at a pair of stations, Campillo and Paul (2003) opened a new chapter in the history of seismology. Following this work, cross-correlation was successfully applied to microseismic noise recordings (Shapiro and Campillo, 2004) so that seismic tomography with no earthquake suddenly became possible (Shapiro et al., 2005; Yao et al., 2006; Lin et al., 2007; Brenguier et al., 2007; Cho et al., 2007; Yang et al., 2007; Lin et al., 2008; Bensen et al., 2008). Microseismic noise is generated by the ocean activity in the 5-20 s period band, mainly (e.g. Longuet-Higgins, 1950; Friedrich et al., 1998; Stehly et al., 2006). Therefore, tomographic models derived from microseismic noise correlations primarily consist in short-period surface wave velocity maps. In the recent years, inversion of such maps in densely instrumented regions led to crustal and upper mantle models with unprecedented resolution (e.g. Lu et al., 2018; Nimiya et al., 2020). Moreover, body waves (e.g. Poli et al., 2012; Nakata et al., 2015) and higher frequency noise sources such as roads (Nakata et al., 2015) or railways (Brenguier et al., 2019) have been successfully explored.

1.4.1 Basics

Consider two stations A and B that are separated by a distance Δ . These stations respectively record signals $A(t)$ and $B(t)$ due to random noise sources acting in the medium. We denote by $A_p(t)$, $B_p(t)$ the signal received in A , B , respectively, from a single point-source P . The correlation between $A(t)$ and $B(t)$ can thus be written

$$C_{AB}(t) = \sum_{p,p'} \int A_p(\tau) B_{p'}(t + \tau) d\tau. \quad (1.19)$$

We assume that the signals produced by two distinct sources are uncorrelated. Therefore, the cross terms $p \neq p'$ in the double sum $\sum_{p,p'}$ vanish and expression (1.19) reduces to

$$C_{AB}(t) = \sum_p C_{AB}^p(t), \quad (1.20)$$

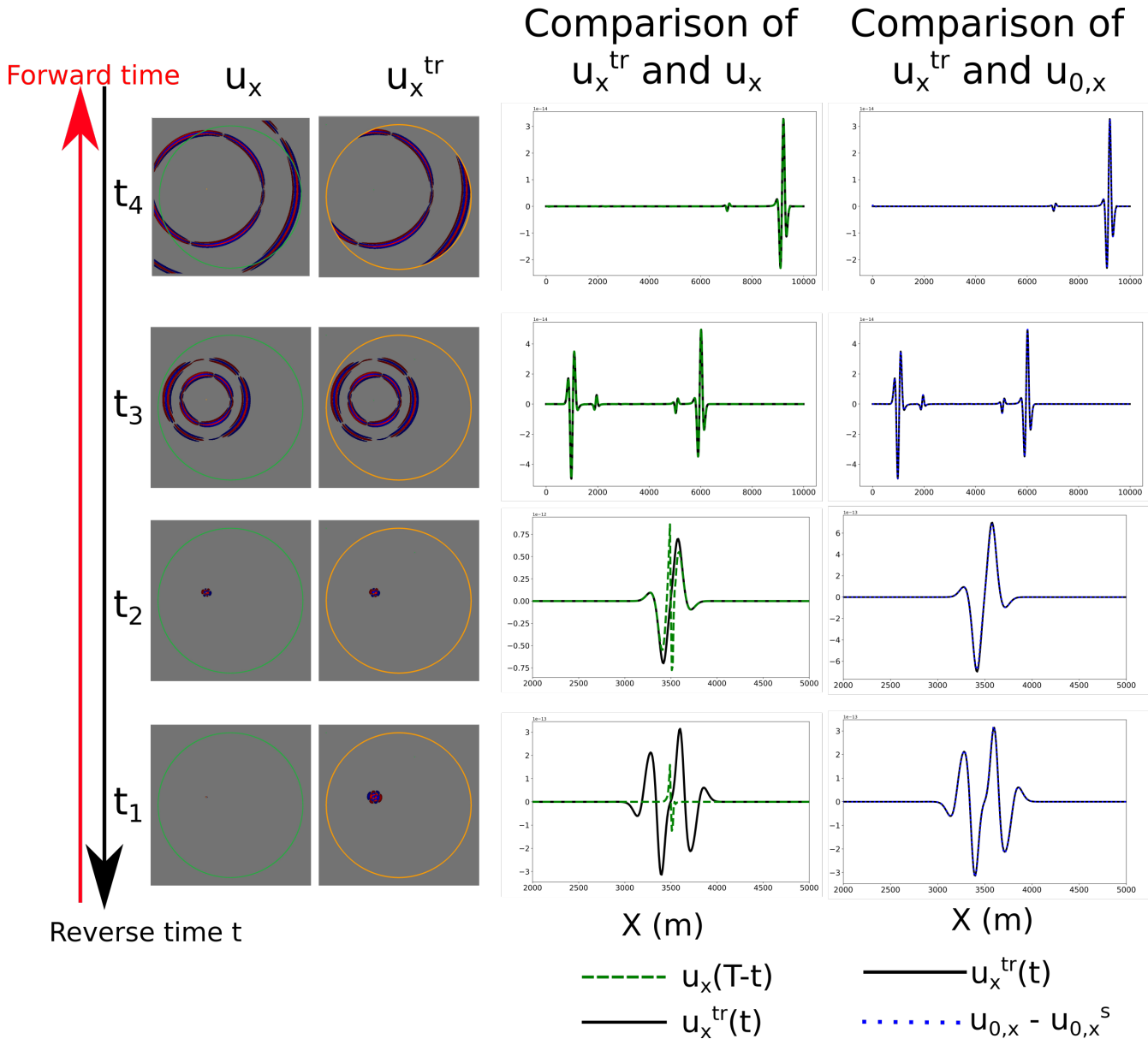


Figure 1.19: Comparison of the x-component of the backward wavefield \mathbf{u}^{tr} (black curve) to the forward wavefield \mathbf{u} (green dashed curve) in a 2D homogeneous medium. At the focal spot (e.g. at t_2 and t_1), the two wavefields are different. This is because the 'sink wavefield' \mathbf{u}^S is missing. When taking such a wavefield into account, i.e. comparing \mathbf{u}^{tr} to $\mathbf{u} - \mathbf{u}^S$, the match is perfect (not shown in the figure). We demonstrate that the zeroth-order forward and sink wavefields (\mathbf{u}_0 and \mathbf{u}_0^S , respectively) generated by a homogenized point-source are sufficient for matching \mathbf{u}^{tr} (right-hand side column), which prevents from representing the small scales of the source by a fine mesh. Such a result provides a new interpretation of the focal spot and an efficient forward way to compute it. Modified from Renat et al. (2022).

with

$$C_{AB}^p(t) = \int A_p(\tau) B_p(t + \tau) d\tau. \quad (1.21)$$

To study what can emerge from C_{AB} , let us place the noise sources and the stations on the surface of a layered medium, so that we can use analytical expressions. For the sake of simplicity, we work with fundamental mode surface waves, but overtones could be included by introducing a summation (e.g. Snieder, 2004; Halliday and Curtis, 2008). In the frequency domain, the vertical-vertical component of the fundamental mode surface wave Green's tensor between two points U and V is given by (Aki and Richards, 1980)

$$G_{UV}(\omega) = \frac{\exp\left[-i\left(kd_{UV} + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}}kd_{UV}}, \quad (1.22)$$

where ω is the angular frequency, $k(\omega)$ is the wave-number and d_{UV} is the horizontal distance between U and V . Using equation (1.22) and denoting by $|S_p(\omega)|^2$ the power spectral density of source P , the Fourier transform of cross-correlation (1.21) becomes

$$C_{AB}^p(\omega) = \frac{2}{\pi} |S_p(\omega)|^2 \frac{\exp[ik(d_{AP} - d_{BP})]}{k\sqrt{d_{AP}d_{BP}}}. \quad (1.23)$$

It follows that

$$C_{AB}(\omega) = \frac{2}{\pi} \sum_p |S_p(\omega)|^2 \frac{\exp[ik(d_{AP} - d_{BP})]}{k\sqrt{d_{AP}d_{BP}}}. \quad (1.24)$$

Replacing the summation over discrete sources by a surface integration (the variables with index p become functions of the source location \mathbf{r}), we obtain

$$C_{AB}(\omega) = \frac{2}{\pi} \iint |S(\mathbf{r}; \omega)|^2 \frac{\exp[ik(d_A(\mathbf{r}) - d_B(\mathbf{r}))]}{k\sqrt{d_A(\mathbf{r})d_B(\mathbf{r})}} d\mathbf{r}. \quad (1.25)$$

When $|S(\mathbf{r}; \omega)|^2$ is a smooth function of \mathbf{r} , the integral in equation (1.25) can be evaluated using the stationary phase approximation (Snieder, 2004). Let us use this approximation in a Cartesian coordinate system. We position receiver A at the origin and receiver B on the positive x axis. Then

$$C_{AB}(\omega) = \frac{2}{ik} \frac{\exp\left[i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}}k\Delta} \int_{\Delta}^{\infty} |S(x, y=0; \omega)|^2 dx - \frac{2}{ik} \frac{\exp\left[-i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}}k\Delta} \int_{-\infty}^0 |S(x, y=0; \omega)|^2 dx. \quad (1.26)$$

This result shows that the Green's function (GF) between A and B can emerge from the correlation of random noise records. The first term on the right-hand side contains the acausal GF and arises because of the noise sources at $x > \Delta$. The second term contains the causal GF and arises because of the sources at $x < 0$. Other demonstrations of this result include an analogy with time-reversal (Derode et al., 2003), the fluctuation-dissipation theorem (van Tiggelen, 2003), the reciprocity theorem (Wapenaar, 2004) and equipartition (Sánchez-Sesma and Campillo, 2006; Sánchez-Sesma et al., 2006).

Denoting by $v(\omega)$ the phase velocity of the fundamental mode Rayleigh wave and considering a uniform distribution of noise sources such that $|S(x, y=0; \omega)|^2 = |S(\omega)|^2$, equation (1.26) becomes

$$C_{AB}(\omega) = 2D \frac{v(\omega)|S(\omega)|^2}{i\omega} \left\{ \frac{\exp\left[i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}}k\Delta} - \frac{\exp\left[-i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}}k\Delta} \right\}. \quad (1.27)$$

We limit the integration over x to a range bounded by a finite distance $D \gg \Delta$. This is to prevent this integration to diverge. In practice, there is no problem of divergence because of intrinsic attenuation. Equation (1.27) shows that one has to differentiate the correlation in time to get the GF. Not only the waveform but also the amplitude decay of the GF with distance Δ is retrieved by the correlation.

1.4.2 Contributions

Numerical experiments for investigating the amplitude of noise correlations

Noise correlations depend on the spatial distribution of the noise sources (e.g. equation (1.25)). In other words, the signal which emerges from noise correlation corresponds to the GF when favorable conditions are met. Most of the theoretical demonstrations require a uniform distribution of sources (Lobkis and Weaver, 2001; Derode et al., 2003; van Tiggelen, 2003; Snieder, 2004; Wapenaar, 2004; Sánchez-Sesma and Campillo, 2006; Sánchez-Sesma et al., 2006). Although this condition is not satisfied in practice, noise correlations arrival-times were shown to be robust in many cases (e.g. Sabra et al., 2005; Harmon et al., 2008; Yang and Ritzwoller, 2008; Tsai, 2009) so that both group and phase velocity maps were established (e.g. Lin et al., 2008; Bensen et al., 2008; Yao and van der Hilst, 2009; Ekström et al., 2009). By contrast, the behavior of the amplitude looked more complicated (Larose et al., 2007; Gouédard et al., 2008; Prieto et al., 2009). In this context, I carried out a set of numerical experiments to better understand it (Cupillard and Capdeville, 2010). The experiments consist in computing synthetic noise recordings in a spherical Earth using a normal mode summation technique (e.g. Woodhouse and Girnius, 1982).

Figure 1.20 shows the result of the first experiment, which aims at retrieving what the aforementioned theories establish in the case of a uniform distribution of noise sources. In this experiment, three hundred sources are randomly positioned on the surface of the Earth. Each source generates a 24-hr random signal filtered between 66 and 200 s. An array of twelve receivers ($n = 0, \dots, 11$) records the wavefield produced by the noise sources. The correlations between the vertical displacement at station 0 and the vertical displacement at the other stations are performed. The Earth model is PREM (Dziewonski and Anderson, 1981). The configuration of the experiment is shown in figure 1.20a. It is repeated 5 120 times (64 processors perform it 80 times each) and all the obtained correlations are then stacked. Figure 1.20b compares the time-derivative of the correlation between stations 0 and 6 with the corresponding fundamental mode GF. As predicted by equation (1.27), the two curves match very well. We also plot the comparison between the amplitude decay of the correlation along the array and the amplitude decay of the GF. Again, the curves match very well, which confirms the theory.

Further numerical experiments were conducted to investigate the effect of non-uniform distributions of noise sources. Figure 1.21 illustrates one of these experiments. The sources are contained in a 50° -radius patch. Taking the same stations as those involved in figure 1.20, we computed the correlations using three different processing techniques commonly applied to noise records: raw, one-bit normalization and spectral whitening (e.g. Bensen et al., 2007). We observed that the amplitude decay of the fundamental mode GF is retrieved by the raw noise correlations. However, for the one-bit and the whitened noise correlations, this is no longer true: the amplitude decay corresponding to these two processes is less steep than the decay of the fundamental mode.

The overall conclusions of our paper are:

1. When the source distribution is uniform, both geometrical spreading and intrinsic attenuation are recovered by the correlations, regardless the processing applied to noise.
2. Intrinsic attenuation is always contained in raw noise correlations. As for geometrical spreading, it is dependent of the source distribution.
3. An amplitude decay is always observed in the one-bit and whitened noise correlations. In the case of non-uniform distributions of sources, this decay corresponds neither to geometrical spreading nor to intrinsic attenuation.

A theory for understanding the one-bit noise correlation

To enhance multiple scattering in coda records (e.g. Campillo and Paul, 2003; Larose et al., 2004) or to get rid off non-stationary signals like earthquakes or spikes in long noise records (e.g. Shapiro and Campillo, 2004; Shapiro et al., 2005; Bensen et al., 2007; Larose et al., 2007; Yao and van der Hilst, 2009), one-bit normalization is often applied prior to correlation. This process consists of retaining only the sign (+1 or -1) of each sample in the records. Interestingly, such a strong non-linear operation is able

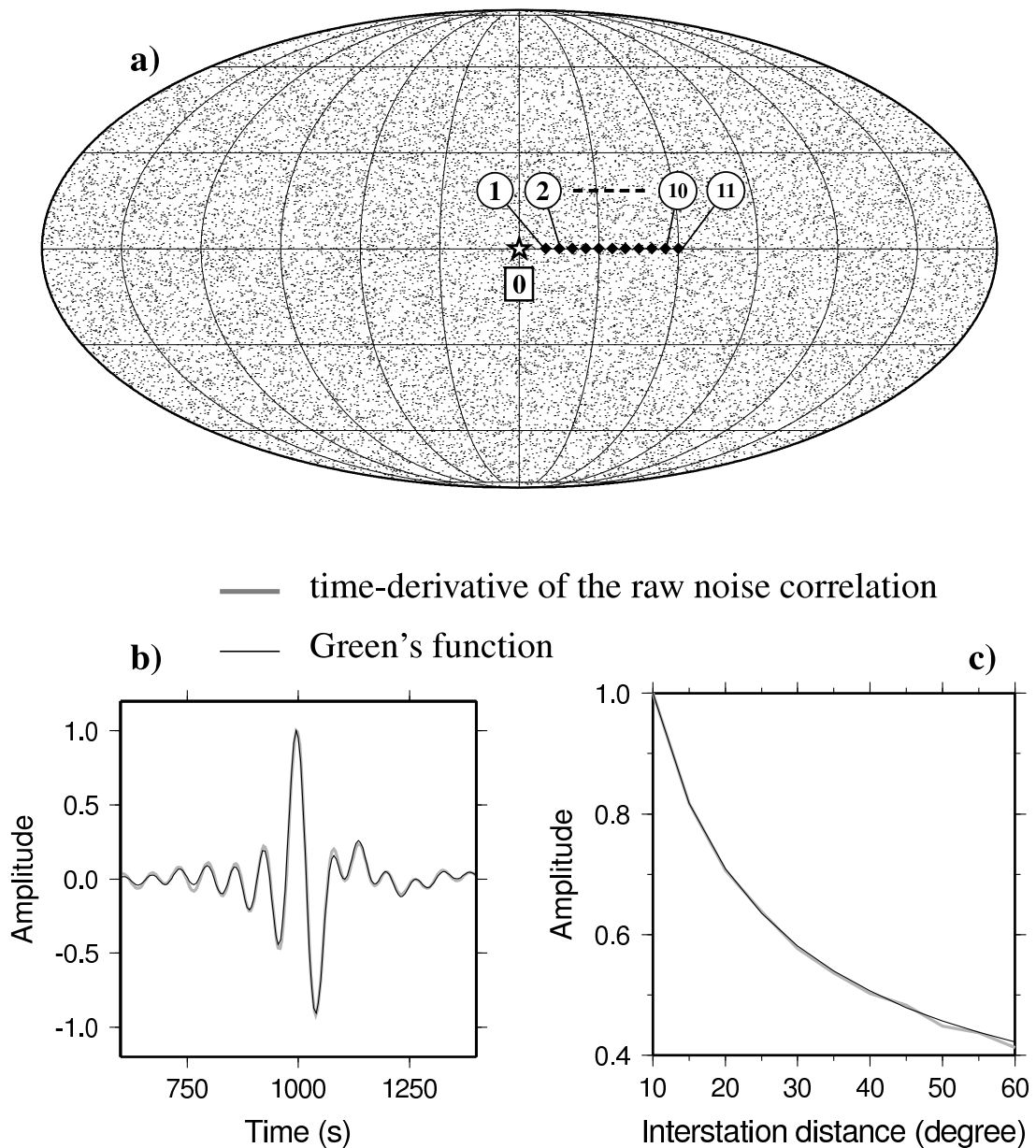


Figure 1.20: Noise simulations and correlations in the case of a uniform distribution of sources. a) Tiny pixels indicate the location of 24 000 noise sources generated by a single processor. All the sources are on the surface of the Earth. Diamonds represent eleven receivers ($n = 1, \dots, 11$). Two consecutive receivers are separated by 5° . The synthetic noise record from each station is correlated with the noise recorded at receiver 0 (white star). The distance between stations 0 and 6 is 10° . We plot the time-derivative of the correlation between stations 0 and 6 and the corresponding fundamental mode Rayleigh wave GF (b). We also compare the amplitude decay of the correlation with the amplitude decay of the GF (c). We see that the correlation fits both waveform and relative amplitude of the GF. Modified from Cupillard and Capdeville (2010).

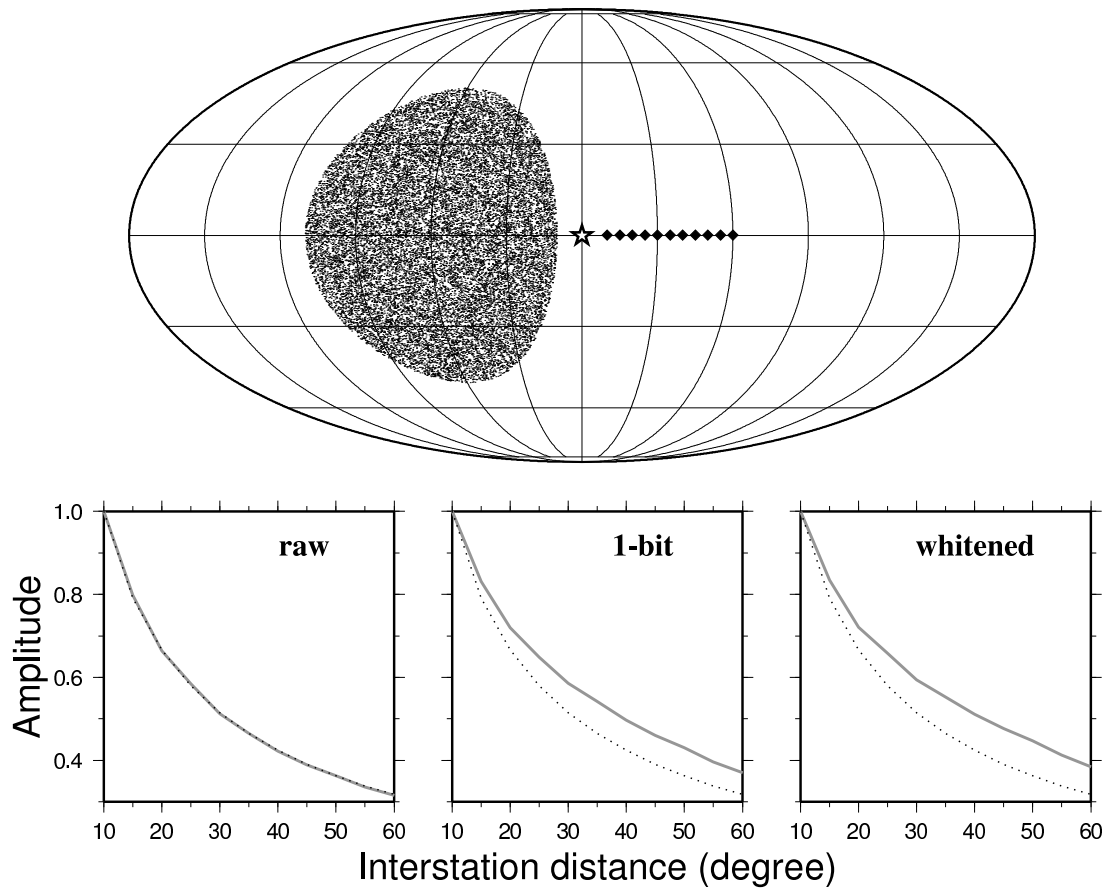


Figure 1.21: Noise simulations and correlations in the case of a non-uniform distribution of sources. Top: The sources belong to a 50° -radius patch centered on the equator at longitude -60° . The stations are the same as those involved in figure 1.20. Bottom: Comparison of the amplitude decay of the correlation (gray line) with the amplitude decay of the GF (dotted line) for three different noise processing techniques (raw, one-bit normalization and whitening). Modified from Cupillard and Capdeville (2010).

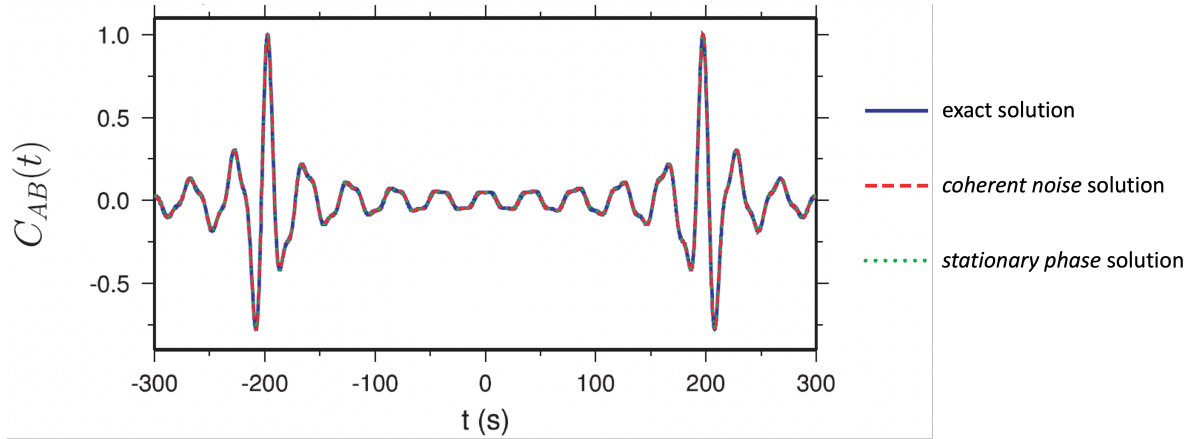


Figure 1.22: Plot of the raw correlation $C_{AB}(t)$ using three different equations. The exact solution (equation (1.25)) is in blue, the solution involving the coherent noise is in dashed red and the solution given by the stationary phase approximation (equation (1.27)) is in dotted green. The three curves match very well.

to preserve some information on the amplitude (e.g. Cupillard and Capdeville, 2010). To understand how that can work, I developed a theory based on the concept of coherent and incoherent noise (Cupillard et al., 2011).

For each lag time t in the correlation $C_{AB}(t)$ between two raw noise records $A(\tau)$ and $B(\tau)$, we first show that there exist two Gaussian signals $A^t(\tau)$ and $B^t(\tau)$ with zero mean and variance $\sigma_{A^t}^2$ and $\sigma_{B^t}^2$ respectively, such that $A^t(\tau)$ and $B^t(t + \tau)$ are perfectly correlated and give rise to $C_{AB}(t)$. In other words,

$$C_{AB}(t) = \int A^t(\tau)B^t(t + \tau)d\tau = \sigma_{A^t}\sigma_{B^t}. \quad (1.28)$$

$A^t(\tau)$ and $B^t(\tau)$ are called *coherent noise at lag time t* . We also define *incoherent noise at lag time t* :

$$\overline{A^t}(\tau) = A(\tau) - A^t(\tau) \quad \text{and} \quad \overline{B^t}(\tau) = B(\tau) - B^t(\tau), \quad (1.29)$$

with variances $\sigma_{\overline{A^t}}^2$ and $\sigma_{\overline{B^t}}^2$, respectively.

Equation (1.28) has been demonstrated in the case of a uniform distribution of noise sources, but one can reasonably think that coherent and incoherent noise also exist in the case of other distributions. Moreover, equation (1.28) holds for broad band sources. In the case of narrow band sources, it is slightly different. Figure 1.22 shows a comparison of the expression obtained in this last case (equation 46 in Cupillard et al. (2011)) with the exact solution (1.25) and the approximation (1.27). For the calculation, we take the real part of the expressions and $|S(\omega)|$ as a boxcar function $H(\omega_0 + \frac{\Delta\omega}{2}) - H(\omega_0 - \frac{\Delta\omega}{2})$, where H is the Heaviside step function. In this case, we can show that

$$\sigma_{A^t}\sigma_{B^t} = 2Dv \left| \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{1}{\omega} \left[\frac{\cos(\omega t + k\Delta - \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} + \frac{\cos(\omega t - k\Delta + \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} \right] d\omega \right| \quad (1.30)$$

Taking $\Delta\omega = \omega_0 = 0.1\pi \text{ rad.s}^{-1}$, $v = 3 \text{ km.s}^{-1}$ and $\Delta = 600 \text{ km}$, we obtain the curves in figure 1.22. They match very well, which illustrates the validity of our theory.

Based on our definitions of coherent and incoherent noise, we derive a general expression for the one-bit noise correlation:

$$|C_{AB}^{ob}(t)| = n \left[1 - \frac{2}{\pi} \arctan\left(\frac{\sigma_{\overline{A^t}}}{\sigma_{A^t}}\right) \right] \left[1 - \frac{2}{\pi} \arctan\left(\frac{\sigma_{\overline{B^t}}}{\sigma_{B^t}}\right) \right], \quad (1.31)$$

where n is the number of samples in the correlation. The involved physical parameters are the standard deviations of coherent and incoherent noise. More precisely, the ratio $w_{R^t} = \sigma_{\overline{R^t}}/\sigma_{R^t}$ at each receiver R

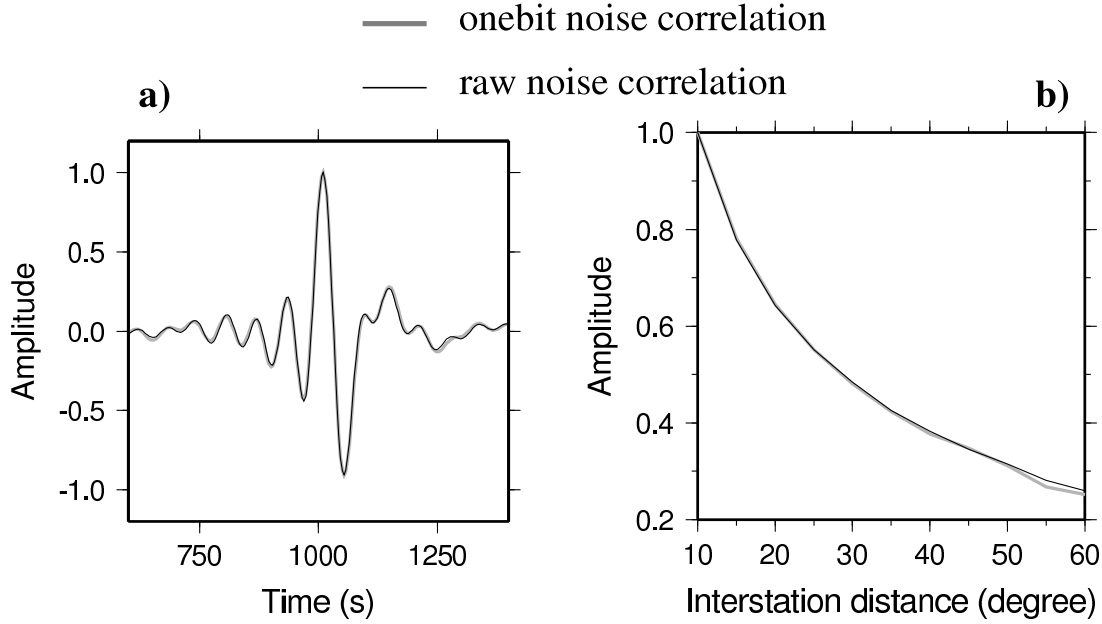


Figure 1.23: Comparison between raw noise correlations and one-bit noise correlations obtained by numerical simulation in an anelastic Earth (cf figure 1.20a). The time-derivatives of the two waveforms from the pair of stations 0 and 6 (a) and the amplitude decays along the array of receivers (b) are plotted.

is the argument of an inverse tangent function. At a given time t , if there is no coherent noise at one of the receivers, then no signal emerges from the correlation: $C_{AB}^{ob}(t) = 0$ because $w_{R'}$ tends to infinity so $1 - \frac{2}{\pi} \arctan(w_{R'}) = 0$. On the contrary, if the coherent energy is large with respect to the incoherent energy at both receivers, then most of the n samples contribute to the correlation so $|C_{AB}^{ob}(t)|$ is large.

In the case of a uniform distribution of noise sources, equation (1.31) simplifies and we show that

$$|C_{AB}^{ob}(t)| = \left(\frac{2}{\pi\sigma}\right)^2 |C_{AB}(t)|, \quad (1.32)$$

where σ^2 is the autocorrelation peak at any point in the surface of the medium. This equation sets the equality between the raw and the one-bit noise correlations in the case of a uniform distribution of noise sources. It is valid in both purely elastic and anelastic media. Obviously, the absolute amplitudes of the two correlations are not the same, but the waveforms and the relative amplitudes are identical. This is confirmed by the numerical experiment presented in figure 1.20a). Looking at the raw noise correlation and the one-bit noise correlation from the pair of stations 0 and 6, a good match is observed (figure 1.23a). The amplitude decays along the array of receivers also match very well (figure 1.23b).

1.4.3 Perspective: Modeling full waveforms of ambient noise correlations

Studying the behavior of the amplitude of noise correlations allowed me to consider a method to simulate the full waveform of these signals (Cupillard, 2008). The method is based on the time reversal of the amplitude spectrum of the noise data. This enables to build a virtual source capable of radiating the anisotropic flux of the seismic noise which crosses a given station. Indeed, the amplitude of the correlations holds this flux, so it is necessary to take it into account to correctly simulate the full waveforms. Figure 1.24 illustrates the propagation of the wavefield generated by the virtual source associated with the AIGLE station in the Swiss Alps. Figure 1.25 shows the result of this propagation: at each station in the network, we find the waveform corresponding to the correlation between AIGLE and the station.

Although very promising, this method needs to be improved in the light of significant advances in our understanding of noise correlation waveforms within the last decade (Curtis and Halliday, 2010; Tromp et al., 2010; Fichtner, 2014, 2015; Fichtner et al., 2016; Sager et al., 2018). Building proper virtual

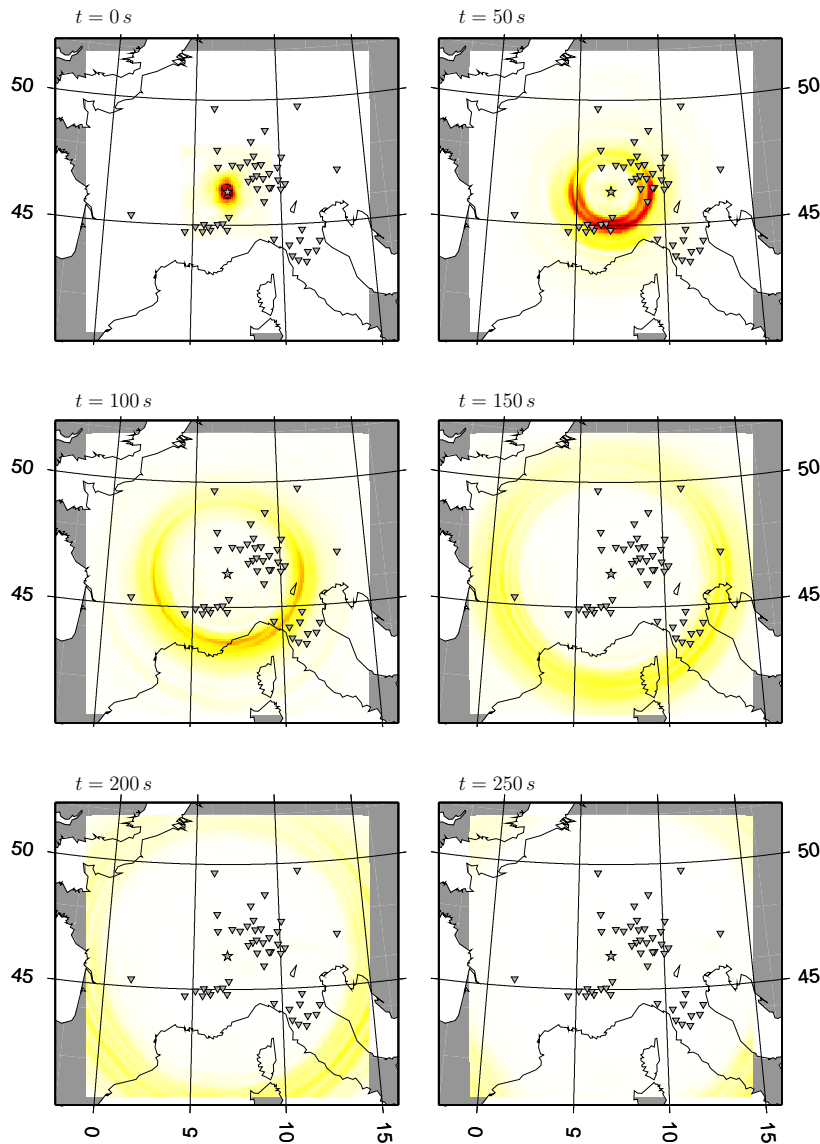


Figure 1.24: SEM simulation of the wavefield generated by the virtual source associated to station AIGLE (star). The L^2 -norm of the displacement at the surface is represented. Triangles mark the other stations of the network.

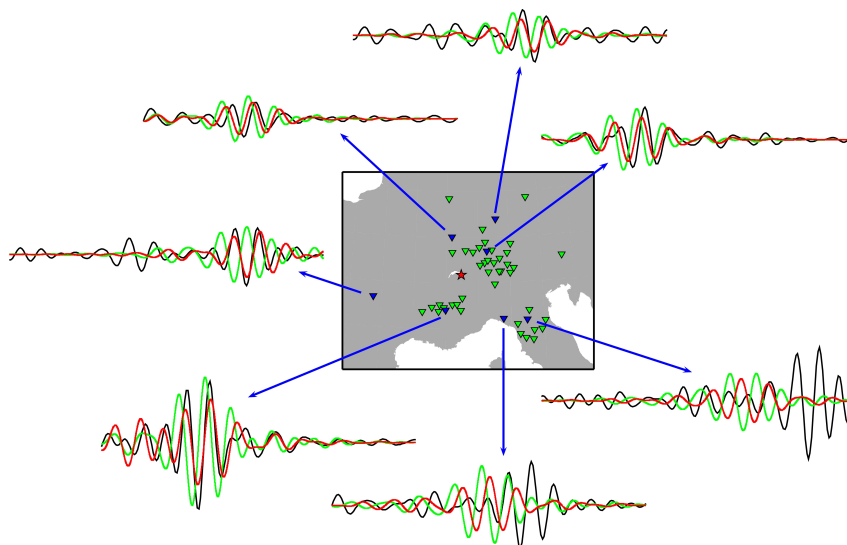


Figure 1.25: Seismograms obtained from the simulation presented in figure 1.24. The red star is station AIGLE. Green waveforms are obtained in a 1D model (PREM, Dziewonski and Anderson, 1981); red waveforms are obtained in a 3D model of the lithosphere (CUB, Shapiro and Ritzwoller, 2002); black waveforms are the data (i.e. noise correlations).

sources would open the path to the modeling and inversion of correlation full waveforms and the overcoming of current limitations in ambient noise tomography such as the great-circle path approximation and the use of phase information only.

References

- Aki, K. and P. Richards (1980). *Quantitative Seismology: Theory and Methods*. Freeman, San Francisco.
- Alder, C., T. Bodin, Y. Ricard, Y. Capdeville, E. Debayle, and J. P. Montagner (2017). Quantifying seismic anisotropy induced by small-scale chemical heterogeneities. *Geophys. J. Int.* 211(3), 1585–1600.
- Allaire, G. (1992). Homogenization and two-scale convergence. *SIAM J. Math. Anal.* 23, 1482–1518.
- Allaire, G. and Y. Capdeboscq (2000). Homogenization of a spectral problem in neutronic multigroup diffusion. *Comp. Methods Appl. Mech. Engrg.* 187(1-2), 91–117.
- Allaire, G. and Z. Habibi (2013). Homogenization of a conductive, convective, and radiative heat transfer problem in a heterogeneous domain. *SIAM J. Math. Anal.* 45(3), 1136–1178.
- Allègre, C., J.-L. Le Mouél, and A. Provost (1982). Scaling rules in rock fracture and possible implications for earthquake prediction. *Nature* 297(5861), 47–49.
- Alterman, Z. S. and F. C. Karal (1968). Propagation of elastic waves in layered media by finite-difference method. *Bull. Seism. Soc. Am.* 58, 367–398.
- Aminzadeh, F., J. Brac, and T. Kunz (1997). *3-D Salt and Overthrust models*. SEG/EAGE 3-D Modeling Series No.1.
- Anquez, P., N. Glinsky, P. Cupillard, and G. Caumon (2021). Impacts of geometric model simplifications on wave propagation - Application to ground motion simulation in the lower Var valley basin (France). *Geophys. J. Int.* 229(1), 110–137.
- Anquez, P., J. Pellerin, M. Irakarama, P. Cupillard, B. Lévy, and G. Caumon (2019). Automatic correction and simplification of geological maps and cross-sections for numerical simulations. *Comptes Rendus Geoscience* 351(1), 48–58.
- Bacigalupo, A. and L. Gambarotta (2014). Second-gradient homogenized model for wave propagation in heterogeneous periodic media. *Int. J. Solids Struct.* 51(5), 1052–1065.
- Backer, G. (1976). Error estimates for the finite element method for second order hyperbolic equations. *SIAM J. Num. Anal.* 13, 564–575.
- Bard, P.-Y. (1997). Local effects on strong ground motion: Basic physical phenomena and estimation methods for microzoning studies. *Advanced study course on Seismic Risk (SERINA), Thessaloniki, Greece*, 229–299.
- Baudouin, T., J.-F. Remacle, E. Marchandise, F. Henrotte, and C. Geuzaine (2014). A frontal approach to hex-dominant mesh generation. *Adv. Model. Simul. Eng. Sci.* 1(1), 8.
- Beller, S., V. Monteiller, L. Combe, S. Operto, and G. Nolet (2017). On the sensitivity of teleseismic full-waveform inversion to earth parametrization, initial model and acquisition design. *Geophys. J. Int.* 212(2), 1344–1368.
- Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophys. J. Int.* 169, 1239–1260.

- Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2008). Broad-band ambient noise surface wave tomography across the United States. *J. Geophys. Res.* 113, B05306.
- Bensoussan, A., J.-L. Lions, and G. Papanicolaou (1978). *Asymptotic Analysis for Periodic Structures*. Amsterdam: North-Holland.
- Bielak, J., O. Ghattas, and E. Kim (2005). Parallel octree-based finite element method for large-scale earthquake ground motion simulation. *Computer Modeling in Engineering and Sciences* 10(2), 99.
- Bielak, J., R. C. MacCamy, D. S. McGhee, and A. Barry (1991). Unified symmetric BEM-FEM for site effects on ground motion - SH waves. *Journal of Engineering Mechanics* 117(10), 2265–2285.
- Bodin, T., Y. Capdeville, B. Romanowicz, and J.-P. Montagner (2015). Interpreting radial anisotropy in global and regional tomographic models. In *The Earth's Heterogeneous Mantle*, pp. 105–144. Springer.
- Bonneau, F., V. Henrion, G. Caumon, P. Renard, and J. Sausse (2013). A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Computers & Geosciences* 56, 12–22.
- Bonnet, E., O. Bour, N. E. Odling, P. Davy, I. Main, P. Cowie, and B. Berkowitz (2001). Scaling of fracture systems in geological media. *Reviews of geophysics* 39(3), 347–383.
- Boore, D. M. (1972). Finite-difference method for seismic wave propagation in heterogeneous materials. In B. A. Bolt (Ed.), *Methods in Computational Physics*, Volume 11, pp. 1–37. Academic Press, New York.
- Botella, A., B. Lévy, and G. Caumon (2016). Indirect unstructured hex-dominant mesh generation using tetrahedra recombination. *Computational Geosciences* 20(3), 437–451.
- Bouchon, M., M. Campillo, and S. Gaffet (1989). A boundary integral equation-discrete wavenumber representation method to study wave propagation in multilayered media having irregular interfaces. *Geophysics* 54(9), 1134–1140.
- Bour, O. and P. Davy (1997). Connectivity of random fault networks following a power law fault length distribution. *Water Resources Research* 33(7), 1567–1583.
- Bour, O., P. Davy, C. Darcel, and N. Odling (2002). A statistical scaling model for fracture network geometry, with validation on a multiscale mapping of a joint network (hornelen basin, norway). *Journal of Geophysical Research: Solid Earth* 107(B6), ETG–4.
- Boutin, C. and J. L. Auriault (1993). Rayleigh scattering in elastic composite materials. *Int. J. Engrg. Sci.* 31(12), 1669–1689.
- Bozdağ, E., D. Peter, M. Lefebvre, D. Komatitsch, J. Tromp, J. Hill, N. Podhorszki, and D. Pugmire (2016). Global adjoint tomography: first-generation model. *Geophys. J. Int.* 207(3), 1739–1766.
- Bozdağ, E., J. Trampert, and J. Tromp (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophys. J. Int.* 185(2), 845–870.
- Brenguier, F., P. Boué, Y. Ben-Zion, F. Vernon, C. Johnson, A. Mordret, O. Coutant, P.-E. Share, E. Beaucé, D. Hollis, and T. Lecocq (2019). Train traffic as a powerful noise source for monitoring active faults with seismic interferometry. *Geophys. Res. Lett.* 46(16), 9529–9536.
- Brenguier, F., N. M. Shapiro, M. Campillo, A. Nercessian, and V. Ferrazzini (2007). 3D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlation. *Geophys. Res. Lett.* 34, L02305.

- Brossier, R., S. Operto, and J. Virieux (2009). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics* 74(6), WCC105–WCC118.
- Brossier, R., S. Operto, and J. Virieux (2015). Velocity model building from seismic reflection data by Full Waveform Inversion. *Geophys. Prosp.* 63(2), 354–367.
- Budiansky, B. and R. J. O’connell (1976). Elastic moduli of a cracked solid. *International journal of Solids and structures* 12(2), 81–97.
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995). Multiscale seismic waveform inversion. *Geophysics* 60(5), 1457–1473.
- Campillo, M. and M. Bouchon (1985). Synthetic SH seismograms in a laterally varying medium by the discrete wavenumber method. *Geophys. J. Int.* 83(1), 307–317.
- Campillo, M. and A. Paul (2003). Long-range correlations in the diffuse seismic coda. *Science* 299, 547–549.
- Cao, W., S. M. Hanafy, G. T. Schuster, G. Zhan, and C. Boonyasirawat (2012). High-resolution and super stacking of time-reversal mirrors in locating seismic sources. *Geophys. Prospect.* 60(1), 1–17.
- Capdeville, Y. (2005). An efficient Born normal mode method to compute sensitivity kernels and synthetic seismograms in the Earth. *Geophys. J. Int.* 163(2), 639–646.
- Capdeville, Y. (2021). Homogenization of seismic point and extended sources. *Geophys. J. Int.* 226(2), 1390–1416.
- Capdeville, Y., E. Chaljub, J. P. Vilotte, and J. P. Montagner (2003). Coupling the Spectral Element Method with a modal solution for elastic wave propagation in global Earth models. *Geophys. J. Int.* 152, 34–66.
- Capdeville, Y., P. Cupillard, and S. Singh (2020). Chapter six - an introduction to the two-scale homogenization method for seismology. In B. Moseley and L. Krischer (Eds.), *Machine Learning in Geosciences*, Volume 61 of *Advances in Geophysics*, pp. 217–306. Elsevier.
- Capdeville, Y., L. Guillot, and J. Marigo (2010a). 1-D non-periodic homogenization for the seismic wave equation. *Geophys. J. Int.* 181, 907–910.
- Capdeville, Y., L. Guillot, and J. Marigo (2010b). 2-D non-periodic homogenization to upscale elastic media for P-SV waves. *Geophys. J. Int.* 182, 903–922.
- Capdeville, Y. and J. Marigo (2007). Second order homogenization of the elastic wave equation for non-periodic layered media. *Geophys. J. Int.* 170, 823–838.
- Capdeville, Y. and J. Marigo (2008). Shallow layer correction for spectral element like methods. *Geophys. J. Int.* 172, 1135–1150.
- Capdeville, Y. and L. Métivier (2018). Elastic full waveform inversion based on the homogenization method: theoretical framework and 2-D numerical illustrations. *Geophys. J. Int.* 213(2), 1093–1112.
- Capdeville, Y., M. Zhao, and P. Cupillard (2015). Fast Fourier homogenization for elastic wave propagation in complex media. *Wave Motion* 54, 170–186.
- Casadei, F., E. Gabellini, F. Fotia, F. Maggio, and A. Quarteroni (2002). A mortar spectral/finite element method for complex 2D and 3D elastodynamic problems. *Computer Methods in Applied Mechanics and Engineering* 191, 5119–5148.

- Castañeda, P. and J. Willis (1995). The effect of spatial distribution on the effective behavior of composite materials and cracked media. *Journal of the Mechanics and Physics of Solids* 43(12), 1919–1951.
- Caumon, G., P. Collon, C. Le Carlier de Veslud, J. Sausse, and S. Viseur (2009). Surface-based 3D modeling of geological structures. *Mathematical Geosciences* 41(8), 927–945.
- Cerveny, V. . (2001). *Seismic ray theory*. Cambridge University Press, Cambridge, UK.
- Chaillat, S., M. Bonnet, and J.-F. Semblat (2009). A new fast multi-domain bem to model seismic wave propagation and amplification in 3-d geological structures. *Geophys. J. Int.* 177(2), 509–531.
- Chaljub, E. (2000). *Modélisation numérique de la propagation d’ondes sismiques à l’échelle du globe*. Ph. D. thesis, Université Paris 7.
- Chaljub, E., Y. Capdeville, and J. Vilotte (2003). Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids. *J. Comp. Physics* 183, 457–491.
- Chaljub, E., P. Moczo, S. Tsuno, P.-Y. Bard, J. Kristek, M. Käser, M. Stupazzini, and M. Kristekova (2010). Quantitative comparison of four numerical predictions of 3D ground motion in the Grenoble valley, France. *Bull. Seism. Soc. Am.* 100(4), 1427–1455.
- Chaljub, E. and B. Valette (2004). Spectral element modelling of three-dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core. *Geophys. J. Int.* 183, 131–141.
- Cherpeau, N. and G. Caumon (2015). Stochastic structural modelling in sparse data situations. *Petroleum Geoscience* 21(4), 233–247.
- Cherpeau, N., G. Caumon, J. Caers, and B. Lévy (2012). Method for stochastic inverse modeling of fault geometry and connectivity using flow data. *Mathematical Geosciences* 44(2), 147–168.
- Cherpeau, N., G. Caumon, and B. Lévy (2010). Stochastic simulations of fault networks in 3d structural modeling. *Comptes Rendus Geoscience* 342(9), 687–694.
- Cho, K., R. B. Herrmann, C. J. Ammon, and K. Lee (2007). Imaging the upper crust of the Korean peninsula by surface-wave tomography. *Bull. Seism. Soc. Am.* 67, 198–207.
- Cho, Y., R. L. Gibson Jr, J. Lee, and C. Shin (2019). Linear-slip discrete fracture network model and multiscale seismic wave simulation. *Journal of Applied Geophysics* 164, 140–152.
- Courant, R., K. Friedrichs, and H. Lewy (1928). Über die partiellen Differenzgleichungen der mathematischen Physik. *Mathematische Annalen* 100, 32–74.
- Cros, E., P. Roux, J. Vandemeulebrouck, and S. Kedar (2011). Locating hydrothermal acoustic sources at Old Faithful Geyser using Matched Field Processing. *Geophys. J. Int.* 187(1), 385–393.
- Cupillard, P. (2008). *Simulation par la méthode des éléments spectraux des formes d’onde obtenues par corrélation de bruit sismique*. Ph. D. thesis, Université Paris 7.
- Cupillard, P. and Y. Capdeville (2010). On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach. *Geophys. J. Int.* 181, 1687–1700.
- Cupillard, P. and Y. Capdeville (2017). Performance and convergence of the non-periodic homogenization for the 3D elastic wave equation. In *Proc. 79th EAGE Conference and Exhibition*. European Association of Geoscientists & Engineers.
- Cupillard, P. and Y. Capdeville (2018). Non-periodic homogenization of 3-D elastic media for the seismic wave equation. *Geophys. J. Int.* 213(2), 983–1001.

- Cupillard, P., Y. Capdeville, and A. Botella (2015). Homogenization of 3D geological models for seismic wave propagation. In *SEG Technical Program Expanded Abstracts*, New Orleans, pp. 3656–3660.
- Cupillard, P., E. Delavaud, G. Burgos, G. Festa, J.-P. Vilotte, Y. Capdeville, and J.-P. Montagner (2012). RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale. *Geophys. J. Int.* 188, 1203–1220.
- Cupillard, P., W. Mulder, P. Anquez, A. Mazuyer, and J.-F. Barthélémy (2020). The apparent anisotropy of the SEG-EAGE overthrust model". In *Proc. 82nd EAGE Conference and Exhibition*. European Association of Geoscientists & Engineers.
- Cupillard, P., L. Stehly, and B. Romanowicz (2011). The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise. *Geophys. J. Int.* 183, 1397–1414.
- Curtis, A., B. Dost, J. Trampert, and R. Snieder (1998). Eurasian fundamental mode surface wave phase velocities and their relationship to tectonic structures. *J. Geophys. Res.* 103, 26,919–26,947.
- Curtis, A. and D. Halliday (2010). Directional balancing for seismic and general wavefield interferometry. *Geophysics* 75(1), SA1–SA14.
- Delavaud, E. (2007). *Simulation numérique de la propagation d'ondes en milieu géologique complexe : application à l'évaluation de la réponse sismique du bassin de Caracas (Vénézuéla)*. Ph. D. thesis, Université Paris 7.
- Delavaud, E., P. Cupillard, G. Festa, and J. P. Vilotte (2006). 3D Spectral Element Method simulations of the seismic response in the Caracas basin. In *Proc. of the Third International Symposium on the Effects of Surface Geology on Seismic Motion*, Volume 1, Grenoble, France, pp. 515–522.
- Derode, A., E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink (2003). Recovering the Green's function from field-field correlations in an open scattering medium (L). *J. Acoust. Soc. Am.* 113, 2973–2976.
- Dessa, J. X., S. Operto, S. Kodaira, A. Nakanishi, G. Pascal, J. Virieux, and Y. Kaneda (2004). Multiscale seismic imaging of the eastern nankai trough by full waveform inversion. *Geophys. Res. Lett.* 31, L18606.
- Dewaide, L., J. M. Baele, P. Collon-Drouaillet, Y. Quinif, G. Rochez, S. Vandycke, and V. Hallet (2014). Karstification in dolomitized Waulsortian mudmounds (Belgium). *Geologica Belgica* 17(1), 43–51.
- Diaz, J. and P. Joly (2005). Robust high order non-conforming finite element formulation for time domain fluid-structure interaction. *Journal of computational acoustics* 13(03), 403–431.
- Douglas, J. (2003). Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews* 61(1), 43–104.
- Dumbser, M. and M. Käser (2006). An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - II. The three-dimensional isotropic case. *Geophys. J. Int.* 167(1), 319–336.
- Dumbser, M., M. Käser, and E. F. Toro (2007). An arbitrary high-order Discontinuous Galerkin method for elastic waves on unstructured meshes - V. Local time stepping and p-adaptivity. *Geophys. J. Int.* 171(2), 695–717.
- Dumontet, H. (1990). *Homogénéisation et effets de bords dans les matériaux composites*. Ph. D. thesis, Université Paris 6.

- Dziewonski, A. M. and D. L. Anderson (1981). Preliminary reference Earth model. *Phys. Earth Planet. Inter.* 25, 297–356.
- Ekström, G., G. A. Abers, and S. C. Webb (2009). Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation. *Geophys. Res. Lett.* 36(18), L18301.
- Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proc. R. Soc. Lond. A* 241, 376–396.
- Etienne, V., E. Chaljub, J. Virieux, and N. Glinsky (2010). An hp-adaptive discontinuous Galerkin finite-element method for 3-D elastic wave modelling. *Geophys. J. Int.* 183, 941–962.
- Faccioli, E., F. Maggio, R. Paolucci, and A. Quarteroni (1997). 2D and 3D elastic wave propagation by a pseudospectral domain decomposition method. *J. Seismology* 1, 237–251.
- Ferreira, A. M. G., J. H. Woodhouse, K. Visser, and J. Trampert (2010). On the robustness of global radially anisotropic surface wave tomography. *J. Geophys. Res.* 115, B04313, doi:10.1029/2009JB006716.
- Festa, G. and J. P. Vilotte (2005). The Newmark scheme as velocity-stress time-staggering : an efficient PML for spectral element simulations of elastodynamics. *Geophys. J. Int.* 161(3), 789–812.
- Fichtner, A. (2014). Source and processing effects on noise correlations. *Geophysical Journal International* 197(3), 1527–1531.
- Fichtner, A. (2015). Source-structure trade-offs in ambient noise correlations. *Geophysical Journal International* 202(1), 678–694.
- Fichtner, A., H.-P. Bunge, and H. Igel (2006a). The adjoint method in seismology: I - Theory. *Phys. Earth Planet. Inter.* 157, 86–104.
- Fichtner, A., H.-P. Bunge, and H. Igel (2006b). The adjoint method in seismology: II - Applications: traveltimes and sensitivity functionals. *Phys. Earth Planet. Inter.* 157, 105–123.
- Fichtner, A., B. L. N. Kennett, H. Igel, and H.-P. Bunge (2008). Theoretical background for continental and global scale full-waveform inversion in the time-frequency domain. *Geophys. J. Int.* 175(8), 665–685.
- Fichtner, A., B. L. N. Kennett, H. Igel, and H.-P. Bunge (2009). Full waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophys. J. Int.* 179(3), 1703–1725.
- Fichtner, A., E. Saygin, T. Taymaz, P. Cupillard, Y. Capdeville, and J. Trampert (2013). The deep structure of the North Anatolian Fault Zone. *Earth Planet. Sci. Lett.* 373, 109–117.
- Fichtner, A. and S. Simutè (2018). Hamiltonian Monte Carlo Inversion of seismic sources in complex media. *J. Geophys. Res.* 123(4), 2984–2999.
- Fichtner, A., L. Stehly, L. Ermert, and C. Boehm (2016). Generalized interferometry – I: theory for interstation correlations. *Geophys. J. Int.* 208(2), 603–638.
- Fichtner, A., J. Trampert, P. Cupillard, E. Saygin, T. Taymaz, Y. Capdeville, and A. Villasenor (2013). Multi-scale full waveform inversion. *Geophys. J. Int.* 194, 534–556.
- Fichtner, A., D.-P. van Herwaarden, M. Afanasiev, S. Simutè, L. Krischer, Y. Çubuk Sabuncu, T. Taymaz, L. Colli, E. Saygin, A. Villaseñor, J. Trampert, P. Cupillard, H.-P. Bunge, and H. Igel (2018). The Collaborative Seismic Earth Model: Generation 1. *Geophys. Res. Lett.* 45(9), 4007–4016.
- Fichtner, A. and A. Zunino (2019). Hamiltonian nullspace shuttles. *Geophys. Res. Lett.* 46(2), 644–651.

- Fichtner, A., A. Zunino, and L. Gebraad (2019). Hamiltonian Monte Carlo solution of tomographic inverse problems. *Geophys. J. Int.* 216(2), 1344–1363.
- Fish, J. and W. Chen (2001). Higher-order homogenization of initial/boundary-value problem. *J. Eng. Mech.* 127(12), 1223–1230.
- Fish, J. and W. Chen (2004). Space-time multiscale model for wave propagation in heterogeneous media. *Comp. Methods Appl. Mech. Engrg.* 193, 4837–4856.
- Friedrich, A., F. Krüger, and K. Klinge (1998). Ocean-generated microseismic noise located with the Gräfenberg array. *J. Seismol.* 2, 47–64.
- Gajewski, D. and E. Tessmer (2005). Reverse modelling for seismic event characterization. *Geophys. J. Int.* 163(1), 276–284.
- Gebraad, L., C. Boehm, and A. Fichtner (2020). Bayesian elastic full-waveform inversion using hamiltonian monte carlo. *J. Geophys. Res.* 125(3), e2019JB018428.
- Geevers, S., W. A. Mulder, and J. J. van der Vegt (2019). Efficient quadrature rules for computing the stiffness matrices of mass-lumped tetrahedral elements for linear wave problems. *SIAM journal on scientific computing* 41(2), A1041–A1065.
- Geuzaine, C. and J.-F. Remacle (2007). Gmsh: a three-dimensional finite element mesh generator with built-in pre-and post-processing facilities. In *Proceedings of the Second Workshop on Grid Generation for Numerical Computations, Tetrahedron II*.
- Gilbert, F. (1971). Excitation of normal modes of the Earth by earthquake sources. *Geophys. J. R. Astron. Soc.* 22, 223–226.
- Godefroy, G., G. Caumon, G. Laurent, and F. Bonneau (2019). Structural interpretation of sparse fault data using graph theory and geological rules. *Mathematical Geosciences* 51(8), 1091–1107.
- Gouédard, P., P. Roux, M. Campillo, and A. Verdel (2008). Convergence of the two-points correlation function toward the Green’s function in the context of a prospecting dataset. *Geophysics* 73(6), V47–V53.
- Grechka, V. and M. Kachanov (2006a). Effective elasticity of fractured rocks: A snapshot of the work in progress. *Geophysics* 71(6), W45–W58.
- Grechka, V. and M. Kachanov (2006b). Effective elasticity of rocks with closely spaced and intersecting cracks. *Geophysics* 71(3), D85–D91.
- Guillot, L., Y. Capdeville, and J. Marigo (2010). 2-D non-periodic homogenization of the elastic wave equation: SH case. *Geophys. J. Int.* 182, 1438–1454.
- Guo, P., G. Visser, and E. Saygin (2020). Bayesian trans-dimensional full waveform inversion: synthetic and field data application. *Geophys. J. Int.* 222(1), 610–627.
- Halliday, D. and A. Curtis (2008). Seismic interferometry, surface waves and source distribution. *Geophys. J. Int.* 175(3), 1067–1087.
- Harmon, N., P. Gerstoft, C. A. Rychert, G. A. Abers, M. S. de la Cruz, and K. M. Fischer (2008). Phase velocities from seismic noise using beamforming and cross correlation in Costa Rica and Nicaragua. *Geophys. Res. Lett.* 35, L19303.
- Hashin, Z. (1988). The differential scheme and its application to cracked materials. *Journal of the Mechanics and Physics of Solids* 36(6), 719–734.

- Hedjazian, N., Y. Capdeville, and T. Bodin (2021). Multiscale seismic imaging with inverse homogenization. *Geophys. J. Int.* 226(1), 676–691.
- Heney, F. S. and N. Pomphrey (1982). Self-consistent elastic moduli of a cracked solid. *Geophys. Res. Lett.* 9(8), 903–906.
- Hornung, U. (1997). *Homogenization and Porous Media*, Volume 6 of *Interdisciplinary Applied Mathematics Series*. Springer Verlag.
- Hudson, J. A. (1981). Wave speeds and attenuation of elastic waves in material containing cracks. *Geophys. J. Int.* 64(1), 133–150.
- Ibourichène, A., P. Cupillard, and J.-F. Barthélémy (2021). Investigating the effective properties of multiscale fractured media using non-periodic homogenization. In *82nd EAGE Conference and Exhibition 2016*. European Association of Geoscientists & Engineers.
- Ichimura, T., M. Hori, and J. Bielak (2009). A hybrid multiresolution meshing technique for finite element three-dimensional earthquake ground motion modelling in basins including topography. *Geophys. J. Int.* 177(3), 1221–1232.
- Igel, H. (2017). *Computational seismology: a practical introduction*. Oxford University Press.
- Igel, H., P. Mora, and B. Riollet (1995). Anisotropic wave propagation through finite-difference grids. *Geophysics* 60(4), 1203–1216.
- Irakarama, M., P. Cupillard, G. Caumon, P. Sava, and J. Edwards (2019). Appraising structural interpretations using seismic data — theoretical elements. *Geophysics* 84(2), N29–N40.
- Irakarama, M., G. Laurent, J. Renaudeau, and G. Caumon (2021). Finite difference implicit structural modeling of geological structures. *Mathematical Geosciences* 53(5), 785–808.
- Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2007). Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array. *J. Geophys. Res.* 112(B11).
- Julio, C., G. Caumon, and M. Ford (2015). Sampling the uncertainty associated with segmented normal fault interpretation using a stochastic downscaling method. *Tectonophysics* 639, 56–67.
- Kao, H. and S.-J. Shan (2004). The Source-Scanning Algorithm: mapping the distribution of seismic sources in time and space. *Geophys. J. Int.* 157(2), 589–594.
- Kelly, K. R., R. W. Ward, S. Treitel, and R. M. Alford (1976). Synthetic seismograms, a finite difference approach. *Geophysics* 41, 2–27.
- Kim, Y., Q. Liu, and J. Tromp (2011). Adjoint centroid-moment tensor inversions. *Geophys. J. Int.* 186(1), 264–278.
- Komatitsch, D. (1997). *Méthodes spectrales et éléments spectraux pour l'équation de l'élastodynamique 2D et 3D en milieu hétérogène*. PhD thesis, Université Paris 7.
- Komatitsch, D., Q. Liu, J. Tromp, P. Süß, C. Stidham, and J. H. Shaw (2004). Simulations of ground motion in the Los Angeles basin based upon the Spectral-Element Method. *Bull. Seism. Soc. Am.* 94(1), 187–206.
- Komatitsch, D., J. Ritsema, and J. Tromp (2002). The Spectral-Element Method, Beowulf computing, and global seismology. *Science* 298, 1737–1742.

- Komatitsch, D. and J. Tromp (1999). Introduction to the Spectral Element Method for three-dimensional seismic wave propagation. *Geophys. J. Int.* 139, 806–822.
- Komatitsch, D. and J. Tromp (2002a). Spectral-element simulations of global seismic wave propagation, part I: Validation. *Geophys. J. Int.* 149, 390–412.
- Komatitsch, D. and J. Tromp (2002b). Spectral-element simulations of global seismic wave propagation, part II: 3-D models, oceans, rotation, and gravity. *Geophys. J. Int.* 150, 303–318.
- Komatitsch, D. and J. P. Vilotte (1998). The Spectral Element Method: an efficient tool to simulate the seismic response of 2d and 3d geological structures. *Bull. Seism. Soc. Am.* 88, 368–392.
- Kononov, A., S. Minisini, E. Zhebel, and W. A. Mulder (2012). A 3d tetrahedral mesh generator for seismic problems. In *Proc. 74th EAGE Conference and Exhibition*. European Association of Geoscientists & Engineers.
- Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In J. Bednar, R. Redner, E. Robinson, and A. Weglein (Eds.), *Conference on Inverse Scattering: Theory and Application*. Soc. Industr. appl. Math., Philadelphia, PA.
- Larose, E., A. Derode, M. Campillo, and M. Fink (2004). Imaging from one-bit correlations of wide-band diffuse wavefields. *J. Appl. Phys.* 95(12), 8393–8399.
- Larose, E., P. Roux, and M. Campillo (2007). Reconstruction of Rayleigh-Lamb dispersion spectrum based on noise obtained from an air-jet forcing. *J. Acoust. Soc. Am.* 122(6), pp 3437.
- Lee, S.-J., H.-W. Chen, Q. Liu, D. Komatitsch, B.-S. Huang, and J. Tromp (2008). Three-dimensional simulations of seismic-wave propagation in the Taipei basin with realistic topography based upon the Spectral Element Method. *Bull. Seism. Soc. Am.* 98(1), 253–264.
- Legentil, C., J. Pellerin, P. Cupillard, A. Froehly, and G. Caumon (2022). Testing scenarios on geological models: Local interface insertion in a 2d mesh and its impact on seismic wave simulation. *Computers & Geosciences* 159, 105013.
- Li, X. D. and B. Romanowicz (1995). Comparison of global waveform inversions with and without considering cross-branch modal coupling. *Geophys. J. Int.* 121, 695–709.
- Li, X. D. and B. Romanowicz (1996). Global mantle shear velocity model developed using nonlinear asymptotic coupling theory. *J. Geophys. Res.* 101, 22 245–22 271.
- Li, Z. and M. van der Baan (2016). Microseismic event localization by acoustic time reversal extrapolation. *Geophysics* 81(3), KS123–KS134.
- Lin, F., M. P. Moschetti, and M. H. Ritzwoller (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophys. J. Int.* 173, 281–298.
- Lin, F., M. H. Ritzwoller, J. Townend, M. Savage, and S. Bannister (2007). Ambient noise Rayleigh wave tomography of New Zealand. *Geophys. J. Int.* 170, 649–666.
- Liu, Q. and Y. Gu (2012). Seismic imaging: From classical to adjoint tomography. *Tectonophysics* 566–567, 31–66.
- Liu, Q., J. Polet, D. Komatitsch, and J. Tromp (2004). Spectral-element moment tensor inversions for earthquakes in southern California. *Bull. Seism. Soc. Am.* 94(5), 1748–1761.
- Lobkis, O. I. and R. L. Weaver (2001). On the emergence of the Green’s function in the correlation of a diffuse field. *J. Acoust. Soc. Am.* 110, 3011–3017.

- Longuet-Higgins, M. S. (1950). A theory on the origin of microseisms. *Philos. Trans. R. Soc. Lond. A* 243, 1–35.
- Lu, Y., L. Stehly, A. Paul, and A. W. Group (2018). High-resolution surface wave tomography of the European crust and uppermost mantle from ambient seismic noise. *Geophys. J. Int.* 214(2), 1136–1150.
- Lysmer, J. and L. A. Drake (1972). A finite element method for seismology. In *Methods in Computational Physics*, Volume 11. New York: Academic Press.
- Ma, Y. and D. Hale (2013). Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion. *Geophysics* 78(6), R223–R233.
- Maday, Y. and A. Patera (1989). Spectral Element Methods for the incompressible Navier-Stokes equations. In A. Noor and J. Oden (Eds.), *State of the Art Survey in Computational Mechanics*, pp. 71–143. New-York: ASME.
- Magali, J. K., T. Bodin, N. Hedjazian, Y. Ricard, Y. Capdeville, and E. Debayle (2021). Quantifying intrinsic and extrinsic contributions to seismic anisotropy in tomographic models. *Earth and Space Science Open Archive*, 43.
- Marfurt, K. J. (1984). Accuracy of finite-difference and finite-element modeling of the scalar wave equation. *Geophysics* 49, 533–549.
- Marone, F. and B. Romanowicz (2007). Non-linear crustal corrections in high-resolution regional waveform seismic tomography. *Geophys. J. Int.* 170, 460–467.
- Masson, Y., P. Cupillard, Y. Capdeville, and B. Romanowicz (2014). On the numerical implementation of time-reversal mirrors for tomographic imaging. *Geophys. J. Int.* 196(3), 1580–1599.
- Mercerat, E. D. and N. Glinisky (2015). A nodal high-order discontinuous Galerkin method for elastic wave propagation in arbitrary heterogeneous media. *Geophysical Journal International* 201(2), 1101–1118.
- Merland, R., G. Caumon, B. Lévy, and P. Collon-Drouaillet (2014). Voronoi grids conforming to 3d structural features. *Computational Geosciences* 18(3-4), 373–383.
- Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux (2016). Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. *Geophys. J. Int.* 205(1), 345–377.
- Moczo, P., E. Bystrický, J. Kristek, J. M. Carcione, and M. Bouchon (1997). Hybrid modeling of P-SV seismic motion at inhomogeneous viscoelastic topographic structures. *Bull. Seism. Soc. Am.* 87(5), 1305–1323.
- Moczo, P., J. Kristek, and M. Gális (2014). *The finite-difference modelling of earthquake motions: Waves and ruptures*. Cambridge University Press.
- Moczo, P., J. O. A. Robertsson, and L. Eisner (2007). The finite-difference time-domain method for modeling of seismic wave propagation. In R. Wu, V. Maupin, and R. Dmowska (Eds.), *Advances in Wave Propagation in Heterogeneous Earth*, Volume 48, pp. 421–516. Elsevier/Academic Press, San Diego.
- Montagner, J. P. and T. Tanimoto (1991). Global upper mantle tomography of seismic velocities and anisotropies. *J. Geophys. Res.* 96, 20337–20351.
- Montelli, R., G. Nolet, F. Dahlen, G. Masters, E. Engdahl, and S. Hung (2004). Finite-frequency tomography reveals a variety of mantle plumes. *Science* 303, 338–343.

- Mori, T. and K. Tanaka (1973). Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta metallurgica* 21(5), 571–574.
- Nakata, N., J. P. Chang, J. F. Lawrence, and P. Boué (2015). Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry. *J. Geophys. Res.* 120(2), 1159–1173.
- Neuman, S. P. (2008). Multiscale relationships between fracture length, aperture, density and permeability. *Geophys. Res. Lett.* 35(22).
- Nieto-Samaniego, A., S. Alaniz-Alvarez, G. Tolson, K. Oleschko, G. Korvin, S. Xu, and J. Perez-Venzor (2005). Spatial distribution, scaling and self-similar behavior of fracture arrays in the Los Planes Fault, Baja California Sur, Mexico. *Pure and Applied Geophysics* 162(5), 805–826.
- Nimiya, H., T. Ikeda, and T. Tsuji (2020). Three-dimensional S wave velocity structure of central Japan estimated by surface-wave tomography using ambient noise. *J. Geophys. Res.* 125(4), e2019JB019043.
- Noble, M., A. Gesret, and N. Belayouni (2014). Accurate 3-D finite difference computation of travel-times in strongly heterogeneous media. *Geophys. J. Int.* 199(3), 1572–1585.
- O’Connell, R. J. and B. Budiansky (1974). Seismic velocities in dry and saturated cracked solids. *J. Geophys. Res.* 79(35), 5412–5426.
- Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Métivier, V. Monteiller, A. Ribodetti, and J. Virieux (2015). Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the viscoacoustic vertical transverse isotropic approximation. *Geophys. J. Int.* 102, 1362–1391.
- Orlowsky, B., E. H. Saenger, Y. Guéguen, and S. Shapiro (2003). Effects of parallel crack distributions on effective elastic properties—a numerical study. *International Journal of Fracture* 124(3-4), L171–L178.
- Pageot, D., S. Operto, M. Vallée, R. Brossier, and J. Virieux (2013). A parametric analysis of two-dimensional elastic full waveform inversion of teleseismic data for lithospheric imaging. *Geophys. J. Int.* 193(3), 1479–1505.
- Papanicolaou, G. C. and S. R. S. Varadhan (1981). Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, Volume 27 of *Colloq. Math. Soc. János Bolyai*, pp. 835–873. North-Holland, Amsterdam.
- Parnell, W. and I. Abrahams (2008). Homogenization for wave propagation in periodic fibre-reinforced media with complex microstructure. I - Theory. *J. Mech. Phys. Solids* 56(7), 2521–2540.
- Patera, A. T. (1984). A Spectral Element Method for fluid dynamics: laminar flow in a channel expansion. *J. Comp. Phys.* 54, 468–488.
- Pellerin, J., A. Botella, F. Bonneau, A. Mazuyer, B. Chauvin, B. Lévy, and G. Caumon (2017). RINGMesh: A programming library for developing mesh-based geomodeling applications. *Computers & Geosciences* 104, 93–100.
- Pellerin, J., B. Lévy, G. Caumon, and A. Botella (2014). Automatic surface remeshing of 3d structural models at specified resolution: A method based on Voronoi diagrams. *Computers & Geosciences* 62, 103–116.
- Poli, P., M. Campillo, H. Pedersen, and the LAPNET Working Group (2012). Body-wave imaging of earth’s mantle discontinuities from ambient seismic noise. *Science* 338(6110), 1063–1065.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics* 64(3), 888–901.

- Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009). Anelastic Earth structure from the coherency of the ambient seismic field. *J. Geophys. Res.* *144*, B07303, doi:10.1029/2008JB006067.
- Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana (2004). Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt. *Geophys. J. Int.* *159*(3), 1032–1056.
- Renat, Z., P. Cupillard, and Y. Capdeville (2022). Interpretation of time reversal focal spot based on point-source homogenization. *Geophys. J. Int.*, submitted.
- Ritsema, J., P. Cupillard, B. Tauzin, W. Xu, L. Stixrude, and C. Lithgow-Bertelloni (2009). Joint mineral physics and seismic wave travel time analysis of upper mantle temperature. *Geology* *37*, 363–366.
- Ritsema, J., H. vanHeijst, and J. H. Woodhouse (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science* *286*, 1925–1928.
- Robertsson, J. O. A., J. O. Blanch, and W. W. Symes (1994). Viscoelastic finite-difference modeling. *Geophysics* *59*(9), 1444–1456.
- Romijn, R. (2017). Groningen full elastic velocity model. Technical report, Nederlandse Aardolie Maatschappij.
- Ronchi, C., R. Iacono, and P. S. Paolucci (1996). The ‘Cubed Sphere’: a new method for the solution of partial differential equations in spherical geometry. *J. Comp. Phys.* *124*, 93–114.
- Ruiu, J., G. Caumon, and S. Viseur (2016). Modeling channel forms and related sedimentary objects using a boundary representation based on non-uniform rational b-splines. *Mathematical Geosciences* *48*(3), 259–284.
- Saade, M., J. P. Montagner, P. Roux, P. Cupillard, S. Durand, and F. Brenguier (2015). Influence of seismic anisotropy on the cross correlation tensor: numerical investigations. *Geophys. J. Int.* *201*(2), 595–604.
- Sabra, K. G., P. Roux, and W. A. Kuperman (2005). Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide. *J. Acoust. Soc. Am.* *117*(1), 164–174.
- Saenger, E. H. and S. A. Shapiro (2002). Effective velocities in fractured media: a numerical study using the rotated staggered finite-difference grid. *Geophysical Prospecting* *50*(2), 183–194.
- Sævik, P. N., I. Berre, M. Jakobsen, and M. Lien (2013). A 3d computational study of effective medium methods applied to fractured media. *Transport in porous media* *100*(1), 115–142.
- Sager, K., C. Boehm, L. Ermert, L. Krischer, and A. Fichtner (2018). Sensitivity of seismic noise correlation functions to global noise sources. *Journal of Geophysical Research: Solid Earth* *123*(8), 6911–6921.
- Sanchez-Palencia, E. (1980). *Non-homogenous media and vibration theory*, Volume 127 of *Lecture notes in physics*. Berlin, Heidelberg: Springer-Verlag.
- Sánchez-Sesma, F. J. and M. Campillo (2006). Retrieval of the Green’s function from cross-correlation: The canonical elastic problem. *Bull. Seism. Soc. Am.* *96*, 1182–1191.
- Sánchez-Sesma, F. J., J. Pérez-Ruiz, M. Campillo, and F. Lúzon (2006). Elastodynamic 2D Green function retrieval from cross-correlation: Canonical inclusion problem. *Geophys. Res. Lett.* *33*, L13305, doi:10.1029/2006GL026454.
- Santugini-Repiquet, K. (2007). Homogenization of the demagnetization field operator in periodically perforated domains. *J. Math. Anal. Appl.* *334*(1), 502–516.

- Schenk, O. and K. Gärtner (2006). On fast factorization pivoting methods for symmetric indefinite systems. *Elec. Trans. Numer. Anal.* 23, 158–179.
- Schoenberg, M. (1980). Elastic wave behavior across linear slip interfaces. *The Journal of the Acoustical Society of America* 68(5), 1516–1521.
- Seriani, G. (1998). 3D large-scale wave propagation modeling by a Spectral Element Method on a Cray T3E multiprocessor. *Comp. Methods Appl. Mech. Engrg.* 164(1), 235–247.
- Seriani, G. and E. Priolo (1994). Spectral Element Method for acoustic wave simulation in heterogeneous media. *Finite Elements in Analysis and Design* 16, 337–348.
- Sevostianov, I. and M. Kachanov (2012). Effective properties of heterogeneous materials: Proper application of the non-interaction and the “dilute limit” approximations. *International Journal of Engineering Science* 58, 124–128.
- Shapiro, N. M. and M. Campillo (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophys. Res. Lett.* 31, L07614.
- Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller (2005). High-resolution surface wave tomography from ambient seismic noise. *Science* 307, 1615–1618.
- Shapiro, N. M. and M. H. Ritzwoller (2002). Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle. *Geophys. J. Int.* 151, 88–105.
- Si, H. (2015). Tetgen, a delaunay-based quality tetrahedral mesh generator. *ACM Trans. Math. Softw.* 41(2).
- Snieder, R. (2004). Extracting the Green’s function from the correlation of coda waves: A derivation based on stationary phase. *Phys. Rev. E* 69, 046610.
- Somala, S. N., J.-P. Ampuero, and N. Lapusta (2018). Finite-fault source inversion using adjoint methods in 3-D heterogeneous media. *Geophys. J. Int.* 214(1), 402–420.
- Stehly, L., M. Campillo, and N. Shapiro (2006). A study of the seismic noise from its long-range correlation properties. *J. Geophys. Res.* 111, B10306.
- Stehly, L., P. Cupillard, and B. Romanowicz (2011). Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise. *C. R. Geoscience* 343, 591–599.
- Stupazzini, M., R. Paolucci, and H. Igel (2009). Near-fault earthquake ground-motion simulation in the Grenoble valley by a high-performance spectral element code. *Bull. Seism. Soc. Am.* 99(1), 286–301.
- Suzuki, S., G. Caumon, and J. Caers (2008). Dynamic data integration for structural modeling: model screening approach using a distance-based model parameterization. *Computational Geosciences* 12(1), 105–119.
- Tape, C., Q. Liu, A. Maggi, and J. Tromp (2009). Adjoint tomography of the southern California crust. *Science* 325, 988–992.
- Tape, C., Q. Liu, A. Maggi, and J. Tromp (2010). Seismic tomography of the southern California crust based on spectral-element and adjoint methods. *Geophys. J. Int.* 180, 433–462.
- Tape, C., Q. Liu, and J. Tromp (2007). Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves. *Geophys. J. Int.* 168(3), 1105–1129.

- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics* 49, 1259–1266.
- Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure Appl. Geophys.* 128(1/2), 365–399.
- Terrana, S., J. P. Vilotte, and L. Guillot (2017). A spectral hybridizable discontinuous Galerkin method for elastic-acoustic wave propagation. *Geophys. J. Int.* 213(1), 574–602.
- Thore, P., A. Shtuka, M. Lecour, T. Ait-Ettajer, and R. Cognot (2002). Structural uncertainties: Determination, management, and applications. *Geophysics* 67(3), 840–852.
- Toshinawa, T. and T. Ohmachi (1992). Love wave propagation in three-dimensional sedimentary basin. *Bull. Seism. Soc. Am.* 82, 1661–1667.
- Tromp, J., Y. Luo, S. Hanasoge, and D. Peter (2010). Noise cross-correlation sensitivity kernels. *Geophys. J. Int.* 183, 791–819, doi:10.1111/j.1365-246X.2010.04721.x.
- Tromp, J., C. Tape, and Q. Liu (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.* 160, 195–216.
- Tsai, V. C. (2009). On establishing the accuracy of noise tomography travel-time measurements in a realistic medium. *Geophys. J. Int.* 178(3), 1555–1564.
- van Leeuwen, T. and W. A. Mulder (2010). A correlation-based misfit criterion for wave-equation traveltime tomography. *Geophys. J. Int.* 182(3), 1383–1394.
- van Tiggelen, B. A. (2003). Green function retrieval and time-reversal in a disordered world. *Phys. Rev. Lett.* 91(24), 243904.
- Vasilyeva, M., J. De Basabe, Y. Efendiev, and R. Gibson Jr (2019). Multiscale model reduction of the wave propagation problem in viscoelastic fractured media. *Geophys. J. Int.* 217(1), 558–571.
- Vavakin, A. and R. Salganik (1975). Effective characteristics of nonhomogeneous media with isolated nonhomogeneities. *Mech. Solids* 10, 65–75.
- Virieux, J. (1984). SH wave propagation in heterogeneous media: velocity-stress finite difference method. *Geophysics* 49, 1933–1942.
- Virieux, J. (1986). P-SV wave propagation in heterogeneous media: velocity-stress finite difference method. *Geophysics* 51, 889–901.
- Virieux, J., H. Calandra, and R.-E. Plessix (2011). A review of the spectral, pseudo-spectral, finite-difference and finite-element modelling techniques for geophysical imaging. *Geophys. Prosp.* 59(5), 794–813.
- Virieux, J. and S. Operto (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74(6), WCC1–WCC26.
- Visser, G., P. Guo, and E. Saygin (2019). Bayesian transdimensional seismic full-waveform inversion with a dipping layer parameterization. *Geophysics* 84(6), R845–R858.
- Wapenaar, K. (2004). Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross-correlation. *Phys. Rev. Lett.* 93, 254301.
- Wellmann, F. and G. Caumon (2018). 3-D Structural geological models: Concepts, methods, and uncertainties. *Advances in Geophysics* 59, 1–121.

- Willacy, C., E. van Dedem, S. Minisini, J. Li, J.-W. Blokland, I. Das, and A. Droujinine (2019). Full-waveform event location and moment tensor inversion for induced seismicity. *Geophysics* 84(2), KS39–KS57.
- Woodhouse, J. H. and A. M. Dziewonski (1984). Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms. *J. Geophys. Res.* 89, 5953–5986.
- Woodhouse, J. H. and T. P. Girnius (1982). Surface waves and free oscillations in a regionalized earth model. *Geophys. J. R. Astron. Soc.* 78, 641–660.
- Yamakawa, S. and K. Shimada (2003). Fully-automated hex-dominant mesh generation with directionality control via packing rectangular solid cells. *International Journal for Numerical Methods in Engineering* 57(15), 2099–2129.
- Yang, Y. and M. H. Ritzwoller (2008). Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochemistry, Geophysics, Geosystems* 9(2), Q02008.
- Yang, Y., M. H. Ritzwoller, A. L. Levshin, and N. M. Shapiro (2007). Ambient noise Rayleigh wave tomography across Europe. *Geophys. J. Int.* 168, 259–274.
- Yao, H. and R. D. van der Hilst (2009). Analysis of ambient noise energy distribution and phase velocity bias in ambient noise tomography, with application to SE Tibet. *Geophys. J. Int.* 179, 1113–1132.
- Yao, H., R. D. van der Hilst, and M. V. de Hoop (2006). Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps. *Geophys. J. Int.* 166, 732–744.
- Yuan, H., S. French, P. Cupillard, and B. Romanowicz (2014). Lithospheric expression of geological units in central and eastern North America from full waveform tomography. *Earth Planet. Sci. Lett.* 402, 176–186.
- Zunino, A., A. Khan, P. Cupillard, and K. Mosegaard (2016). Constitution and Structure of Earth's Mantle : Insights from Mineral Physics and Seismology. In *Integrated Imaging of the Earth : Theory and Applications*, AGU Monograph Series, pp. 219–243.

Chapter 2

Personal records

2.1 Curriculum vitae

2.1.1 Contact

Nancy School of Geology (ENSG)
2 rue du Doyen Marcel Roubault
54518 Vandœuvre-lès-Nancy
France

Phone: +33 (0)3 72 74 45 24
E-mail: paul.cupillard@univ-lorraine.fr
Website: georesources.univ-lorraine.fr/content/cupillard

2.1.2 Biodata

- Date of birth: March 12, 1981
- Place of birth: Besançon, France
- Citizenship: French
- Languages: French, English

2.1.3 Appointments

- 2012 - now **Associate Professor at Université de Lorraine**, Nancy, France
GeoRessources Lab (UMR 7359)
Nancy School of Geology (ENSG)
- 2011 - 2012 **Postdoctoral fellow in the Institut de Physique du Globe de Paris**, Université de Paris VII, Paris, France
- 2008 - 2010 **Postdoctoral fellow in the Berkeley Seismological Laboratory**, UC Berkeley, California, USA
- 2005 - 2008 **PhD in Geophysics, Institut de Physique du Globe de Paris**, Université de Paris VII, Paris, France
Spectral element simulation of waveforms obtained by seismic ambient noise correlation
Advisors: Jean-Paul Montagner and Yann Capdeville
- 2001 - 2004 **MSc in Geophysics, Institut de Physique du Globe de Strasbourg**, Université Louis Pasteur, Strasbourg, France
Graduation project conducted in the Institut de Physique du Globe de Paris, Université de Paris VII, Paris, France
Ecole de Physique du Globe de Strasbourg, Université Louis Pasteur, Strasbourg, France
Graduation project conducted in the Geophysics Laboratory of the Atomic Energy Agency (CEA), Tahiti, French Polynesia

2.2 Teaching data

2.2.1 Courses

- 2016 - now
- Applied Geophysics** (103,5 h/year)
 - Finite Element and C++ programming** (24 h/year)
 - Geomodeling** (16 h/year)
 - General physics** (16 h/year)
 - Optimization** (13,5 h/year)
 - Inverse problem** (3 h/year)
- 2012 - 2016
- Applied Geophysics** (103,5 h/year)
 - Geophysics field camp** (70 h/year)
 - Geomodeling** (16 h/year)
 - Finite Element** (6 h/year)

2.2.2 Supervision

- **Research projects:**
 - **2nd-year ENSG students:** S. Carniato, M. Schera, D. Lavandier (2013-2014), A. Pouille (2014-2015), M. Thierry-Coudon (2017-2018), M. Vivès (2018-2019), C. Garayt, H. Marie, J.-C. Gouzou (2019-2020), E. Despinoy, T. Fleury, A. Cormier (2020-2021)
 - **3rd-year ENSG students:** M. Lerat (2013-2014), M. Irakarama, A. Guyot, M. Raguanel (2014-2015), E. Lavoine (2015-2016), Y. Frantz, J. Langanay, M. Schuh-Senlis (2016-2017), T. Béraud, Z. Renat (2017-2018), M. Vivès, C. Barrière (2019-2020), M. Rapenne (2020-2021)
- **Industrial internships:**
 - **Reviewer of internship reports** of ENSG students (6 per year on average)
 - **Jury member of internship defenses** of ENSG students (16 per year on average)

2.2.3 Responsibilities

- Organization of the **Applied Geophysics program** for the 2nd-year ENSG students
- Organization of the **GIS - Remote Sensing - Image Processing program** for the 2nd-year ENSG students
- Organization of the **Geophysics field camp** for the 2nd-year ENSG students from 2012 to 2016
- Organization of **geoscientific conferences** for the 1st-year ENSG students from 2013 to 2015
- Member of the **ENSG council** from 2017 to 2021

2.3 Research data

2.3.1 Indicators

- **h-index: 13**; citations: 934 (Google Scholar, January 31, 2022)
- **21 journal articles** (4 as first author)
- **2 book chapters**
- **12 conference proceedings** (3 as first author)
- **52 conference communications** (17 as a presenter: 9 talks and 8 posters)

2.3.2 Supervision

- **6 PhD students:**
 - Marius Rapenne (2021-):
Homogenization of quaternary basins for the simulation of lithological site effects
 - Zoé Renat (2019-):
Seismic event characterization using time-reversal
 - Capucine Legentil (2019-):
Local updating of geomodels
 - Melchior Schuh-Senlis (2018-2021):
Using Stokes flow equations for the geomechanical restoration of geological models
 - Modeste Irakarama (2015-2019):
Towards reducing structural interpretation uncertainties using seismic data
 - Antoine Mazuyer (2014-2018):
Stress estimation in reservoirs based on an inverse approach
- **11 MSc students:**
 - 2021-2022: Thibaut Leibel
 - 2020-2021: Marius Rapenne, Vivien Belin
 - 2019-2020: Morgane Vivès
 - 2017-2018: Thomas Béraud, Zoé Renat
 - 2016-2017: Yves Frantz, Jean Langanay, Melchior Schuh-Senlis
 - 2015-2016: Etienne Lavoine
 - 2014-2015: Arthur Guyau, Modeste Irakarama, Margaux Ragueneil
 - 2013-2014: Marine Lerat

2.3.3 Responsibilities

- Member of the **board of the RING-GOCAD consortium**
- Member of the **organization committee of the RING Annual Meeting**
- Member of the **organization committee of the 21st IAMG Annual Conference**
- Scientific partner in the **HIWAI ANR project** (ANR-16-CE31-0022)
- Scientific partner in the **DONUTS Géodénergies project**
- Scientific manager of the **Numerical Modeling platform** of the GeoRessources Lab
- Member of the **scientific committee of the Grand archeological site** (Vosges, France)
- Member of the **EXPLOR scientific committee** (Université de Lorraine supercomputing facilities) from 2014 to 2021
- Member of the **UL-CNRS-SAYENS Software Detection group** from 2017 to 2019
- Member of the **GeoRessources Lab council** from 2013 to 2019
- Examiner of **Maria Saade's PhD** progress and **Charline Julio's PhD** defense

2.3.4 Collaborations

- **Academic:**

- B. Romanowicz (UC Berkeley)
- Y. Capdeville (CNRS, Université de Nantes)
- G. Caumon (Université de Lorraine)
- A. Fichtner (ETH Zürich)
- A. Khan (ETH Zürich)
- Y. Masson (INRIA, Université de Pau et des Pays de l'Adour)
- L. Métivier (Université de Grenoble Alpes)
- W. Mulder (TU Delft)
- H. Yuan (Macquarie University)

- **Industrial:**

- P. Anquez and A. Botella (Geode-solutions)
- J.-F. Barthélémy (CEREMA)
- G. Daniel and P. Gouédard (Baker Hugues)
- N. Glinsky (CEREMA)
- C. Gouache and P. Tinard (Caisse Centrale de Réassurance)
- J. Pellerin (TotalEnergies)
- P. Thore and Y. Leroy (TotalEnergies)

2.3.5 Mobility

- **15 seminars** outside my institution:

- ISTerre, Université de Grenoble Alpes, France (2020)
- Laboratoire de Géologie de Lyon, Université Claude Bernard & ENS de Lyon, France (2013, 2018, 2019)
- Department of Geophysics, Tel Aviv University, Israel (2018)
- Center for Wave Phenomena, Colorado School of Mines, CO, USA (2017)
- ExxonMobil, Houston Campus, TX, USA (2017)
- Laboratoire de Planétologie et Géodynamique, Université de Nantes, France (2017)
- Institute of Geophysics, ETH Zürich, Switzerland (2013, 2016)
- CGG, Massy, France (2016)
- Total E&P, Centre Scientifique et Technique Jean-Féger, Pau, France (2016)
- Niels Bohr Institute, University of Copenhagen, Denmark (2013)
- Géoazur, Université Nice Sophia Antipolis, France (2012)
- Department of Earth and Environmental Sciences, Ludwig Maximilian University of Munich, Germany (2012)

- **3 stays abroad** (apart from seminars and conferences):

- June 2019: A week at Neubrandenburg University of Applied Sciences (Germany) in the frame of the ERASMUS+ program

- Sep. - Dec. 2018: Sabbatical semester at Tel Aviv University (Israël) supported by the Université de Lorraine Widen Horizons program
- April 2016: A week at Neubrandenburg University of Applied Sciences (Germany) in the frame of the ERASMUS+ program

- **2 fieldworks:**

- April 2017: Geophysical prospection in the Cleebourg quarry, France
- September 2015: Geophysical prospection in the Ifni region, Marocco

2.3.6 Miscellaneous

- **14 paper reviews** for international journals:

- *Geophysical Journal International* (IF = 2.528)
- *Computers & Geosciences* (IF = 2.567)
- *Journal of Geophysical Research* (IF = 3.44)
- *Solid Earth* (IF = 2.083)
- *Journal of Applied Geophysics* (IF = 1.975)

- Learned societies:

- American Geophysical Union (**AGU**)
- European Geophysical Union (**EGU**)
- Society of Exploration Geophysicists (**SEG**)
- European Association of Geoscientists and Engineers (**EAGE**)

- Prime d'Encadrement Doctoral et de Recherche (**PEDR**)

2.4 List of publications and communications

Journal Articles

- [A1] Capucine Legentil, Jeanne Pellerin, Paul Cupillard, Algiane Froehly, Guillaume Caumon. “Testing scenarios on geological models: Local interface insertion in a 2D mesh and its impact on seismic wave simulation”. In: *Computers & Geosciences* 159 (2022), p. 105013. DOI: 10.1016/j.cageo.2021.105013.
- [A2] Zoé Renat, Paul Cupillard, Yann Capdeville. “Interpretation of time reversal focal spot based on point-source homogenization”. In: *Geophysical Journal International* (2022), submitted.
- [A3] Pierre Anquez, Nathalie Glinsky, Paul Cupillard, Guillaume Caumon. “Impacts of geometric model simplifications on wave propagation—application to ground motion simulation in the lower Var valley basin (France)”. In: *Geophysical Journal International* 229.1 (2021), pp. 110–137. DOI: 10.1093/gji/ggab447.
- [A4] François Desquilbet, Jian Cao, Paul Cupillard, Ludovic Métivier, Jean-Marie Mirebeau. “Single Pass Computation of First Seismic Wave Travel Time in Three Dimensional Heterogeneous Media With General Anisotropy”. In: *Journal of Scientific Computing* 89 (2021), p. 23. DOI: 10.1007/s10915-021-01607-8.
- [A5] Melchior Schuh-Senlis, Cedric Thieulot, Paul Cupillard, Guillaume Caumon. “Towards the application of Stokes flow equations to structural restoration simulations”. In: *Solid Earth* 11.5 (2020), pp. 1909–1930. DOI: 10.5194/se-11-1909-2020.
- [A6] Pierre Anquez, Jeanne Pellerin, Modeste Irakarama, Paul Cupillard, Bruno Lévy, Guillaume Caumon. “Automatic correction and simplification of geological maps and cross-sections for numerical simulations”. In: *Comptes Rendus Géoscience* 351.1 (2019), pp. 48–58. DOI: 10.1016/j.crte.2018.12.001.
- [A7] Modeste Irakarama, Paul Cupillard, Guillaume Caumon, Paul Sava, Jonathan Edwards. “Appraising structural interpretations using seismic data-theoretical elements”. In: *Geophysics* 84.2 (2019), N29–N40. DOI: 10.1190/geo2018-0128.1.
- [A8] Paul Cupillard, Yann Capdeville. “Non-periodic homogenization of 3-D elastic media for the seismic wave equation”. In: *Geophysical Journal International* 213.2 (2018), pp. 983–1001. DOI: 10.1093/gji/ggy032.
- [A9] Andreas Fichtner, Dirk-Philip Van Herwaarden, Michael Afanasiev, Saulé Simuté, Lion Krischer, Yeşim Çubuk-Sabuncu, Tuncay Taymaz, Lorenzo Colli, Erdinc Saygin, Antonio Villaseñor, Jeannot Trampert, Paul Cupillard, Hans-Peter Bunge, Heiner Igel. “The Collaborative Seismic Earth Model: Generation 1”. In: *Geophysical Research Letters* 45.9 (2018), pp. 4007–4016. DOI: 10.1029/2018GL077338.
- [A10] Antoine Mazuyer, Paul Cupillard, Richard Giot, Marianne Conin, Yves Leroy, Pierre Thore. “Stress estimation in reservoirs using an integrated inverse method”. In: *Computers & Geosciences* 114 (2018), pp. 30–40. DOI: 10.1016/j.cageo.2018.01.004.
- [A11] Yann Capdeville, Ming Zhao, Paul Cupillard. “Fast Fourier homogenization for elastic wave propagation in complex media”. In: *Wave Motion* 54 (2015), pp. 170–186. DOI: 10.1016/j.wavemoti.2014.12.006.
- [A12] M. Saade, J. P. Montagner, P. Roux, Paul Cupillard, S. Durand, F. Brenguier. “Influence of seismic anisotropy on the cross correlation tensor: numerical investigations”. In: *Geophysical Journal International* 201.2 (2015), pp. 595–604. DOI: 10.1093/gji/ggu470.
- [A13] Yder J. Masson, Paul Cupillard, Yann Capdeville, Barbara Romanowicz. “On the numerical implementation of time-reversal mirrors for tomographic imaging”. In: *Geophysical Journal International* 196.3 (2014), pp. 1580–1599. DOI: 10.1093/gji/ggt459.
- [A14] Huaiyu Yuan, Scott French, Paul Cupillard, Barbara Romanowicz. “Lithospheric expression of geological units in central and eastern North America from full waveform tomography”. In: *Earth and Planetary Science Letters* 402.SI (2014), pp. 176–186. DOI: 10.1016/j.epsl.2013.11.057.
- [A15] A. Fichtner, E. Saygin, T. Taymaz, P. Cupillard, Y. Capdeville, J. Trampert. “The deep structure of the North Anatolian Fault Zone”. In: *Earth and Planetary Science Letters* 373 (2013), pp. 109–117. DOI: 10.1016/j.epsl.2013.04.027.
- [A16] A. Fichtner, J. Trampert, P. Cupillard, E. Saygin, T. Taymaz, Y. Capdeville, A. Villaseñor. “Multiscale full waveform inversion”. In: *Geophysical Journal International* 194.1 (2013), pp. 534–556. DOI: 10.1093/gji/ggt118.

- [A17] Paul Cupillard, Elise Delavaud, Gaël Burgos, Geatano Festa, Jean-Pierre Vilotte, Yann Capdeville, Jean-Paul Montagner. “RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale”. In: *Geophysical Journal International* 188.3 (2012), pp. 1203–1220. DOI: 10.1111/j.1365-246X.2011.05311.x.
- [A18] P. Cupillard, L. Stehly, B. Romanowicz. “The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise”. In: *Geophysical Journal International* 184.3 (2011), pp. 1397–1414. DOI: 10.1111/j.1365-246X.2010.04923.x.
- [A19] L. Stehly, P. Cupillard, B. Romanowicz. “Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise”. In: *Comptes Rendus Géoscience* 343.8-9 (2011), pp. 591–599. DOI: 10.1016/j.crte.2011.03.005.
- [A20] Paul Cupillard, Yann Capdeville. “On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach”. In: *Geophysical Journal International* 181.3 (2010), pp. 1687–1700. DOI: 10.1111/j.1365-246X.2010.04586.x.
- [A21] J. Ritsema, Paul Cupillard, B. Tauzin, Wei-Jiang Xu, L. Stixrude, C. Lithgow-Bertelloni. “Joint mineral physics and seismic wave travelttime analysis of upper mantle temperature”. In: *Geology* 37 (2009), pp. 363–366. DOI: 10.1130/G25428A.1.

Book Chapters

- [BC1] Yann Capdeville, Paul Cupillard, Sneha Singh. “An introduction to the two-scale homogenization method for seismology”. In: *Advances in Geophysics*. Ed. by B. Moseley and L. Krischer. Vol. 61. Elsevier, 2020, pp. 217–306. DOI: 10.1016/bs.agph.2020.07.001.
- [BC2] A. Zunino, A. Khan, Paul Cupillard, K. Mosegaard. “Constitution and Structure of Earth’s Mantle : Insights from Mineral Physics and Seismology”. In: *Integrated Imaging of the Earth : Theory and Applications*. AGU Monograph Series. 2016, pp. 219–243. DOI: 10.1002/9781118929063.

Conferences with Proceedings

- [iP1] Ibourichène Anais, Paul Cupillard, Jean-François Barthélémy. “Investigating the Effective Properties of Multiscale Fractured Media Using Non-Periodic Homogenization”. In: *82nd EAGE Conference and Exhibition*. Amsterdam, Dec. 2021. DOI: 10.3997/2214-4609.202011943.
- [iP2] Capucine Legentil, Jeanne Pellerin, Paul Cupillard, Guillaume Caumon. “Testing Scenarios on Geological Models: Local Interface Insertion in a 2D Mesh and its Impact on Seismic Wave Simulation”. In: *10th International Conference on Adaptative Modeling and Simulation*. June 2021. DOI: 10.23967/admos.2021.077.
- [iP3] Paul Cupillard, Wim A. Mulder, Pierre Anquez, Antoine Mazuyer, Jean-François Barthélémy. “The Apparent Anisotropy of the SEG-EAGE Overthrust Model”. In: *82nd EAGE Conference and Exhibition*. Online, Dec. 2020. DOI: 10.3997/2214-4609.202011955.
- [iP4] Schuh-Senlis Melchior, Cedric Thieulot, Paul Cupillard, Guillaume Caumon. “Structural Restoration of Geological Structures with Viscous Stokes Flow - Principle and First Results”. In: *82nd EAGE Conference and Exhibition*. Online, Dec. 2020. DOI: 10.3997/2214-4609.202010733.
- [iP5] François Bonneau, M. Ragueneil, L. Scholtes, Paul Cupillard. “Simulating Micro-seismic Activity with a Discrete Geomechanical Model”. In: *79th EAGE Conference and Exhibition*. Paris, France, June 2017. DOI: 10.3997/2214-4609.201701223.
- [iP6] Paul Cupillard, Y. Capdeville. “Performance and Convergence of the Non-periodic Homogenization for the 3D Elastic Wave Equation”. In: *79th EAGE Conference and Exhibition*. Paris, France, June 2017. DOI: 10.3997/2214-4609.201700524.
- [iP7] M. Irakarama, Paul Cupillard, G. Caumon, P. Sava. “Appraising Structural Interpretations Using Seismic Data Misfit Functionals”. In: *79th EAGE Conference and Exhibition*. Paris, France, June 2017. DOI: 10.3997/2214-4609.201700536.

- [iP8] Modeste Irakarama, Paul Cupillard, Guillaume Caumon, Paul Sava. “Appraising structural models using seismic data: Problem and challenges”. In: *SEG Technical Program Expanded Abstracts*. Houston, TX, United States: Society of Exploration Geophysicists, Sept. 2017. DOI: 10.1190/segam2017-17791860.1.
- [iP9] A. Mazuyer, Paul Cupillard, R. Giot, M. Conin, P. Thore, Y. Leroy. “Integrated Inverse Method to Estimate Virgin Stress State in Reservoirs and Overburden”. In: *79th EAGE Conference and Exhibition*. Paris, France, June 2017. DOI: 10.3997/2214-4609.201700935.
- [iP10] Modeste Irakarama, Paul Cupillard, Guillaume Caumon. “Reduction of Fault Uncertainty Using Vertical Seismic Profiling Data”. In: *78th EAGE Conference and Exhibition*. Vienna, Austria, May 2016. DOI: 10.3997/2214-4609.201600681.
- [iP11] Paul Cupillard, Yann Capdeville, Arnaud Botella. “Homogenization of 3d geological models for seismic wave propagation”. In: *SEG Technical Program Expanded Abstracts*. New Orleans, LA, United States, Oct. 2015. DOI: 10.1190/segam2015-5907841.1.
- [iP12] Philippe Roux, Jean-Paul Montagner, Florent Brenguier, Stéphanie Durand, Paul Cupillard, Lucia Zaccarelli. “Monitoring of Fractured Media by Temporal Changes of Anisotropy Using Ambient Seismic Noise”. In: *74th EAGE Conference and Exhibition*. Copenhagen, Denmark, June 2012. DOI: 10.3997/2214-4609.20148753.

Conferences without Proceedings

- [C1] Zoé Renat, Paul Cupillard, Yann Capdeville. “Interpretation of the focal spot obtained with time reversal”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2021.
- [C2] Paul Cupillard, Wim A. Mulder, Modeste Irakarama, Antoine Mazuyer, Pierre Anquez. “Small-scale-induced anisotropy of a 3D subsurface model: quantitative analysis and numerical simulations of waves within”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2019.
- [C3] Corentin Gouache, Paul Cupillard, Pierre Tinard. “Calibration de lois d’atténuation par simulation de propagation d’ondes sismiques”. In: *Colloque National de l’AFPS*. Strasbourg, France, Sept. 2019.
- [C4] Melchior Schuh-Senlis, Guillaume Caumon, Paul Cupillard. “Representing faults in a geomechanical restoration scheme using creeping flow equations”. In: *20th Annual Conference of the IAMG*. State College, PA, United States, Aug. 2019.
- [C5] Paul Cupillard, Yann Capdeville, Ming Zhao. “Non-periodic homogenization for the elastic wave equation in 3D”. In: *13ème Congrès Français d’Acoustique*. Le Mans, France, Apr. 2016.
- [C6] Antoine Mazuyer, Richard Giot, Paul Cupillard, Marianne Conin, Pierre Thore. “Stress Estimation in Reservoirs by a Stochastic Inverse Approach”. In: *7th International Symposium on In-Situ Rock Stress*. Tampere, Finland, May 2016.
- [C7] Yder J. Masson, Barbara Romanowicz, Pierre Clouzet, Paul Cupillard, Scott French. “Continental scale waveform tomography using both global and regional data: Application to the North American craton”. In: *EGU General Assembly*. Vienna, Austria, Apr. 2015.
- [C8] Antoine Mazuyer, Marianne Conin, Richard Giot, Paul Cupillard. “Could faults be modelled with elastic materials ?” In: *The Geology of Geomechanics*. London, United Kingdom, Oct. 2015.
- [C9] Laurent Stehly, Pierre Boué, Paul Cupillard. “Computing sensitivity kernels of noise correlations with respect to noise sources”. In: *26th IUGG General Assembly*. Prague, Czech Republic, June 2015.
- [C10] Paul Cupillard, Yann Capdeville, Arnaud Botella. “Numerical homogenization for seismic wave propagation in 3D geological media”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2014.
- [C11] Paul Cupillard, Yann Capdeville. “Upscaling 3D Complex Geological Media for the Elastic Wave Equation”. In: *SIAM Conference on Mathematical and Computational Issues in the Geosciences*. Padova, Italy, June 2013.
- [C12] Andreas Fichtner, Paul Cupillard, Erdinc Saygin, Jeannot Trampert, Tuncay Taymaz, Yann Capdeville. “The deep structure of the North Anatolian Fault Zone”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2013.
- [C13] M. Saade, Jean-Paul Montagner, Philippe Roux, Paul Cupillard, F. Brenguier, Bogdan Enescu, K. Shiomi. “Spatio-temporal changes of seismic anisotropy in seismogenic zones”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2013.

- [C14] Paul Cupillard. “Heterogeneities and Anisotropy in the Earth”. In: *3rd QUEST workshop*. Tatranska-Lomnica, Slovakia, May 2012.
- [C15] Paul Cupillard, Yann Capdeville. “Upscaling small heterogeneities for seismic wave propagation in 3D complex media”. In: *EGU General Assembly*. Vienna, Austria, Apr. 2012.
- [C16] Andreas Fichtner, Jeannot Trampert, Paul Cupillard, Erdinc Saygin, Tuncay Taymaz, Antonio Villasenor. “Imaging the North Anatolian Fault Zone with multi-scale full waveform inversion”. In: *AGU Fall Meeting*. San Francisco, United States, Dec. 2012.
- [C17] Shravan M. Hanasoge, L. Stehly, Tarje Nissen-Meyer, Paul Cupillard. “Non-linear iterative inversions for the distribution of noise sources”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2012.
- [C18] Tarje Nissen-Meyer, Shravan M. Hanasoge, Laurent Stehly, Paul Cupillard. “Non-linear inversions for the origin of ambient noise”. In: *EGU General Assembly*. Vienna, Austria, Apr. 2012.
- [C19] Nian Wang, Jean-Paul Montagner, Yann Capdeville, Gael Burgos, Paul Cupillard. “How to separate intrinsic and artificial anisotropy”. In: *EGU General Assembly*. Vienna, Austria, Apr. 2012.
- [C20] Lucia Zaccarelli, M. Saade, Paul Cupillard, Philippe Roux, Jean-Paul Montagner, F. Brenguier. “Crustal anisotropy viewed by noise cross-correlations”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2012.
- [C21] Paul Cupillard, Yann Capdeville. “Implementation of the homogenization technique for wave propagation in 3D elastic media”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2011.
- [C22] Paul Cupillard, Yann Capdeville. “Implementation of the homogenization technique for wave propagation in 3D elastic media”. In: *2nd QUEST workshop*. Hveragerdi, Iceland, July 2011.
- [C23] Barbara Romanowicz, Huaiyu Yuan, Paul Cupillard. “Refining Upper Mantle Structure in the North American continent using Spectral Element method”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2011.
- [C24] Paul Cupillard, Huaiyu Yuan, Barbara Romanowicz, Yann Capdeville, Jean-Paul Montagner, Gaetano Festa. “RegSEM, a flexible regional Spectral Element code: application to continental scale problems”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2010.
- [C25] Stephanie Durand, Jean-Paul Montagner, Philippe Roux, Florent Brenguier, S. Saumet, Paul Cupillard, Gael Burgos. “Passive monitoring of anisotropy change for the Parkfield 2004 earthquake”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2010.
- [C26] Laurent Stehly, Paul Cupillard, Barbara Romanowicz. “Using simultaneously curvelet filters and SEM simulation of seismic ambient noise: a possible way to improve ambient noise tomography”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2010.
- [C27] Huaiyu Yuan, Paul Cupillard, Scott French, Barbara Romanowicz. “Refining the cratonic upper mantle: modeling North American upper mantle and crustal structure using the Spectral Element method”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2010.
- [C28] Sanne Cottaar, Paul Cupillard, Alen McNamara, Barbara Romanowicz, Rudy Wenk. “Forward modeling the origin of seismic anisotropy at the base of the mantle”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2009.
- [C29] Paul Cupillard, Yann Capdeville. “Recovering the attenuation of surface waves from noise correlation: synthetic tests in a spherically symmetric Earth”. In: *CIDER '09 Community Workshop*. Marconi Center, CA, United States, May 2009.
- [C30] Paul Cupillard, Yann P Capdeville, Laurent Stehly, Jean-Paul Montagner, Barbara Romanowicz. “Spectral Element Simulation of Waveforms Obtained by Correlation of Ambient Seismic Noise”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2009.
- [C31] Laurent Stehly, Paul Cupillard, Barbara Romanowicz. “Numerical simulation of seismic ambient noise in the 10-50s period band using the Spectral Element Method”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2009.
- [C32] Paul Cupillard, Yann Capdeville, Laurent Stehly. “Recovering the Attenuation of Surface Waves From Noise Correlation: Synthetic Tests in a Spherically Symmetric Earth.” In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2008.
- [C33] Laurent Stehly, Paul Cupillard, Aimin Cao, Barbara Romanowicz. “Regional 3D tomography of the upper mantle using a summed source approach”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2008.

- [C34] Elise Delavaud, Paul Cupillard, Gaetano Festa, Jean-Pierre Vilotte. “3D Spectral Element Method simulations of the seismic response of Caracas (Venezuela) basin”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2007.
- [C35] Paul Cupillard, Yann P Capdeville, Nikolai M. Shapiro. “Spectral Element Simulation of Green Functions Obtained by Correlation of Ambient Seismic Noise”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2006.
- [C36] Elise Delavaud, Paul Cupillard, Gaetano Festa, Jean-Pierre Vilotte. “3D Spectral Element Method simulations of the seismic response in complex media: application to the Caracas valley”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2006.
- [C37] Elise Delavaud, Paul Cupillard, Gaetano Festa, Jean-Pierre Vilotte. “3D Spectral Element Method Simulations of the Seismic Response in the Caracas Basin”. In: *Third International Symposium on the Effects of Surface Geology on Seismic Motion*. Vol. 1. Grenoble, France, Aug. 2006, pp. 512–522.
- [C38] Marc Panning, Federica Marone, Ahyi Kim, Yann P Capdeville, Paul Cupillard, Yuancheng Gung, Barbara Romanowicz. “Improvements in mode-based waveform modeling and application to Eurasian velocity structure”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2006.
- [C39] Paul Cupillard, Yann Capdeville. “Tomography based on the spectral element method: the stacked sources method”. In: *2nd SPICE workshop*. Smolenice, Slovakia, Sept. 2005.
- [C40] Jeroen Ritsema, Paul Cupillard, Saskia Goes, Lithgow-Bertelloni C. Stixrude L. “Transition zone structure of California from Ps converted waves”. In: *AGU Fall Meeting*. San Francisco, CA, United States, Dec. 2005.

Chapter 3

Selected articles

In this chapter, I provide the five papers on which I have worked the most among the 21 articles I have been involved in (section 2.4):

- Cupillard, P. and Y. Capdeville (2010). On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach. *Geophys. J. Int.* 181, 1687–1700.
- Cupillard, P., L. Stehly, and B. Romanowicz (2011). The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise. *Geophys. J. Int.* 183, 1397–1414.
- Cupillard, P., E. Delavaud, G. Burgos, G. Festa, J.-P. Vilotte, Y. Capdeville, and J.-P. Montagner (2012). RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale. *Geophys. J. Int.* 188, 1203–1220.
- Cupillard, P. and Y. Capdeville (2018). Non-periodic homogenization of 3-D elastic media for the seismic wave equation. *Geophys. J. Int.* 213(2), 983–1001.
- Renat, Z., P. Cupillard, and Y. Capdeville (2022). Interpretation of time reversal focal spot based on point-source homogenization. *Geophys. J. Int.*, submitted.

On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach

Paul Cupillard* and Yann Capdeville

Département de sismologie, IPGP, 4 place Jussieu, 75252 Paris Cédex 05, France. E-mail: paulcup@ipgp.jussieu.fr

Accepted 2010 March 4. Received 2010 March 3; in original form 2009 September 24

SUMMARY

Cross-correlation of ambient seismic noise recorded by a pair of stations is now commonly recognized to contain the Green's function between the stations. Although traveltimes extracted from such data have been extensively used to get images of the Earth interior, very few studies have attempted to exploit the amplitudes. In this work, we investigate the information contained in the amplitudes and we probe the capability of noise correlations to recover anelastic attenuation. To do so, we carry out numerical experiments in which we generate seismic noise at the surface of a 1-D Earth model. One of the advantages of our approach is that both uniform and non-uniform distributions of noise sources can be taken into account. In the case of a uniform distribution, we find that geometrical spreading as well as intrinsic attenuation are retrieved, even after strong non-linear operations such as one-bit normalization and spectral whitening applied to the noise recordings. In the case of a non-uniform distribution of sources, the geometrical spreading of the raw noise correlations depends on the distribution, but intrinsic attenuation is preserved. For the one-bit noise and whitened noise correlations, the interpretation of observed amplitude decays requires further study.

Key words: Surface waves and free oscillations; Seismic attenuation; Computational seismology; Wave propagation.

1 INTRODUCTION

Recent developments have shown that the Green's function (GF) between two distant seismometers can emerge from the cross-correlation of a sufficient duration of seismic noise recorded at both seismometers (Shapiro & Campillo 2004). This fact provides new data of great interest for seismologists, because they enable imaging of Earth structure in aseismic regions. In most applications, the emerging signal is dominated by surface waves. Group-speeds on interstation paths are now widely measured and numerous high-resolution tomographic images have appeared over the last 5 years (e.g. Shapiro *et al.* 2005; Cho *et al.* 2007; Lin *et al.* 2007; Yang *et al.* 2007; Bensen *et al.* 2008; Stehly *et al.* 2009).

The emergence of the GF between two stations is possible because the noise sources create a spatially and temporally incoherent wavefield that carries, nevertheless, a small coherent part. This small coherent part is due to the sources that produce paths, which reach either one of the receivers via the other. A time-averaged correlation of noise recorded at these receivers enables recovery of this coherency. Prior to its use in seismology, this result was successfully applied in other fields such as helioseismology (Duvall *et al.* 1993) and ultrasonics (Lobkis & Weaver 2001; Weaver & Lobkis 2001, 2003). Many theoretical developments subsequently

explained the phenomenon. Lobkis & Weaver (2001) were the first to describe the incoherent wavefield as a sum of equipartitioned modes and show that correlation of two records of such a field yields the GF. More recently, the wavefield was seen as an uncorrelated and isotropic mix of plane waves from all propagation direction (Weaver & Lobkis 2003). With this definition, Sánchez-Sesma & Campillo (2006) and Sánchez-Sesma *et al.* (2006) show that equipartitioning of the wavefield is a necessary condition for retrieving the exact GF. Other developments, based on an analogy with time-reversal experiments (Derode *et al.* 2003), the fluctuation-dissipation theorem (van Tiggelen 2003) or reciprocity (Wapenaar 2004), and stationary-phase derivation (Snieder 2004), also demonstrate the emergence of GFs from ambient noise correlations.

All these theories only take into account the case of uniformly distributed noise sources. Now, as Sánchez-Sesma & Campillo (2006) assess, 'an anisotropic flux as well as the absence of equipartition has to be considered to fully understand the limitations of the method'. Indeed, noise consistently observed in seismic records mainly comes from the oceans (Longuet-Higgins 1950) so that its distribution at the surface of the Earth clearly is non-uniform. Interest in the effects of such a distribution has recently grown (Weaver *et al.* 2009; Yao & van der Hilst 2009). In this paper, we study these effects by computing correlations of numerically generated seismic noise in an attenuating sphere. We also look at the influence on correlations of two processing techniques commonly applied to the noise records: one-bit normalization and spectral whitening

*Now at: Seismological Laboratory, University of California, Berkeley, CA 94720, USA. E-mail: paulcup@seismo.berkeley.edu

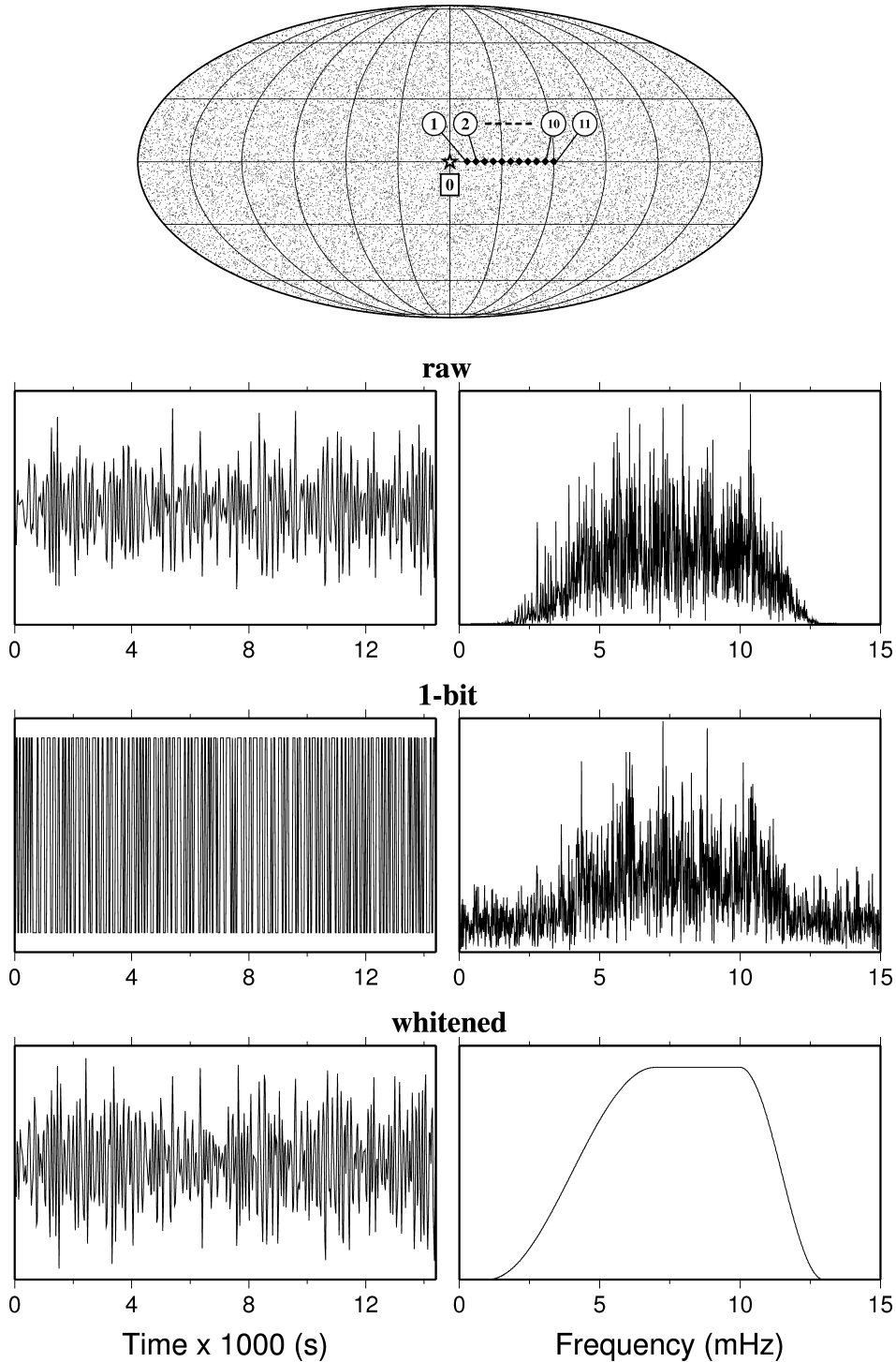


Figure 1. Top panel: source and receiver configuration of our first experiment. Stations $n = 1, \dots, 11$ (black diamonds) are $(n + 1) \times 5^\circ$ far from station 0 (white star). Tiny dot pixels indicate the location of 24 000 noise sources (80 realizations) randomly distributed on the surface of the Earth. Bottom panel: an example of noise recorded at station 0 is presented (left panel) for the three different noise processing we consider (raw, one-bit normalization and whitening). We also plot the amplitude spectrum of these recordings (right panel).

(Bensen *et al.* 2007). One-bit normalization is a procedure for reducing the weight of the earthquakes that inevitably lie in seismic records. It consists of retaining only the sign of the raw signal by replacing all positive amplitudes with a 1 and all negative amplitudes with a -1 . Spectral whitening is often practiced as well to

enhance frequencies with low amplitude. Other noise-processing techniques, not studied in this paper, can be found in Bensen *et al.* (2007).

Traveltimes of ambient noise correlations have been extensively used so far. In this paper, we want to probe the information contained in the amplitude of such data. To do so, we compare our

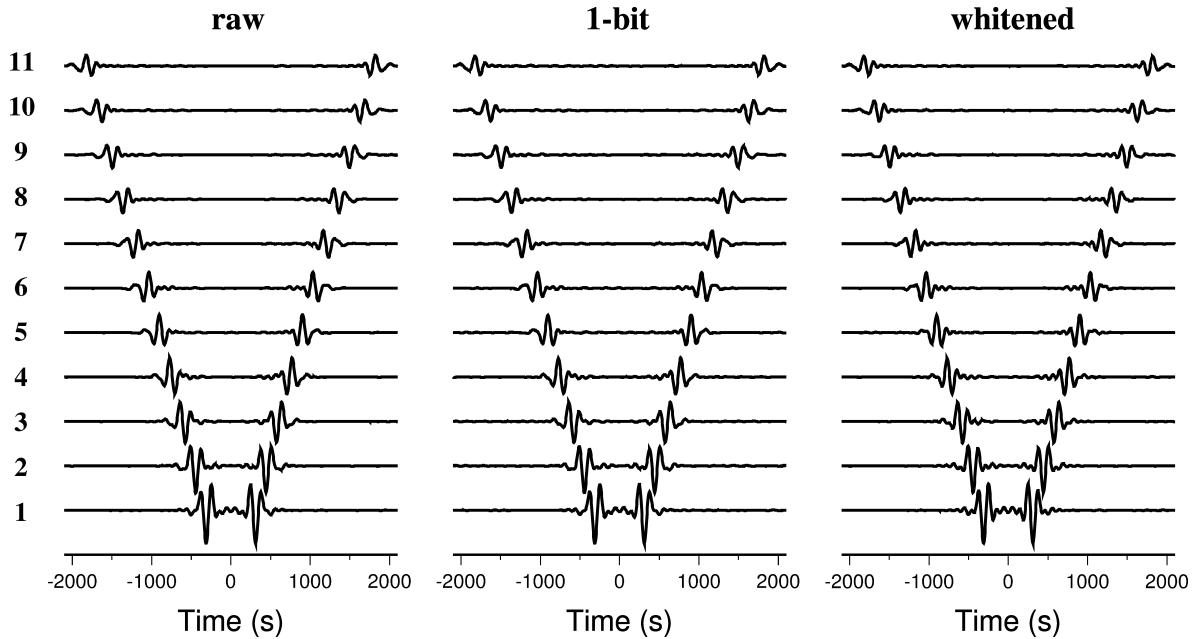


Figure 2. Cross-correlations between the vertical displacement recorded at 0 and the vertical displacement at n (numbers on the left) in the case of a uniform distribution of noise sources. Three data sets are presented corresponding to the three different noise processing investigated in this work: raw, one-bit normalization and whitening.

synthetic correlations with the full waveforms of exact GFs computed with a normal-mode summation technique. The behaviour of our reconstructed Rayleigh waves with respect to the attenuation (geometrical spreading and anelasticity) is addressed by carefully analysing the amplitude decay with increasing interstation distance. Our results are then discussed in the context of previous studies that deal with the same topic (Larose *et al.* 2007; Gouédard *et al.* 2008; Matzel 2008; Prieto *et al.* 2009).

2 UNIFORM DISTRIBUTION OF NOISE SOURCES

2.1 Experimental setup

We compute synthetic noise at the surface of a spherical attenuating Earth to mimic the continuous oscillations that are consistently observed in seismic records and commonly considered to come from interactions between the atmosphere, ocean and seafloor. Such oscillations do not have a flat amplitude spectrum; two peaks generally are present around 14 and 7 s period, which are referred to as primary and secondary microseism, respectively (Longuet-Higgins 1950; Friedrich *et al.* 1998; Stehly *et al.* 2006). Moreover, at longer periods (150–500 s), the so-called Earth ‘hum’ is observed (Nawa *et al.* 1998; Rhie & Romanowicz 2004; Kedar & Webb 2005). In this work, we do not try to simulate the complex mechanisms, which produce this kind of spectra. We will consider simple noise sources with flat spectra in a specified frequency band.

To create our synthetic noise, we randomly position 300 sources on the surface of the Earth. For each spatial component of each source, we generate a 24-hr time-series with random phase and flat spectrum filtered between 100 and 150 s. Using normal-mode summation (e.g. Woodhouse & Girnius 1982) in the Preliminary Reference Earth Model (Dziewonski & Anderson 1981), the effect of all the sources is computed at 12 stations $n = 0, \dots, 11$. All the stations are located on the equator. We distinguish receiver

$n = 0$ at longitude 0° from the other receivers $n = 1, \dots, 11$ at longitude $(n + 1) \times 5^\circ$ (Fig. 1). The correlations between the vertical displacement recorded at 0 and the vertical displacement at the other stations are computed. That is the result of what we call a ‘realization’. Because a single realization is not enough to get a good convergence, we perform 5120 realizations (we use 64 processors; each of them computes 80 realizations; the total number of sources then is 1 536 000). Stacking all the realizations, we obtain the waveforms represented in Fig. 2.

Three different cases are studied corresponding to different processes applied to the noise records: (1) raw noise (no treatment is done); (2) one-bit noise (meaning that only the sign of the waveforms is considered); (3) whitened noise (meaning that the spectral amplitudes of each record are set to 1 in the chosen frequency band). One-bit normalization is widely used to attenuate earthquake signals from real seismic records (e.g. Shapiro & Campillo 2004; Yao *et al.* 2006) or enhance codas relative to first arrivals (e.g. Campillo & Paul 2003; Paul *et al.* 2005). Although there is no earthquake in our experiment, we apply a one-bit normalization because we want to investigate the effect of such a strong non-linear operation on the amplitude of cross-correlations. Frequency whitening is another non-linear operation frequently used to prepare noise prior to correlation. The whitening we adopt in this work is very aggressive: the amplitude spectrum of the original raw noise is replaced by a boxcar smoothed using a cosine-taper. This procedure has been used in previous studies (e.g. Stehly *et al.* 2009). Less aggressive and more sophisticated whitenings are described in Bensen *et al.* (2007).

Fig. 1 shows a 4-hr noise record processed in the three ways mentioned earlier. The effect of the one-bit normalization is clearly visible in the time domain, whereas the whitening effect appears in the frequency domain. Each process provides a set of correlations (Fig. 2). All the correlations are symmetric: the waveforms at negative and positive times are the same in shape and amplitude. This is because the energy flux is the same from station 0 to station n

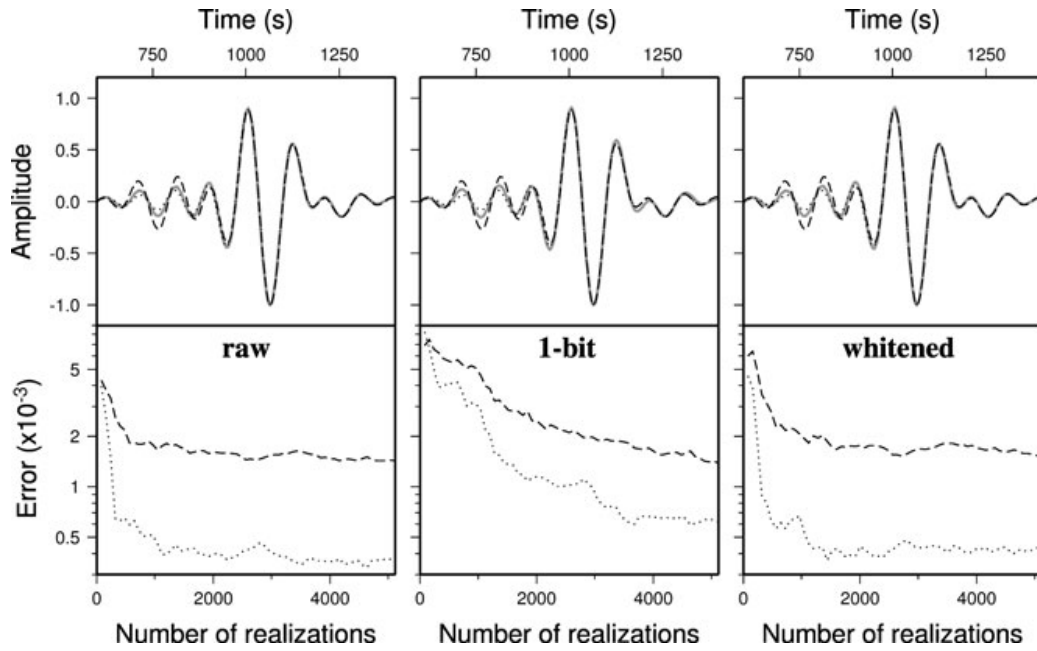
4 *P. Cupillard and Y. Capdeville*

Figure 3. Top panel: for each process applied to the noise records (raw, one-bit normalization and whitening), the correlation between stations 0 and 6 (grey line) is compared with the corresponding Green's function (dashed line) and the fundamental mode (dotted line). To focus on the Rayleigh wave, we use a 800 s time window centred on $t = 1000$ s. The correlations are obtained using 1 536 000 noise sources (5120 realizations). These sources are uniformly distributed on the surface of the Earth. Bottom panel: the dashed (dotted) line is the error between the correlation and the Green's function (the fundamental mode) with respect to the number of realizations. For every process, we converge to a small and stable error. Moreover, the error from the Green's function is systematically larger than the error from the fundamental mode.

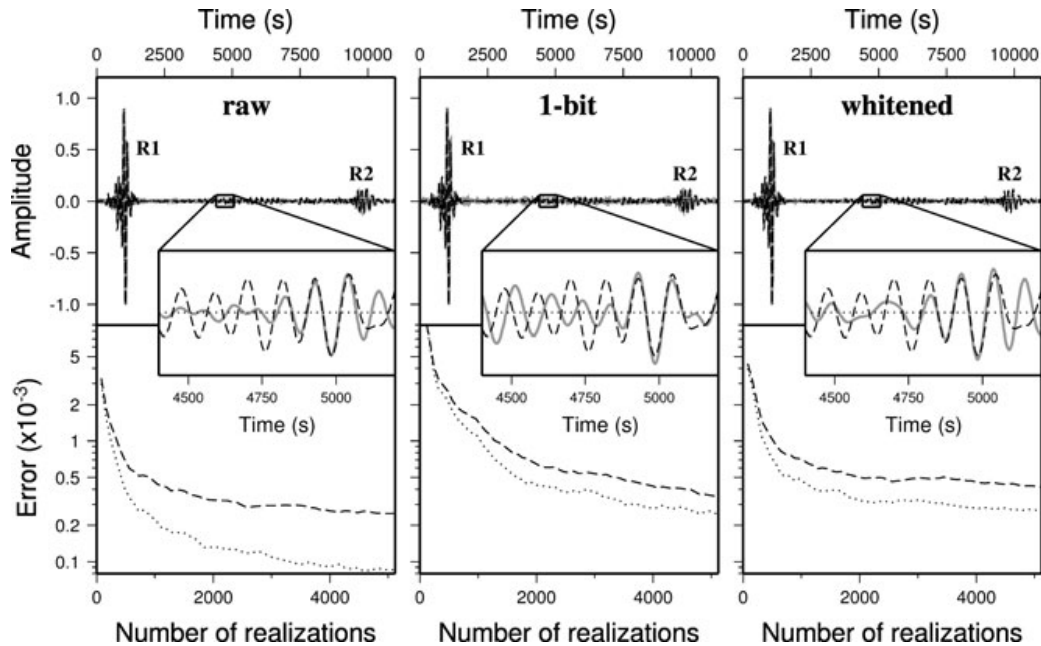


Figure 4. Same as Fig. 3 but the time window is now 11 000 s wide. R1 (the Rayleigh wave travelling along the short arc of the great circle) and R2 (the Rayleigh wave from the long arc) are both visible. Zooming in on a part of the signal where the energy of the fundamental mode is negligible shows that the correlation can retrieve some, but not all, overtones. In the chosen time window, the error from the fundamental mode is very slightly smaller than the error from the Green's function. After a sufficient number of realizations, the two errors are small and stable for the three noise processing.

(negative part) than from n to 0 (positive part). As expected, the arrival time of the wave packets increases as interstation distance becomes larger. On the contrary, the amplitude decreases as interstation distance increases. This decay will be studied more precisely at the end of Section 2.2.

2.2 Comparison with the Green's function

The Green's tensor between two stations in a laterally homogeneous sphere can be easily computed using normal-mode summation. Following Woodhouse & Girnius (1982), we can write the vertical

displacement G_{nm} at a station m due to a vertical Dirac delta function at another station n as a function of time t

$$G_{nm}(t) = \sum_k a_k(n) u_k(m) e^{i\omega_k t}, \quad (1)$$

where u_k and ω_k are the eigenfunctions and eigenfrequencies of the Earth model, and a_k are the excitation coefficients due to the vertical Dirac delta function. Using this equation, we are able to calculate the reference GFs against which we can compare our correlations. To make the comparison possible, we have to differentiate the correlations in time and correct them for a source term, which is the power spectral density of the noise, recorded at one station. Indeed, different authors (Lobkis & Weaver 2001; Snieder 2004; Colin de Verdière 2006; Sánchez-Sesma *et al.* 2008) have demonstrated that, in 3-D, we have

$$i\omega C_{nm}(\omega) \propto |S_n(\omega)|^2 G_{nm}(\omega), \quad (2)$$

where C_{nm} is the side of the correlation between n and m that corresponds to the noise going from n to m , ω is the angular frequency and S_n is the noise recorded at n . $||$ denotes the modulus of a complex quantity.

Fig. 3 shows a comparison of the corrected correlations from station 6 with the corresponding GF on the one hand, and with the fundamental mode (eq. 1 restricted to $k \in \{\text{fundamental mode}\}$) on the other hand. We chose a 800 s time window centred on $t = 1000$ s to focus on the Rayleigh wave. For each kind of noise processing (raw, one-bit normalization and whitening), the correlation fits both the GF and the fundamental mode very well. Nevertheless, the computation of the mean square deviation of the full waveform (hereafter referred to as the error) shows that the fit is a little better for the fundamental mode. This is mainly due to the discrepancy between the GF and the correlation in the 650–850 s time window: it looks like the correlation is not able to retrieve overtones. This is confirmed in Fig. 4. This figure shows the same comparison as in Fig. 3 but using a much larger time window (11 000 s). Zooming in a part of the signal that has overtones only (the curve of the fundamental mode is equal to zero), we see that some of them are recovered whereas some others are not. This means that the correlation of a seismic waveform produced by noise sources, which are uniformly distributed on the surface of a 1-D Earth, does not perfectly converge to the GF. The convergence of such a correlation towards the fundamental mode is better (this is visible on both Figs 3 and 4) but is not perfect either. Nevertheless, the quality of our reconstructed Rayleigh waves is very good for our purpose.

The reason why some overtones are not retrieved in our correlations is that there are no noise sources at depth in our experiment. Including such sources enables to recover all the overtones, as shown in Fig. 5(a). This figure presents the raw noise correlation between stations 0 and 6 when using sources everywhere in the medium. As done before, we use two time windows to compare this correlation with the GF and the fundamental mode. For each of these two windows, we compute the error as a function of realizations. Because a good convergence requires more realizations than in the case of sources distributed on the surface, we performed 9600 realizations. It is clear that the correlation now fits the GF better than the fundamental mode. Nevertheless, the error between the correlation and the GF stabilizes to a non-zero value. This is because the amplitudes are not correctly retrieved (see the zoom window in the overtones). As stated by Snieder (2007), getting an equipartitioned wavefield (and thus an accurate reconstruction of the GF) in an attenuating medium requires a source excitation that is proportional to the local dissipation rate. We do not use such a particular excitation here so

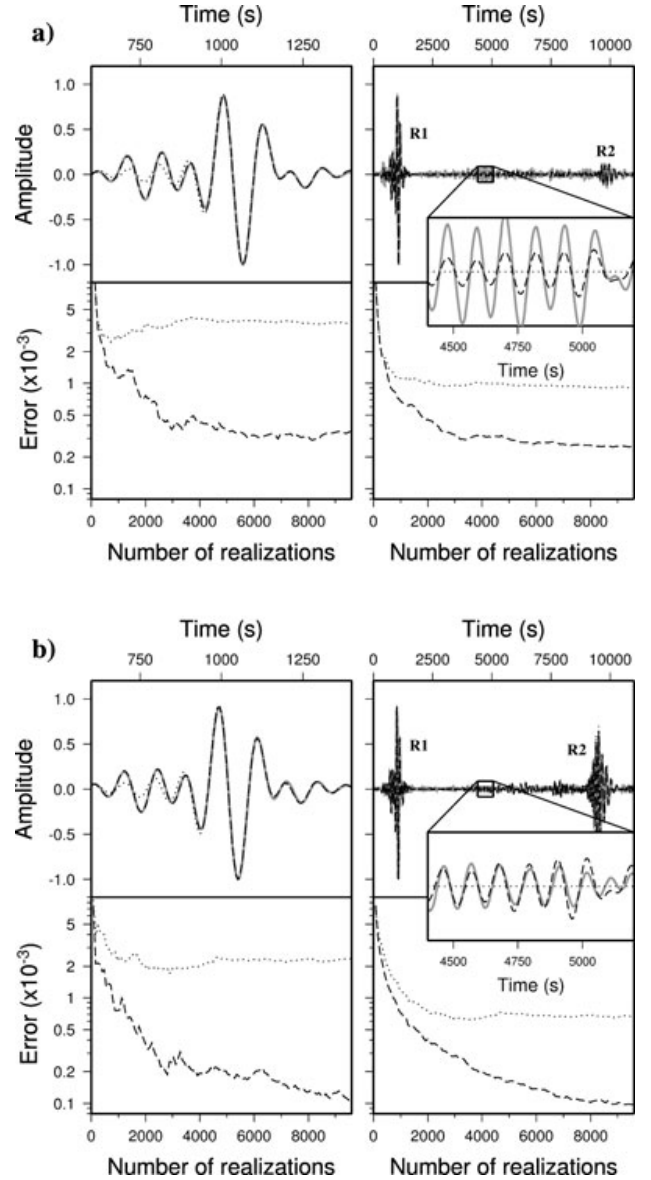


Figure 5. Raw noise correlation between stations 0 and 6 (grey line) emerging from a uniform distribution of noise sources in the interior of the Earth. Two different experiments are carried out: one in an attenuating medium (a) and one in a lossless medium (b). The result of each experiment is compared with the GF (dashed line) and the fundamental mode (dotted line) using two different time windows. The error as a function of realizations is computed for each of these two windows. In both experiment, the correlation fits the GF better than the fundamental mode. In the attenuating Earth, the amplitudes are not very well recovered, so the error between the correlation and the GF stabilizes to a non-zero value. In the lossless Earth, the GF is very well retrieved.

we do not get the exact GF. We check that the anelasticity is indeed the cause of the discrepancy between the GF and the correlation by performing the same experiment in a lossless medium. The result shows that the GF is accurately recovered in this case (Fig. 5b). Similar observations are found for one-bit noise and whitened noise correlations.

The purpose of this paper is to study the amplitude of surface waves reconstructed by noise correlation in the context of noise sources located on the surface of the Earth. Therefore, we do not

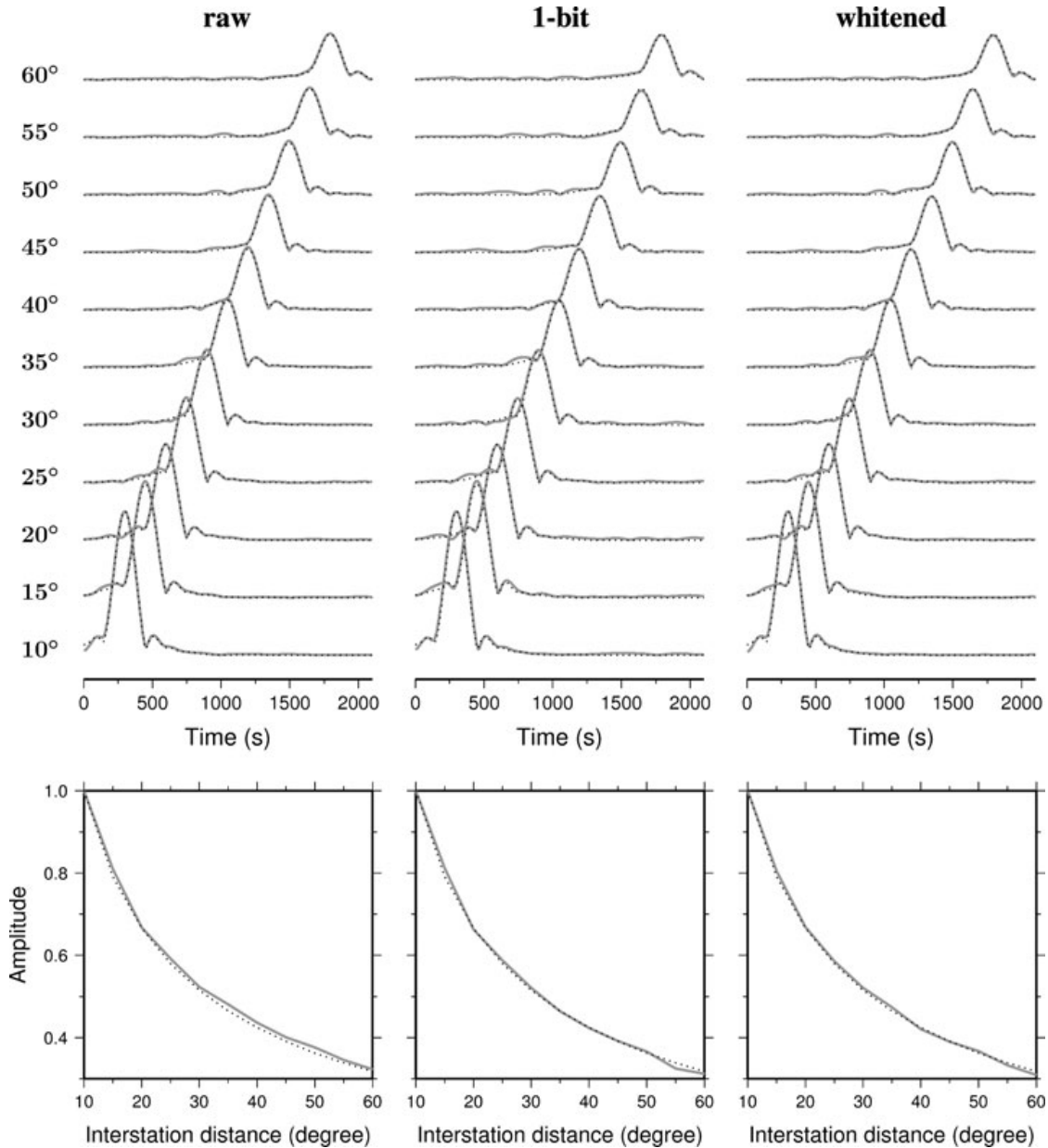


Figure 6. Comparison of the amplitude decay of the envelope of the correlations (grey line) with their corresponding fundamental mode (dotted line) in the case of a uniform distribution of noise sources. Top panel: The waveforms from the three processes (column) and all the stations (row) are presented. Bottom panel: The maximum of amplitude is plotted as a function of interstation distance.

consider the correlations obtained using sources at depth in what follows. As shown in Fig. 3, the quality of the Rayleigh waves obtained using sources on the surface is very good. Because the convergence of these waves towards the fundamental mode is better than the convergence towards the GF, we choose the fundamental mode as the reference in what follows and we compare the amplitude decay of our reconstructed Rayleigh waves with the decay of the fundamental mode (Fig. 6). We match the amplitude maximum of the signals at station 1 so we can see how the relative amplitudes evolve as interstation distance increases. We use the envelope of the signals because it is closely related to the energy, and thus to the attenuation (Aki & Richards 1980). We observe that the signal en-

velopes are very similar for each interstation distance and regardless of the noise processing. Plotting the amplitude maximum as a function of interstation distance, we see that the decay of the correlation is the same as that of the fundamental mode. This means that, in the context of a radially symmetric Earth, cross-correlations contain geometrical spreading as well as intrinsic attenuation, even after applying strong non-linear operations on the amplitude of the noise like one-bit normalization and frequency whitening. As the Rayleigh wave Green's function is highly dominated by the fundamental mode, the results and the conclusion would be the same if considering the GF as the reference.

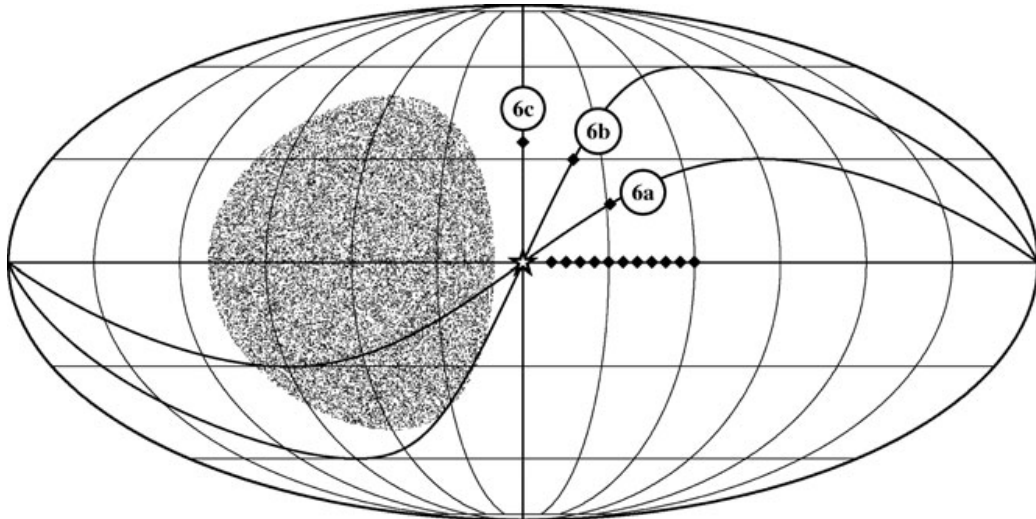


Figure 7. Configuration of the second experiment. Tiny dot pixels indicate the location of 24 000 noise sources (80 realizations). We see that all the sources belong to a 50° -radius patch centred on the equator at longitude -60° . Again, we use station 0 (star) and stations $n = 1, \dots, 11$ (diamonds) as introduced in Fig. 1. We add three new stations (6a, 6b and 6c) to study the anisotropy of the energy flux produced by such a distribution of noise sources.

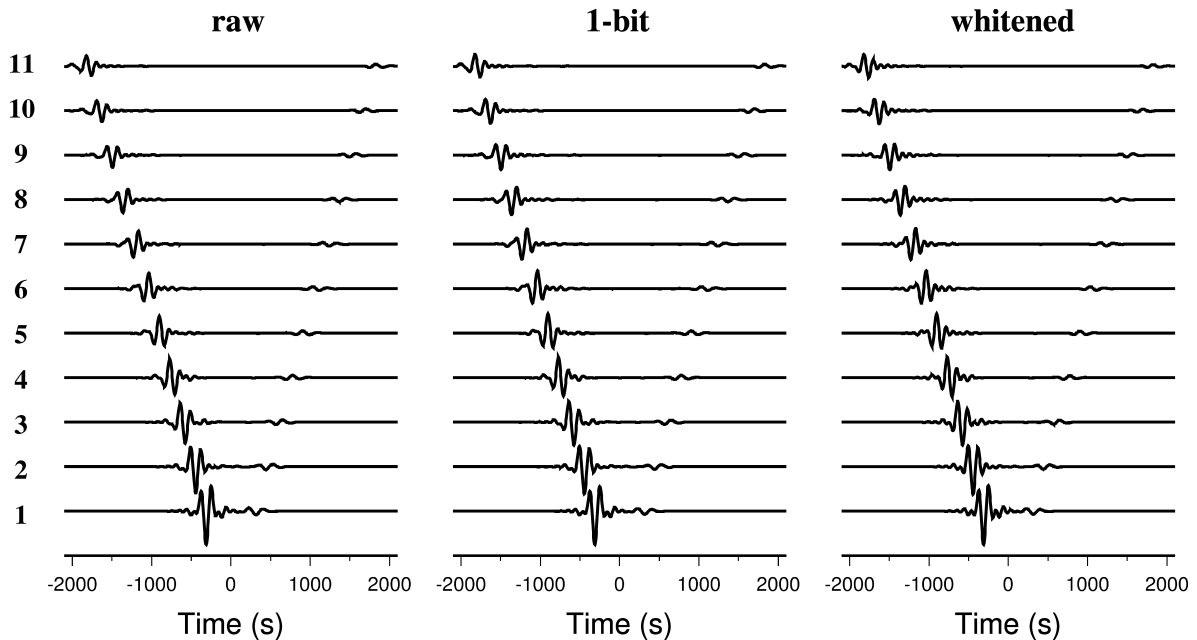


Figure 8. Cross-correlations between the vertical displacement recorded at 0 and the vertical displacement at n (numbers on the left) in the case of noise sources distributed in a big patch.

3 NOISE SOURCES DISTRIBUTED IN A BIG PATCH

3.1 Experimental setup

It is important to study the case of a non-uniform distribution of sources because seismic noise in the Earth mainly comes from the oceans and is therefore non-uniformly distributed. Such a distribution produces an anisotropic energy flux that is difficult to consider analytically and, therefore, has never been clearly taken into account. We probe the effects of wavefield inhomogeneity, by repeating the same experiment as the one in the previous section, except that the

noise sources are now confined in a 50° -radius cap on the surface of the sphere. The centre of this cap is on the equator at longitude -60° (Fig. 7). Reducing the source location area and keeping the same number of realizations increases the source density. To investigate the impact of our space-limited ‘noisy’ area on cross-correlations, we use again a line of 12 stations ($n = 0, \dots, 11$) along the equator. As the azimuths from station 0 are not all equivalent with respect to the patch of noise sources, we add three new stations that are not on the equator, 35° far from station 0 (as station 6 is), and call them 6a, 6b and 6c. They lie at an azimuth of 60° , 30° and 0° from receiver 0.

Fig. 8 shows the correlations between the displacement recorded at station 0 and the displacement from the other stations positioned

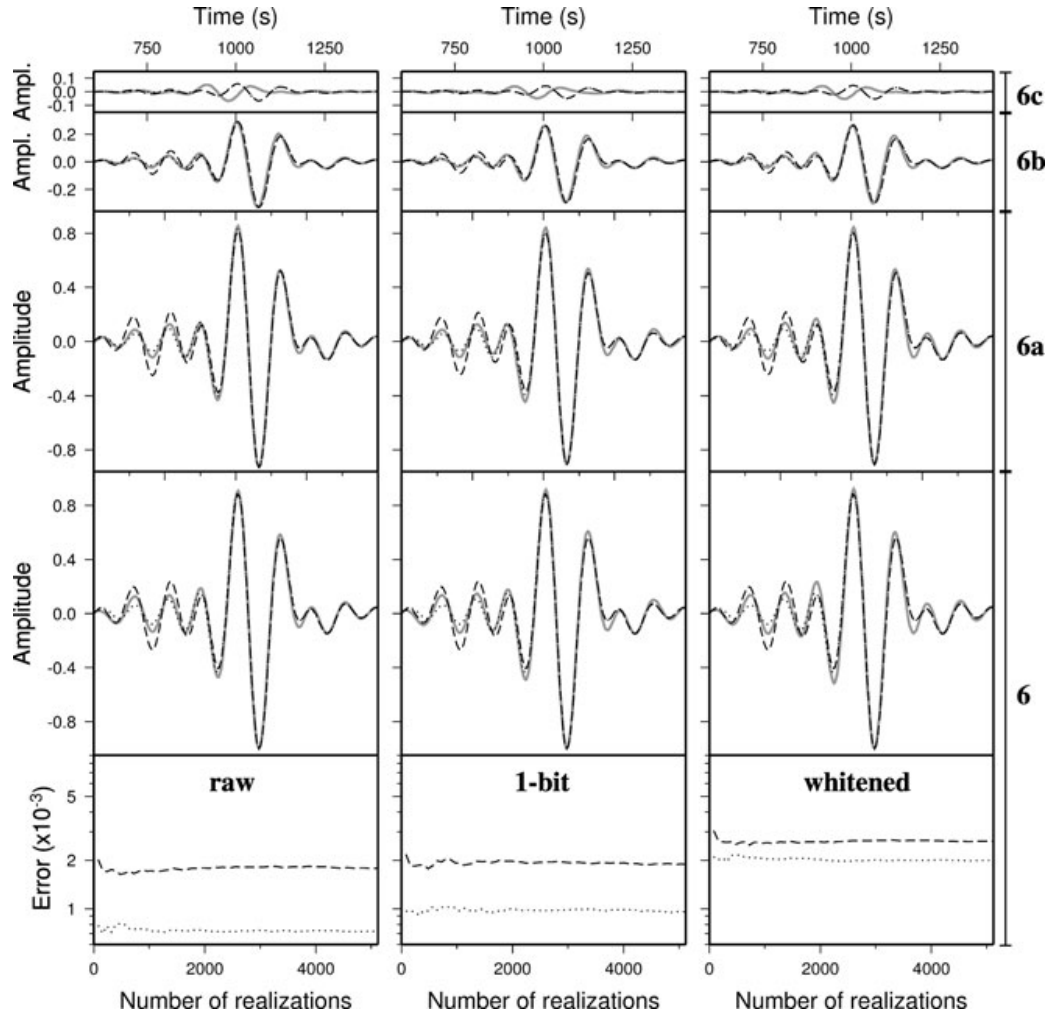


Figure 9. The correlations (grey line) from stations 6c, 6b, 6a and 6 (from the top to the bottom) are compared with the Green's function (dashed line) and the fundamental mode (dotted line) which are the same for the four stations because these latter are all 35° far from station 0. The relative amplitude of the correlations is preserved (and so we adapt the Green's function and the fundamental mode to match the amplitude of the correlations). For station 6, we also plot the convergence as we did in Fig. 3.

on the equator ($n = 1, \dots, 11$). Strong asymmetry appears at all stations and for all processes applied to the noise. This is due to the anisotropy of the energy flux: the energy going from n to 0 is much lower than the energy going from 0 to n . This effect has been reported by many authors working on real data (Shapiro & Campillo 2004; Paul *et al.* 2005; Stehly *et al.* 2006; Yao *et al.* 2006; Nishida & Fukao 2007; Bensen *et al.* 2008). In the following, we will only consider waveforms at negative times (emerging from the noise going from 0 to n). We observe that their amplitude decreases as interstation distance increases. We will study this amplitude decay in the next section.

3.2 Comparison with the Green's function

Using eq. (2), we process our correlations to compare them with the GF. Fig. 9 shows the Rayleigh wave from station 6. As we did before in Fig. 3, we consider the fit to the GF as well as the fit to the fundamental mode. Plotting the error as a function of the number of realizations, we observe that the correlation fits the fundamental mode better than the GF. This is true for all three processes applied

to the noise. We notice that the error becomes stable faster than it did in the previous experiment (Fig. 3). This is simply because the density of sources in the coherent zone, which is the vicinity of the great circle passing through the two stations 0 and 6 (Snieder 2004; Roux *et al.* 2005), is higher in the current experiment. We also notice that the final errors we get are larger than before. This means that the wavefield produced by the current distribution of noise sources is less equipartitioned than the wavefield due to a uniform distribution. Nevertheless, the quality of the reconstructed signal is satisfactory.

Fig. 9 also shows the correlations from stations 6a, 6b and 6c. The relative amplitude of these correlations has been preserved to clearly show that the energy flux is now anisotropic. The amplitude of the reconstructed signal increases with increasing the azimuth. This is because the number of noise sources in the coherent zone is not the same in all azimuths. There are no sources on the great circle defined by 0 and 6c (azimuth = 0°, cf. Fig. 7) so the correlation from 6c has a very small amplitude. On the contrary, the number of noise sources is high for stations 6 and 6a (azimuth = 90° and 60°, respectively) so the corresponding correlations have a large amplitude.

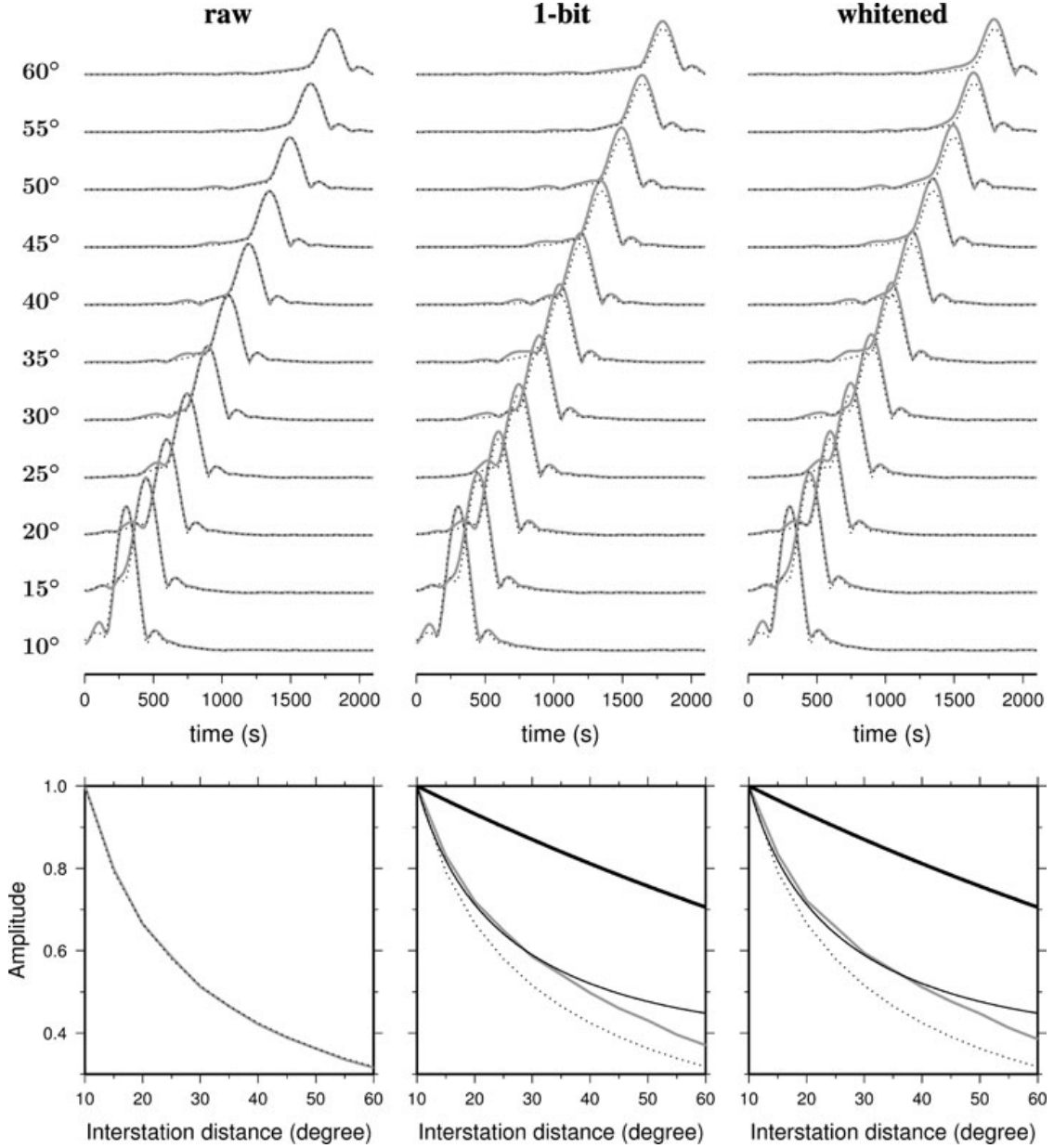


Figure 10. Comparison of the amplitude decay of the envelope of the correlations (grey line) with their corresponding fundamental mode (dotted line) in the case of noise sources distributed in a big patch. Top panel: The waveforms from the three processes (column) and all the stations on the equator (row) are presented. Bottom panel: The maximum of amplitude is plotted as a function of interstation distance. The decay of the one-bit and the whitened noise correlations does not fit the decay of the fundamental mode. It also fits neither geometrical spreading (thin black line) nor intrinsic attenuation (thick black line).

Besides, we note that the phase of the correlation from station 6c has nothing to do with the GF. This is because the noise sources that contribute to the emergence of this correlation are far from the vicinity of the great circle path resulting in erroneous traveltimes. This effect is not explicitly discussed but is substantially contained in the theoretical development of Roux *et al.* (2005). It has been also observed on real data (Stehly *et al.* 2007; Gouédard *et al.* 2008).

We now look at the amplitude decay of the correlations obtained from stations located at the equator. We compare the envelope of these correlations with the envelope of the fundamental mode and we plot the maximum of these envelopes as a function of inter-

station distance (Fig. 10). As was the case in the previous experiment, the decay of the raw noise correlation is the same as that of the fundamental mode. However, for the one-bit and the whitened noise correlations, this is no longer true; the amplitude decay corresponding to these two processes is less steep than the decay of the fundamental mode. We also plot the two terms that correspond to the two different attenuation processes. Writing the amplitude spectrum of the GF restricted to fundamental mode between station 0 and station n as

$$\left| G_{0n}^{k \in \text{fund}}(\omega) \right| \propto \frac{1}{\sqrt{\sin(d_n)}} \exp \left[-\frac{\omega d_n}{2c(\omega)Q(\omega)} \right], \quad (3)$$

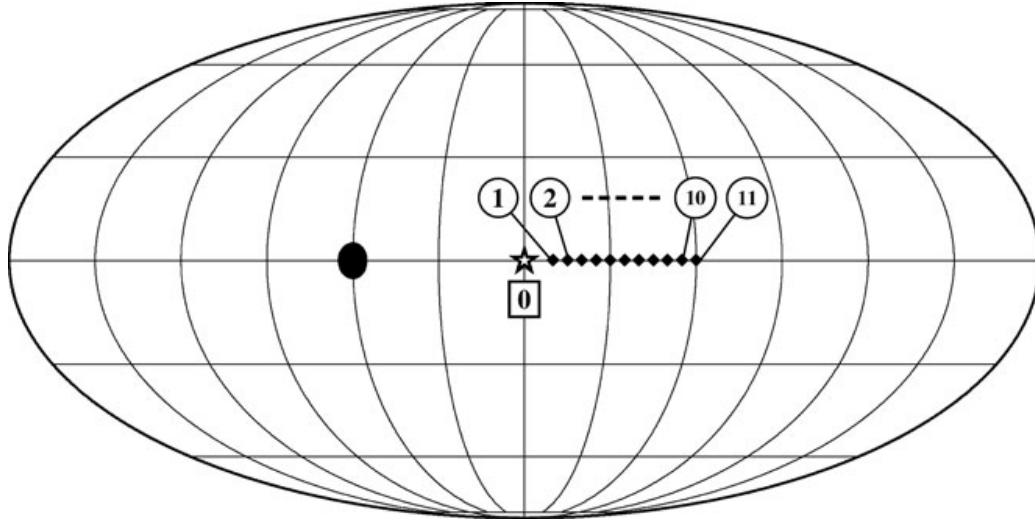


Figure 11. Configuration of the third experiment. The big black dot indicates the cap where the noise sources are. All the sources belong to a 5° -radius patch centred on the equator at longitude -60° . We use twelve receivers on the equator: station 0 (star) and stations $n = 1, \dots, 11$ (diamonds).

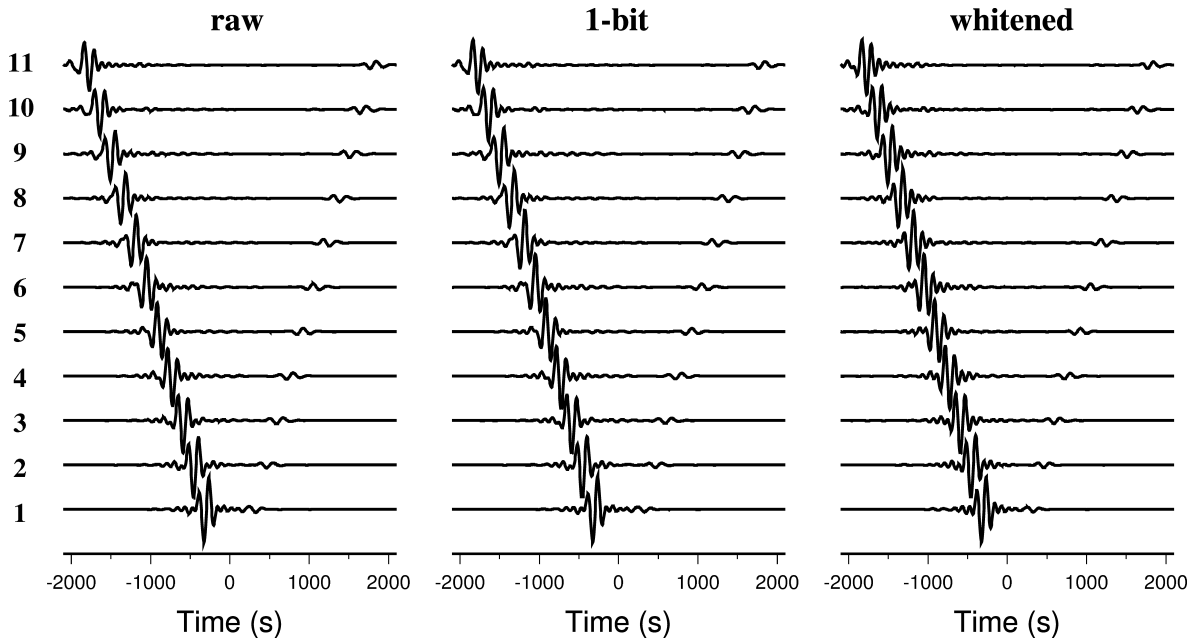


Figure 12. Cross-correlations between the vertical displacement recorded at 0 and the vertical displacement at n (numbers on the left) in the case of noise sources distributed in a small patch.

where d_n is the angular distance between the two stations, c is the Rayleigh wave speed and Q is the quality factor of the medium, we recognize the first term of the product to be the geometrical spreading and the second term to be the intrinsic attenuation. The plot of these two terms in Fig. 10 shows that they both provide a significant contribution to the amplitude of the fundamental mode. The amplitude decay of the one-bit and the whitened noise correlations does not fit any of the two terms. One can reasonably think that the one-bit and the whitened data sets both contain information on the intrinsic attenuation of the medium. However, in the case of the distribution of noise sources studied in this section, it is not possible to treat these data sets as one treats usual records (such as earthquake records) to extract the intrinsic attenuation.

4 NOISE SOURCES DISTRIBUTED IN A SMALL PATCH

4.1 Experimental setup

To investigate extreme source inhomogeneity, one last experiment is carried out. We now distribute the noise sources in a 5° -radius patch centred on the equator at longitude -60° (Fig. 11). The 12-stations line along the equator is used again to compute correlations. As expected, the obtained waveforms are asymmetric (Fig. 12). Moreover, for all three treatments used to process the noise records, we observe that the amplitude decay of the correlations with increasing interstation distance looks less strong than observed in the two

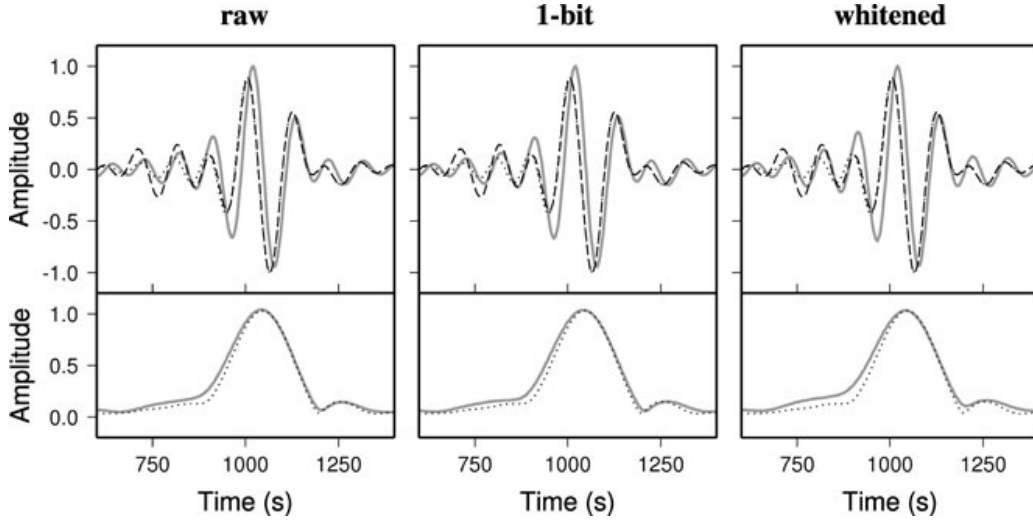


Figure 13. The correlations (grey line) from station 6 are compared with the Green's function (dashed line) and the fundamental mode (dotted line). The noise sources producing the wavefield that we correlate are confined in a small patch. Top panel: A $\pi/4$ phase shift appears between the correlations and the two other curves. Bottom panel: The phase shift is no longer visible when using the envelope of the signals.

previous experiments. We will carefully look at this variation of amplitude in the next section.

4.2 Comparison with the Green's function

There is substantial misfit between the corrected correlation (eq. 2) from station 6 and the corresponding GF and fundamental mode predictions (Fig. 13). This misfit mainly comes from a $\pi/4$ phase shift, which is known to make phase velocity measurements from noise correlations very ambiguous (Harmon *et al.* 2008). Indeed, in the general case (i.e. the case of any given distribution of noise sources), it is not clear whether the extracted GF is 2-D or 3-D. The main difference between these two kinds of function precisely is a $\pi/4$ shift (Aki & Richards 1980; Sánchez-Sesma & Campillo 2006). In spite of the ambiguity, Yao *et al.* (2006) and Lin *et al.* (2008) have succeeded in extracting phase velocity maps from noise correlations. In this work, the $\pi/4$ phase shift is not a problem because we use the envelope of the signals to study variations in amplitude. As shown in Fig. 13, the envelope of the correlation is very similar to the envelope of the fundamental mode, confirming the robustness of group velocity measurements from noise correlations.

Fig. 14 shows how the amplitude of the correlations decreases with increasing interstation distance. For each process, the decay is very different from the decay of the fundamental mode. We think this difference could be due to different geometrical spreading. Indeed, the small patch can be viewed as a point source providing a wavefield that is coherent between station 0 and station n (all the paths going through 0 also go through n) such that

$$|C_{0n}(\omega)| = |S_0(\omega)||S_n(\omega)| \quad (4)$$

with

$$|S_0(\omega)| = \frac{A(\omega)}{\sqrt{\sin(D)}} \exp\left[-\frac{\omega D}{2c(\omega)Q(\omega)}\right] \quad (5)$$

and

$$|S_n(\omega)| = \frac{A(\omega)}{\sqrt{\sin(D+d_n)}} \exp\left[-\frac{\omega(D+d_n)}{2c(\omega)Q(\omega)}\right], \quad (6)$$

where A is the amplitude spectrum of the signal at the source and D is the distance in degrees between station 0 and the point source

(60° in our simulation). We here assume that noise propagates as surface waves. Substituting eqs (5) and (6) in eq. (4) gives

$$|C_{0n}(\omega)| = \frac{\alpha(\omega)}{\sqrt{\sin(D+d_n)}} \exp\left[-\frac{\omega d_n}{2c(\omega)Q(\omega)}\right], \quad (7)$$

where $\alpha = |S_0(\omega)|^2 \sqrt{\sin(D)}$ is a constant with respect to d_n . From expression (7), we see that intrinsic attenuation acts in the correlation as it does in the GF. This is not true for geometrical spreading: the term corresponding to this attenuation in eq. (7) is different from the one in eq. (3). Correcting the amplitude decay of the raw noise correlation using $\sqrt{\sin(D+d_n)}/\sin(d_n)$ confirms our theory: the decay of the fundamental mode is retrieved, meaning that raw noise correlations can be used to measure the attenuation of the medium if the distribution of noise sources is known. This is a major result because it shows that meaningful information can be extracted from the amplitude of raw noise correlations. The same correction applied to the one-bit and the whitened noise correlations does not enable to recover the good amplitude decay. Similarly to the previous experiment, further work is needed to understand what these decays exactly contain.

5 DISCUSSION AND CONCLUSIONS

Records of synthetic noise computed in PREM using the normal-mode summation method enabled us to investigate different features of waveforms obtained by noise correlation. We paid particular attention to the relative amplitude of such waveforms. We find that this amplitude strongly depends on the distribution of noise sources and the processing applied to the noise records prior to correlation. If the sources are uniformly distributed on the surface of the Earth, then the amplitude decay we observe along a line of stations is the same as the decay of the Rayleigh wave Green's function. This is true regardless of the noise processing technique. If the distribution is not uniform, then the amplitude decay depends on the processing. Whereas raw noise correlations contain the intrinsic attenuation of the medium and geometrical spreading that is determined by the position of the noise sources, one-bit noise and whitened noise correlations result in decays that do not fit that of the

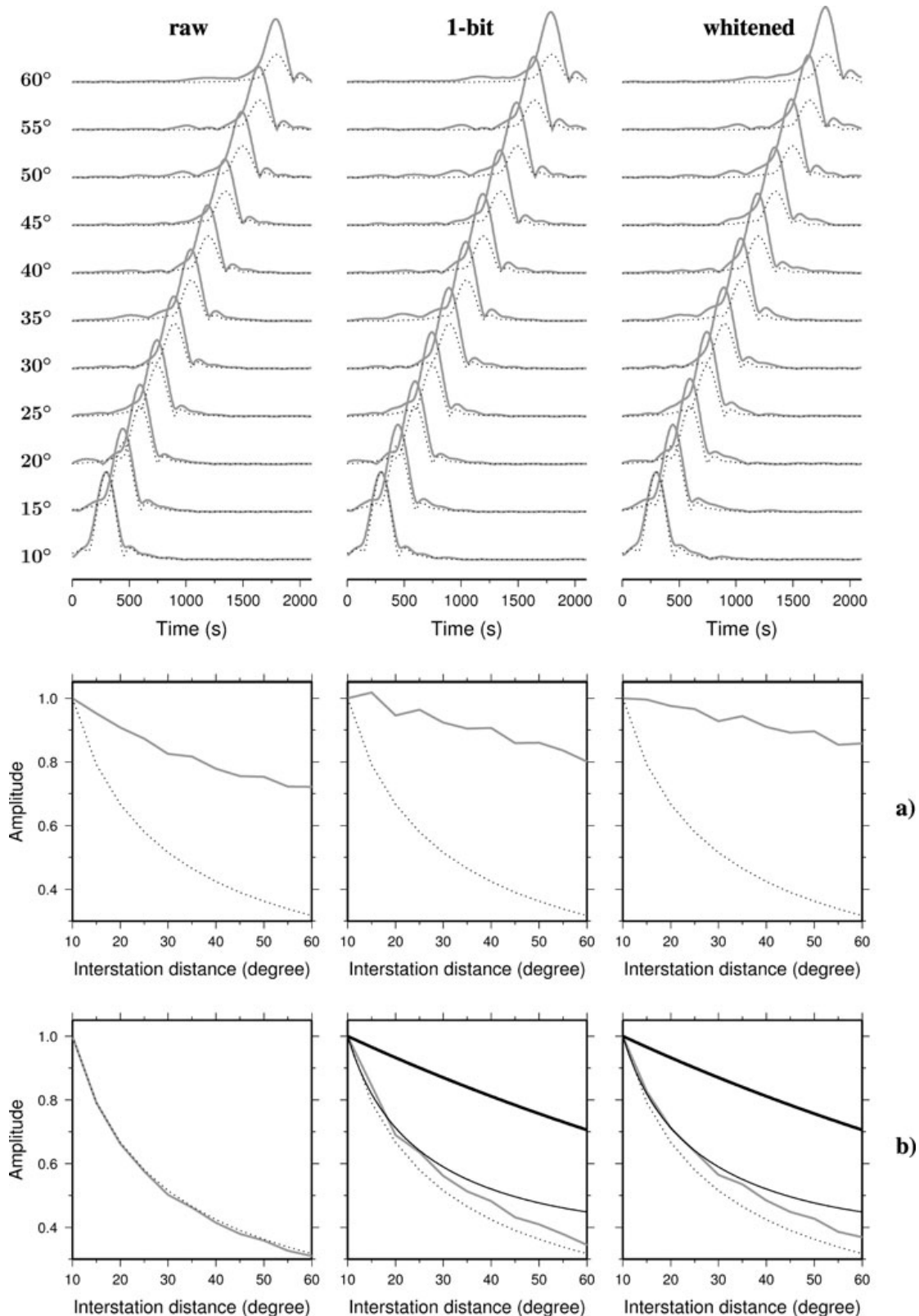


Figure 14. Comparison of the amplitude decay of the correlations envelope (grey line) with their corresponding fundamental mode (dotted line) in the case of noise sources distributed in a small patch. Top panel: Waveforms from the three processes (column) and all the stations (row) are presented. Bottom panel: (a) The maximum of amplitude is plotted as a function of interstation distance. (b) A geometrical correction is applied to the correlations. The decay of the raw noise correlation then fits the decay of the fundamental mode whereas the one-bit and the whitened noise correlations do not fit any decays (geometrical and intrinsic attenuations of the fundamental mode are plotted using thin and thick black lines, respectively.)

Rayleigh wave. These decays require further study to be properly interpreted.

Our results are consistent with previous studies. Larose *et al.* (2007) generates noise using a can of compressed air sprayed on the surface of a plexiglass plate. The amplitude of the correlations they acquire from an array of sensors starting at the air-jet source fits the amplitude of waveforms obtained from an active experiment. This is no longer true when frequency whitening is performed. Furthermore, the analysis of a seismic prospecting data set by Gouédard *et al.* (2008) shows that geometrical spreading retrieved by correlating direct waves (i.e. non-scattered waves) depends on the position of the sources. This observation indicates that the distribution of noise sources has to be known to measure the attenuation of the Earth from noise correlations. At present, our knowledge of this distribution is too poor, but we think that it will be greatly improved as localization of microseismic noise sources becomes a topic of intense research (e.g. Gerstoft & Tanimoto 2007; Kedar *et al.* 2008; Gerstoft *et al.* 2008).

A first important limitation in our experiment is that we do not have any incoherent noise (such as local or acquisition-related noise) or non-stationary phases (such as earthquakes) in our synthetics. These signals are often present in real data and can affect the amplitude of the correlations. Including such signals in our simulation could change our results. For instance, Figs 3 and 4 show that the number of realizations needed to get a stable error is larger for the one-bit than for the raw and the whitened noise correlation. This is the opposite of what is observed in practice in real data (e.g. Larose *et al.* 2004) because the one-bit normalization is usually employed to process incoherent noise or non stationary phases. As we have no such signals in our experiment, we do not take advantage of this process and just lose information, making the convergence slower.

In this work, the ratio of interstation distance to wavelength ranges from 2 to 15, as is the case for the ratio involved in correlations of real seismic noise. For this reason, we believe that our results are valid when using real seismic records. Nevertheless, our numerical experiment has a second important limitation: it involves long-period surface waves in a 1-D model whereas real seismic noise mostly propagates at high frequency (0.05–0.2 Hz) and therefore is very sensitive to strong 3-D structures present in the crust. Such structures scatter the incident wavefield and become secondary noise sources, making the distribution much more uniform and the wavefield much more equipartitioned than in the 1-D case (Hennino *et al.* 2001; Gouédard *et al.* 2008). Because of this phenomenon, the attenuation of the Earth may be easier to recover from noise correlations without knowing the precise location of the primary sources (Matzel 2008; Prieto *et al.* 2009). The great potential of these data in investigating the interior of the Earth is confirmed, although significant effort is still needed to improve our understanding and better extract useful attenuation constraints.

ACKNOWLEDGMENTS

The authors thank Laurent Stehly, Vedran Lekic, Barbara Romanowicz, Jean-Paul Montagner and Nikolai M. Shapiro for fruitful discussions. The authors also thank two anonymous reviewers for comments that helped to improve the paper. We are grateful to the *Service de Calcul Parallèle* of the IPGP and to David Weissenbach for his help in the implementation of our code on the EGEE grid.

REFERENCES

Aki, K. & Richards, P., 1980. *Quantitative Seismology: Theory and Methods*. Freeman, San Francisco.

- Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophys. J. Int.*, **169**, 1239–1260, doi:10.1111/j.1365-246X.2007.03374.x.
- Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2008. Broad-band ambient noise surface wave tomography across the united states. *J. geophys. Res.*, **113**, B05306, doi:10.1029/2007JB005248.
- Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda. *Science*, **299**, 547–549.
- Cho, K., Herrmann, R.B., Ammon, C.J. & Lee, K., 2007. Imaging the upper crust of the Korean peninsula by surface-wave tomography. *Bull. seism. Soc. Am.*, **67**, 198–207.
- Colin de Verdière, Y., 2006. Mathematical models for passive imaging, I: general background, <http://arxiv.org/abs/math-ph/0610043/>.
- Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M. & Fink, M., 2003. Recovering the Green's function from field-field correlations in an open scattering medium (L). *J. acoust. Soc. Am.*, **113**, 2973–2976.
- Duvall, T.L., Jefferies, S.M., Harvey, J. & Pomerantz, M.A., 1993. Time-distance helioseismology. *Nature*, **362**, 430–432.
- Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model. *Phys. Earth planet. Inter.*, **25**, 297–356.
- Friedrich, A., Krüger, F. & Kluge, K., 1998. Ocean-generated microseismic noise located with the Gräfenberg array. *J. Seismol.*, **2**, 47–64.
- Gerstoft, P. & Tanimoto, T., 2007. A year of microseisms in southern California. *Geophys. Res. Lett.*, **34**, L20304, doi:10.1029/2007GL031091.
- Gerstoft, P., Shearer, P.M., Harmon, N. & Zhang, J., 2008. Global P, PP, and PKP wave microseisms observed from distant storms. *Geophys. Res. Lett.*, **35**, L23307, doi:10.1029/2008GL036111.
- Gouédard, P., Roux, P., Campillo, M. & Verdel, A., 2008. Convergence of the two-points correlation function toward the Green's function in the context of a prospecting dataset. *Geophysics*, **73**(6), V47–V53.
- Harmon, N., Gerstoft, P., Rychert, C.A., Abers, G.A., de la Cruz, M.S. & Fischer, K.M., 2008. Phase velocities from seismic noise using beamforming and cross correlation in Costa Rica and Nicaragua. *Geophys. Res. Lett.*, **35**, L19303, doi:10.1029/2008GL035387.
- Hennino, R., Tégouères, N., Shapiro, N.M., Margerin, L., Campillo, M., van Tiggelen, B. & Weaver, R., 2001. Observation of equipartition of seismic waves. *Phys. Rev. Lett.*, **85**(15), 3447–3450.
- Kedar, S. & Webb, F.H., 2005. The ocean's seismic hum. *Science*, **307**, 682–683.
- Kedar, S., Longuet-Higgins, M., Webb, F., Graham, N., Clayton, R. & Jones, C., 2008. The origin of deep ocean microseisms in the North Atlantic Ocean. *Proc. R. Soc. A*, **464**, 777–793.
- Larose, E., Derode, A., Campillo, M. & Fink, M., 2004. Imaging from one-bit correlations of wide-band diffuse wavefields. *J. Appl. Phys.*, **95**(12), 8393–8399.
- Larose, E., Roux, P. & Campillo, M., 2007. Reconstruction of Rayleigh-Lamb dispersion spectrum based on noise obtained from an air-jet forcing. *J. acoust. Soc. Am.*, **122**(6), 3437–3444.
- Lin, F., Ritzwoller, M.H., Townend, J., Savage, M. & Bannister, S., 2007. Ambient noise Rayleigh wave tomography of New Zealand. *Geophys. J. Int.*, **170**, 649–666, doi:10.1111/j.1365-246X.2007.03414.x.
- Lin, F., Moschetti, M.P. & Ritzwoller, M.H., 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophys. J. Int.*, **173**, 281–298, doi:10.1111/j.1365-246X.2008.03720.x.
- Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlation of a diffuse field. *J. acoust. Soc. Am.*, **110**, 3011–3017.
- Longuet-Higgins, M.S., 1950. A theory on the origin of microseisms. *Phil. Trans. R. Soc. Lond., A.*, **243**, 1–35.
- Matzel, E., 2008. Attenuation tomography using ambient noise correlation. *Seism. Res. Lett.*, **79**(2), 358.
- Nawa, K., Suda, N., Fukao, Y., Sato, T., Aoyama, Y. & Shibuya, K., 1998. Incessant excitation of the Earth's free oscillations. *Earth. planet. Space*, **50**, 3–8.

- Nishida, K. & Fukao, Y., 2007. Source distribution of Earth's background free oscillations. *J. geophys. Res.*, **112**, B06306, doi:10.1029/2006JB004720.
- Paul, A., Campillo, M., Margerin, L., Larose, E. & Derode, A., 2005. Empirical synthesis of time-asymmetrical green functions from the correlation of coda waves. *J. geophys. Res.*, **110**, B08302, doi:10.1029/2004JB003521.
- Prieto, G.A., Lawrence, J.F. & Beroza, G.C., 2009. Anelastic Earth structure from the coherency of the ambient seismic field. *J. geophys. Res.*, **144**, B07303, doi:10.1029/2008JB006067.
- Rhie, J. & Romanowicz, B., 2004. Excitation of Earth's incessant free oscillations by atmosphere-ocean-seafloor coupling. *Nature*, **431**, 552–556.
- Roux, P., Sabra, K.G. & Kuperman, W.A., 2005. Ambient noise cross-correlation in free space: Theoretical approach. *J. acoust. Soc. Am.*, **117**(1), 79–84, doi:10.1121/1.1830673.
- Sánchez-Sesma, F.J. & Campillo, M., 2006. Retrieval of the Green's function from cross-correlation: the canonical elastic problem. *Bull. seism. Soc. Am.*, **96**, 1182–1191.
- Sánchez-Sesma, F.J., Pérez-Ruiz, J., Campillo, M. & Lúzon, F., 2006. Elastodynamic 2D Green function retrieval from cross-correlation: canonical inclusion problem. *Geophys. Res. Lett.*, **33**, L13305, doi:10.1029/2006GL026454.
- Sánchez-Sesma, F.J., Pérez-Ruiz, J., Lúzon, F., Campillo, M. & Rodríguez-Castellanos, A., 2008. Diffuse fields in dynamic elasticity. *Wave Motion*, **45**, 641–654.
- Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophys. Res. Lett.*, **31**, L07614, doi:10.1029/2004GL019491.
- Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface wave tomography from ambient seismic noise. *Science*, **307**, 1615–1618.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase. *Phys. Rev. E*, **69**, 046610.
- Snieder, R., 2007. Extracting the green's function of attenuating heterogeneous media from uncorrelated waves. *J. acoust. Soc. Am.*, **121**, 2637–2643.
- Stehly, L., Campillo, M. & Shapiro, N., 2006. A study of the seismic noise from its long-range correlation properties. *J. geophys. Res.*, **111**, B10306, doi:10.1029/2005JB004237.
- Stehly, L., Campillo, M. & Shapiro, N., 2007. Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts. *Geophys. J. Int.*, **171**, 223–230, doi:10.1111/j.1365-246X.2007.03492.x.
- Stehly, L., Fry, B., Campillo, M., Shapiro, N., Guilbert, J., Boschi, L. & Giardini, D., 2009. Tomography of the alpine region from observations of seismic ambient noise. *Geophys. J. Int.*, **178**, 338–350, doi:10.1111/j.1365-246X.2009.04132.x.
- van Tiggelen, B.A., 2003. Green function retrieval and time-reversal in a disordered world. *Phys. Rev. Lett.*, **91**(24), 243904–4.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross-correlation. *Phys. Rev. Lett.*, **93**, 254301–4.
- Weaver, R.L. & Lobkis, O.I., 2001. Ultrasonics without a source: thermal fluctuation correlation at MHz frequencies. *Phys. Rev. Lett.*, **87**, 134301–4.
- Weaver, R.L. & Lobkis, O.I., 2003. Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal phonons. *J. acoust. Soc. Am.*, **113**, 2611–2621, doi:10.1121/1.1564017.
- Weaver, R.L., Froment, B. & Campillo, M., 2009. On the correlation of non-isotropically distributed ballistic scalar diffuse waves. *J. acoust. Soc. Am.*, **126**, 1817–1826.
- Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized earth model. *Geophys. J. R. astr. Soc.*, **78**, 641–660.
- Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across europe. *Geophys. J. Int.*, **168**, 259–274.
- Yao, H. & van der Hilst, R.D., 2009. Analysis of ambient noise energy distribution and phase velocity bias in ambient noise tomography, with application to SE Tibet. *Geophys. J. Int.*, **179**, 1113–1132.
- Yao, H., van der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I. Phase velocity maps. *Geophys. J. Int.*, **166**, 732–744, doi:10.1111/j.1365-246X.2006.03028.x.

The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise

Paul Cupillard, Laurent Stehly* and Barbara Romanowicz

Seismological Laboratory, University of California, Berkeley, CA 94720, USA. E-mail: paulcup@ipgp.jussieu.fr

Accepted 2010 December 13. Received 2010 December 5; in original form 2010 July 31

SUMMARY

Waveforms emerging from correlations of long seismic noise records are extensively used to investigate the crustal and upper-mantle structure of the Earth. To remove the non-stationary events that inevitably lie in seismic records, the so-called one-bit normalization is commonly applied to the noise data. This processing consists of replacing each sample of a record by its sign. Although it is a strong non-linear operation, it preserves the phase of the signal emerging from correlation. Some recent studies show that information can also be extracted from the amplitude of such correlations. In this paper, we develop a theory to understand these non-intuitive results. A statistical approach is used to get an expression of the one-bit noise correlation. This expression involves the standard deviations of coherent and incoherent noise. These two kinds of noise are precisely defined, and explicit expressions of their standard deviations are given in the case of a uniform distribution of noise sources generating surface waves on a layered half-space. In this case, we show that the one-bit noise correlation has the same phase and relative amplitude as the raw noise correlation. This is true in both elastic and anelastic media. Numerical simulations are performed to support our theory.

Key words: Interferometry; Seismic attenuation; Theoretical seismology; Wave propagation.

1 INTRODUCTION

Seismic noise correlations are now widely used to get information about the structure of the Earth. Prior to correlation, different processing steps have to be applied to the noise records. One of the most common processings is the one-bit normalization. It consists of retaining only the sign (+1 or -1) of each sample in the records. It was first introduced in communication theory. Bond & Cahn (1958) demonstrated the possibility of transmitting a continuous signal over a discrete channel by preserving the occurrence of the zero crossings only. Later, Voelcker (1966a,b) and Voelcker & Requicha (1973) extended this result to get modulation procedures for representing signals in terms of real and complex zeros. In seismology, Campillo & Paul (2003) used the one-bit normalization to process seismic coda. In this case, it enhances multiple scattering and thus improves the signal-to-noise ratio of coda correlations (Larose *et al.* 2004). In the context of long noise records, the one-bit normalization enables to remove non-stationary signals like earthquakes or spikes. For examples of applications of this technique, one can refer to Shapiro & Campillo (2004), Shapiro *et al.* (2005), Larose *et al.* (2007) and Yao & van der Hilst (2009). More sophisticated temporal normalizations can be found in Bensen *et al.* (2007) and Brooks & Gerstoft (2009).

The most common way to image the Earth interior from ambient seismic noise consists in evaluating dispersion curves of surface waves emerging from noise correlations. Since most of the noise energy ranges from 5 to 20 s period, the obtained images provide information on crustal and upper-mantle structure (e.g. Shapiro *et al.* 2005; Yao *et al.* 2006; Cho *et al.* 2007; Lin *et al.* 2007; Yang *et al.* 2007; Bensen *et al.* 2008; Lin *et al.* 2008; Yao *et al.* 2008; Stehly *et al.* 2009). Several studies also attempted to use the amplitude of noise correlations to investigate the origin of the seismic noise (Stehly *et al.* 2006; Pedersen *et al.* 2007; Yang & Ritzwoller 2008) and retrieve the attenuation of the Earth (Matzel 2008; Prieto *et al.* 2009). It is extremely interesting to note that all these works apply one-bit normalization or other strong non-linear operations to the noise records. In spite of such operations, information contained in both phase and amplitude of the correlations seems to be preserved. Understanding this non-intuitive phenomenon is the primary motivation for this work.

Voelcker (1966a) showed that zeros are fundamental informational attributes of signals. Therefore, it is not very surprising to retrieve some interesting information in one-bit noise correlations. Nevertheless, because of the lack of a theory, it is not clear so far what is effectively

*Now at: Géoazur, Université de Nice - Sophia Antipolis, 06108 Nice Cedex 02, France.

2 P. Cupillard, L. Stehly and B. Romanowicz

recovered. In this paper, a theoretical development is proposed to explain the full waveform (phase and amplitude) of the one-bit noise correlation. First, we follow Snieder (2004) to get an expression of the raw noise correlation. Using this expression and a statistical approach similar to Derode *et al.* (1999) and Larose *et al.* (2008), we introduce the concepts of coherent and incoherent noise. These two kinds of noise are Gaussian and their variances can be evaluated using the central limit theorem. Then, we look at the one-bit noise correlation: we apply basic laws of probability to the samples of one-bit noise records and we obtain a formula for the correlation. This formula involves the standard deviations of both coherent and incoherent noise. In this work, we provide explicit expressions for these standard deviations in the context of surface waves propagating in a layered medium. Extension to body waves and full 3-D cases is discussed at the end of the paper. Both elastic and anelastic cases are investigated.

2 THE RAW NOISE CORRELATION

2.1 General expression of the raw noise correlation

Consider two stations A and B that are separated by a distance Δ . These stations respectively record signals $A(t)$ and $B(t)$ due to random noise sources acting in the medium. We denote by $A_p(t)$, $B_p(t)$ the signal received in A , B , respectively, from a single point-source P . The correlation between $A(t)$ and $B(t)$ can thus be written as

$$C_{AB}(t) = \sum_{p, p'} \int A_p(\tau) B_{p'}(t + \tau) d\tau. \quad (1)$$

We assume that the signals produced by two distinct sources are uncorrelated. Therefore, the cross terms $p \neq p'$ in the double sum $\sum_{p, p'}$ vanish and expression (1) can be reduced to

$$C_{AB}(t) = \sum_P C_{AB}^P(t), \quad (2)$$

with

$$C_{AB}^P(t) = \int A_p(\tau) B_p(t + \tau) d\tau. \quad (3)$$

In the following, the sources and receivers are placed on the surface of a laterally homogeneous half-space (the elastic parameters of the medium only depend on the depth z). For the sake of simplicity, we work with fundamental mode surface waves, but overtones could be included by introducing a summation, as done by Snieder (2004), Halliday & Curtis (2008) or Kimman & Trampert (2010). Moreover, we consider vertical displacements only, but the whole theory developed in this paper could be easily extended to all components of the Green's tensor. In the frequency domain, the vertical-vertical component of the fundamental mode surface wave Green's tensor between two points U and V is given by (Aki & Richards 2002)

$$G_{UV}(\omega) = \frac{\exp\left[-i\left(k d_{UV} + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}} k d_{UV}}, \quad (4)$$

where ω is the angular frequency, $k(\omega)$ is the wavenumber and d_{UV} is the horizontal distance between U and V . Using eq. (4) and denoting by $|S_p(\omega)|^2$ the power spectral density of source P , the Fourier transform of cross-correlation (3) becomes

$$C_{AB}^P(\omega) = \frac{2}{\pi} |S_p(\omega)|^2 \frac{\exp\left[ik(d_{AP} - d_{BP})\right]}{k\sqrt{d_{AP} d_{BP}}}. \quad (5)$$

It follows that

$$C_{AB}(\omega) = \frac{2}{\pi} \sum_P |S_p(\omega)|^2 \frac{\exp\left[ik(d_{AP} - d_{BP})\right]}{k\sqrt{d_{AP} d_{BP}}}. \quad (6)$$

Replacing the summation over discrete sources by a surface integration (the variables with index p become functions of the source location \mathbf{r}), we obtain

$$C_{AB}(\omega) = \frac{2}{\pi} \iint |S(\mathbf{r}; \omega)|^2 \frac{\exp\left[ik(d_A(\mathbf{r}) - d_B(\mathbf{r}))\right]}{k\sqrt{d_A(\mathbf{r}) d_B(\mathbf{r})}} d\mathbf{r}. \quad (7)$$

2.2 The raw noise correlation in the case of a uniform distribution of sources

When $|S(\mathbf{r}; \omega)|^2$ is a smooth function of \mathbf{r} , the integral in eq. (7) can be evaluated using the stationary phase approximation (Snieder, 2004). Let us use this approximation in a Cartesian coordinate system. We position receiver A at the origin and receiver B on the positive x -axis. Then

$$C_{AB}(\omega) = \frac{2}{ik} \frac{\exp\left[i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}} k \Delta} \int_{\Delta}^{\infty} |S(x, y = 0; \omega)|^2 dx - \frac{2}{ik} \frac{\exp\left[-i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}} k \Delta} \int_{-\infty}^0 |S(x, y = 0; \omega)|^2 dx. \quad (8)$$

This result shows that the Green's function (GF) between A and B can emerge from the correlation of random noise records. The first term on the right-hand side contains the acausal GF and arises because of the noise sources at $x > \Delta$. The second term contains the causal GF and arises because of the sources at $x < 0$. This result is not new. It was first demonstrated by Lobkis & Weaver (2001) using the assumption of equipartitioning of the Earth's normal modes. Other proofs followed, invoking the fluctuation-dissipation theorem (van Tiggelen 2003), an analogy with time-reversal experiments (Derode *et al.* 2003), the reciprocity theorem (Wapenaar 2004) or the stationary-phase approximation (Snieder 2004).

Denoting by $v(\omega)$ the phase velocity of the fundamental mode Rayleigh wave and considering a uniform distribution of noise sources such that $|S(x, y = 0; \omega)|^2 = |S(\omega)|^2$, eq. (8) becomes

$$C_{AB}(\omega) = 2D \frac{v(\omega)|S(\omega)|^2}{i\omega} \left\{ \frac{\exp\left[i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} - \frac{\exp\left[-i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} \right\}. \quad (9)$$

We limit the integration over x to a range bounded by a finite distance $D \gg \Delta$. This is to prevent this integration to diverge. In practice, there is no problem of divergence because of intrinsic attenuation. We study this case in Section 5. Eq. (9) shows that one has to differentiate the correlation in time to get the GF. Not only the waveform but also the amplitude decay of the GF with distance Δ is retrieved by the correlation.

Fig. 1 confirms the theory. This figure shows results from a numerical experiment similar to those carried out by Cupillard & Capdeville (2010). The experiment consists in computing synthetic noise recordings in a spherical earth using a normal mode summation technique (e.g. Woodhouse & Gornius 1982). 300 sources are randomly positioned on the surface of the Earth. Each source generates a 24-hr random signal filtered between 66 and 200 s. An array of 12 receivers ($n = 0, \dots, 11$) records the wavefield produced by the noise sources. The correlations between the vertical displacement at station 0 and the vertical displacement at the other stations are performed. The Earth model is PREM (Dziewonski & Anderson 1981). The configuration of the experiment is shown in Fig. 1(a). It is repeated 5120 times (64 processors perform it 80 times each) and all the obtained correlations are then stacked. You can refer to Cupillard & Capdeville (2010) for more details. Fig. 1(b) compares the time-derivative of the correlation between stations 0 and 6 with the corresponding fundamental mode GF. As predicted by the theory, the two curves match very well. We also plot the comparison between the amplitude decay of the correlation along the array and the amplitude decay of the GF. Again, the curves match very well, which confirms the theory.

3 COHERENT AND INCOHERENT NOISE

3.1 Definition

From a statistical point of view, cross-correlation (3) can be viewed as an ensemble average over time τ . Thus we can write

$$C_{AB}^p(t) = \sigma_{A_p} \sigma_{B_p} \rho_{AB}^p(t), \quad (10)$$

where σ_{A_p} , σ_{B_p} are the standard deviations of the stationary signals $A_p(\tau)$, $B_p(\tau)$, respectively, and $\rho_{AB}^p(t)$ is the correlation coefficient between $A_p(\tau)$ and $B_p(t + \tau)$. The standard deviation σ_{A_p} , σ_{B_p} is the square-root of the energy arriving in A , B , respectively, from a point-source P , so

$$\sigma_{A_p} = \frac{1}{\sqrt{d_{AP}}} \quad (11)$$

and

$$\sigma_{B_p} = \frac{1}{\sqrt{d_{BP}}}. \quad (12)$$

Using the inverse Fourier transform of eq. (5) we find the expression of ρ_{AB}^p .

$$\rho_{AB}^p(t) = \frac{2}{\pi} \int \frac{|S_p(\omega)|^2}{k} \exp\left[i\omega\left(t - \frac{d_{BP} - d_{AP}}{v(\omega)}\right)\right] d\omega. \quad (13)$$

As $\sigma_{A_p}^2$ and $\sigma_{B_p}^2$ are both energies from P , the power spectral density $|S_p(\omega)|^2$ should appear in their definitions. Nevertheless, it is possible and much more practical to put this common factor in the definition of ρ_{AB}^p . We now substitute eq. (10) into eq. (2) and replace the summation \sum_p by a surface integration. This gives us

$$C_{AB}(t) = \iint \sigma_A(\mathbf{r}) \sigma_B(\mathbf{r}) \rho_{AB}(\mathbf{r}; t) d\mathbf{r}, \quad (14)$$

with

$$\sigma_A(\mathbf{r}) = \frac{1}{\sqrt{d_A(\mathbf{r})}}, \quad (15)$$

$$\sigma_B(\mathbf{r}) = \frac{1}{\sqrt{d_B(\mathbf{r})}}, \quad (16)$$

$$\rho_{AB}(\mathbf{r}; t) = \frac{2}{\pi} \int \frac{|S(\mathbf{r}; \omega)|^2}{k} \exp\left[i\omega\left(t - \frac{d_B(\mathbf{r}) - d_A(\mathbf{r})}{v(\omega)}\right)\right] d\omega. \quad (17)$$

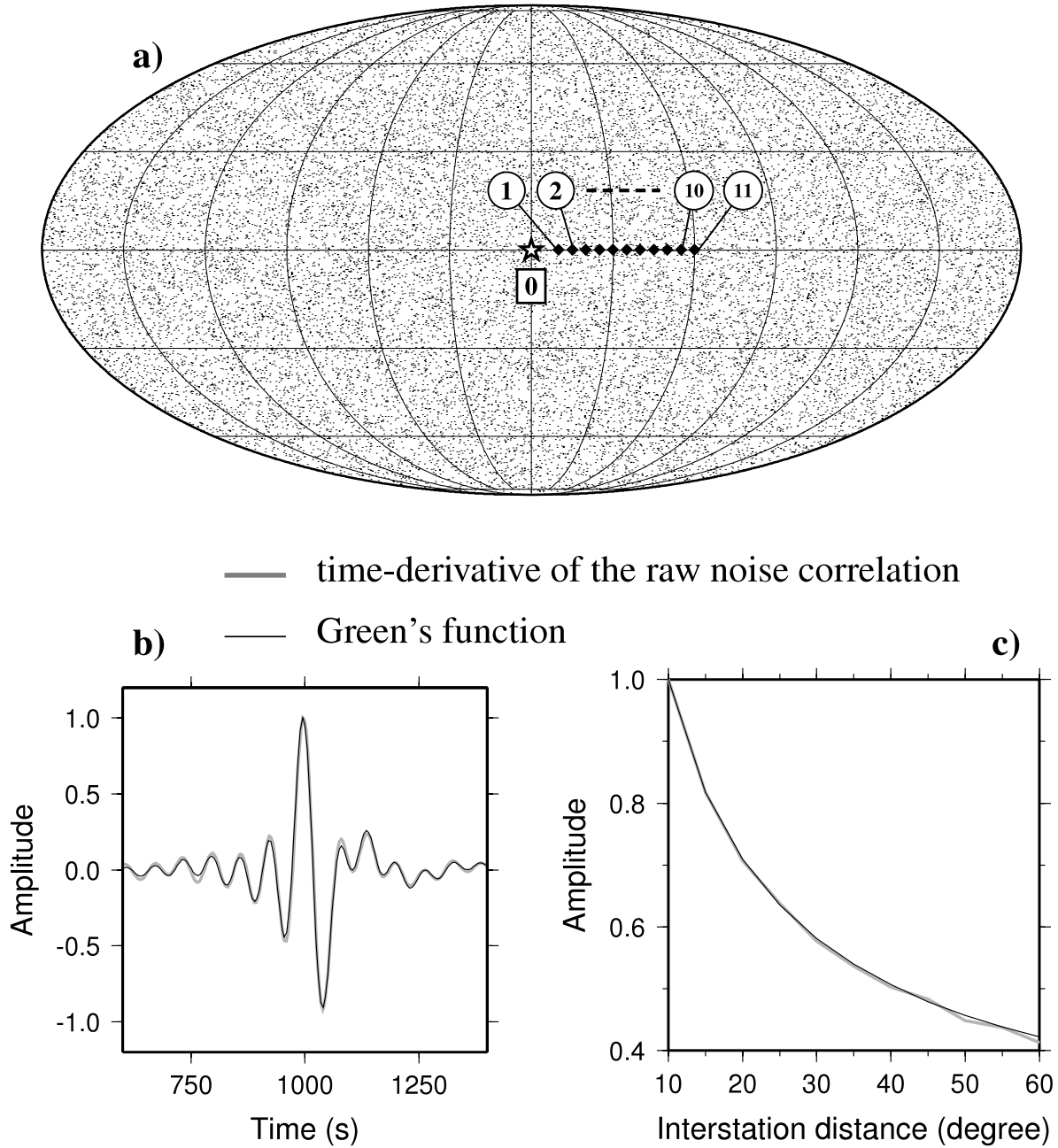
4 *P. Cupillard, L. Stehly and B. Romanowicz*

Figure 1. Source and receiver configuration of the simulation (a). Tiny pixels indicate the location of 24 000 noise sources generated by a single processor. All the sources are on the surface of the Earth. Diamonds represent 11 receivers ($n = 1, \dots, 11$). Two consecutive receivers are separated by 5° . The synthetic noise record from each station is correlated with the noise recorded at receiver 0 (white star). The distance between stations 0 and 1 is 10° . We plot the time-derivative of the correlation between stations 0 and 6 and the corresponding fundamental mode Rayleigh wave GF (b). We also compare the amplitude decay of the correlation with the amplitude decay of the GF (c). We see that the correlation fits both waveform and relative amplitude of the GF.

Taking a uniform distribution of noise sources into account, we easily demonstrate (Appendix A) that

$$C_{AB}^2(t) = \iint \sigma_A^2(\mathbf{r}) \rho_{AB}(\mathbf{r}; t) d\mathbf{r} \iint \sigma_B^2(\mathbf{r}) \rho_{AB}(\mathbf{r}; t) d\mathbf{r}. \quad (18)$$

This last expression shows that the instantaneous energy of the correlation is the product of the two functions

$$\Psi_A(t) = \iint \sigma_A^2(\mathbf{r}) \rho_{AB}(\mathbf{r}; t) d\mathbf{r} \quad (19)$$

and

$$\Psi_B(t) = \iint \sigma_B^2(\mathbf{r}) \rho_{AB}(\mathbf{r}; t) d\mathbf{r}. \quad (20)$$

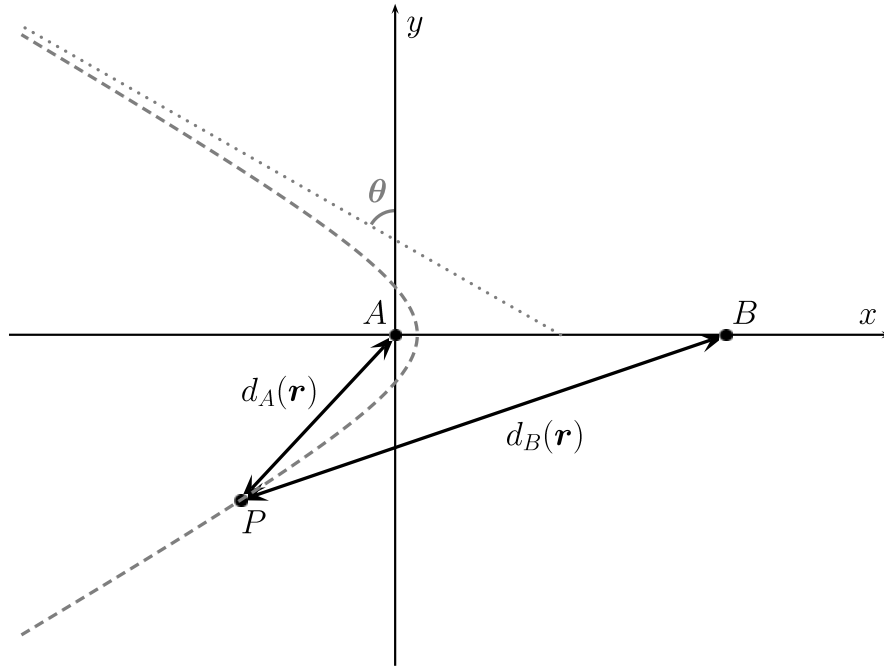


Figure 2. The two coordinate systems used in our derivation. We first consider a simple Cartesian system (x, y) . Receivers A and B are positioned in $(0, 0)$ and $(0, \Delta)$, respectively. The quantity $d_A(\mathbf{r}) - d_B(\mathbf{r})$ is involved in our derivation. This quantity is constant over a hyperbola so we introduce the coordinates (θ, ϕ) to work in a more convenient way. θ is the angle between the asymptote of a given hyperbola and the y -axis. ϕ the curvilinear coordinate along each hyperbola.

Let us study $\Psi_A(t)$ and $\Psi_B(t)$. Without changing our notations, we replace all the complex time-variables by their respective real part. Eq. (18) is still valid in this case and the correlation coefficient becomes

$$\rho_{AB}(\mathbf{r}; t) = \frac{2}{\pi} \int \frac{|S(\mathbf{r}; \omega)|^2}{k} \cos \left[\omega \left(t - \frac{d_B(\mathbf{r}) - d_A(\mathbf{r})}{v(\omega)} \right) \right] d\omega. \quad (21)$$

The value of $d_B(\mathbf{r}) - d_A(\mathbf{r})$ is constant over a hyperbola. Each hyperbola is characterized by θ which is the angle between its asymptotes and the y -axis (Fig. 2). We denote by ϕ the curvilinear coordinate along each hyperbola. Following Roux *et al.* (2005) we make the change of variables

$$\begin{cases} x = \frac{\Delta}{2} \sin \theta \cosh \phi + \frac{\Delta}{2} \\ y = \frac{\Delta}{2} \cos \theta \sinh \phi \end{cases} \quad \text{with} \quad \begin{cases} \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ \phi \in \mathbb{R}. \end{cases} \quad (22)$$

It follows

$$d_A(\mathbf{r}) = \frac{\Delta}{2} (\cosh \phi + \sin \theta), \quad (23)$$

$$d_B(\mathbf{r}) = \frac{\Delta}{2} (\cosh \phi - \sin \theta) \quad (24)$$

and

$$d_B(\mathbf{r}) - d_A(\mathbf{r}) = -\Delta \sin \theta. \quad (25)$$

We neglect the frequency-dependence of the phase velocity [$v(\omega) = v$] and we use the fact that the distribution of noise sources is uniform so we can rewrite (21) as a function of θ .

$$\rho_{AB}(\theta; t) = \frac{2v}{\pi} \int \frac{|S(\omega)|^2}{\omega} \cos \left[\omega \left(t + \frac{\Delta \sin \theta}{v} \right) \right] d\omega. \quad (26)$$

To perform the integration over ω , we need to choose an amplitude spectrum $|S(\omega)|$. For the sake of simplicity, we consider a boxcar function $H(\omega_0 + \frac{\Delta\omega}{2}) - H(\omega_0 - \frac{\Delta\omega}{2})$, where H is the Heaviside step function. We obtain

$$\rho_{AB}(\theta; t) = \frac{2v}{\pi} \left\{ \text{ci} \left[\left(\omega_0 + \frac{\Delta\omega}{2} \right) \left(t + \frac{\Delta \sin \theta}{v} \right) \right] - \text{ci} \left[\left(\omega_0 - \frac{\Delta\omega}{2} \right) \left(t + \frac{\Delta \sin \theta}{v} \right) \right] \right\}, \quad (27)$$

where ci is the cosine-integral special function (Appendix B). This function is defined in the positive real number space \mathbb{R}^{+*} but we can extend it to the set of negative arguments by posing $\text{ci}(u) = \text{ci}(-u) \forall u \in \mathbb{R}^{-*}$. Moreover, ci is not defined at $u = 0$ but the function in braces in eq. (27) has a finite limit when $t \rightarrow -\frac{\Delta \sin \theta}{v}$ so we can define it at 0 (Appendix B). Let us assume that the boxcar width $\Delta\omega$ is very large.

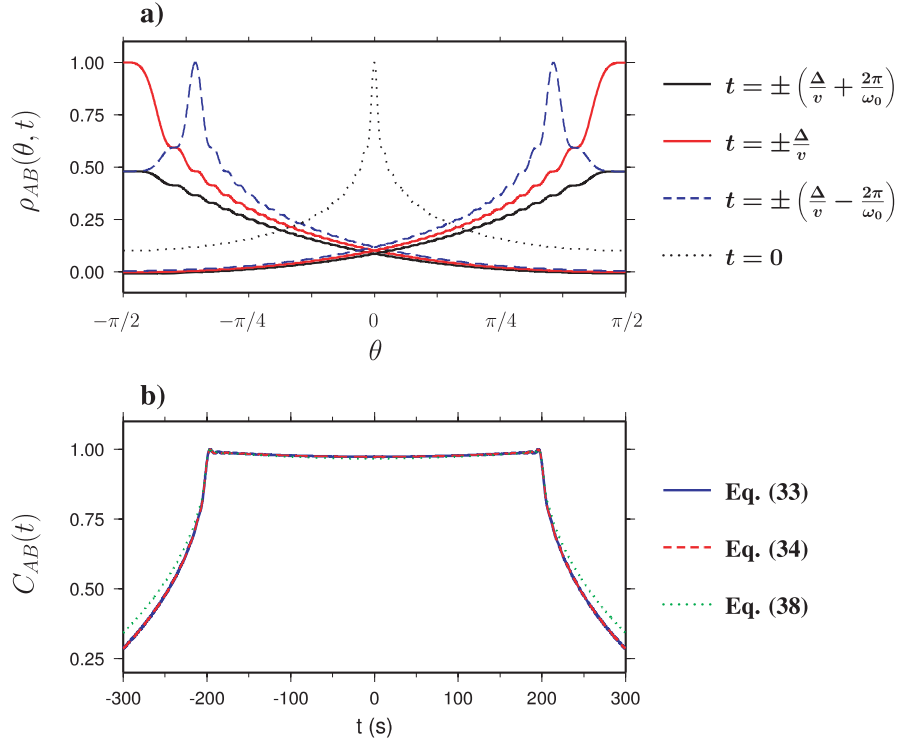
6 *P. Cupillard, L. Stehly and B. Romanowicz*

Figure 3. ρ_{AB} as a function of θ for four different values of t in the case of a large $\Delta\omega = 1.99\omega_0$ (a). The four curves are positive for all θ . We also plot the correlation $C_{AB}(t)$ using three different equations (b). The exact solution (eq. 33) is in blue, the solution involving the coherent noise (eq. 34) is in dashed red and the solution given by the stationary phase approximation (eq. 38) is in dotted green. For the calculation we took $f_0 = 0.05$ Hz, $v = 3$ km s $^{-1}$ and $\Delta = 10\lambda_0 = 600$ km.

Then $\rho_{AB}(\theta; t)$ is always positive (Fig. 3a). This means that $\Psi_A(t)$ and $\Psi_B(t)$ can be defined as sums of variances. Each variance in the sum $\Psi_A(t)$, $\Psi_B(t)$ is the energy $\sigma_A^2(\mathbf{r})$, $\sigma_B^2(\mathbf{r})$, arriving in A , B , respectively, due to a noise point-source \mathbf{r} weighted by a positive value $\rho_{AB}(\theta; t)$. Therefore, according to the central limit theorem, there exist for each lag time t two Gaussian signals $A^t(\tau)$ and $B^t(\tau)$ with zero mean and variance $\sigma_{A^t}^2 = \Psi_A(t)$ and $\sigma_{B^t}^2 = \Psi_B(t)$, respectively, such that $A^t(\tau)$ and $B^t(t + \tau)$ are perfectly correlated and give rise to $C_{AB}(t)$. In other words,

$$C_{AB}(t) = \int A^t(\tau)B^t(t + \tau)d\tau = \sigma_{A^t}\sigma_{B^t}. \quad (28)$$

We call $A^t(\tau)$ and $B^t(\tau)$ coherent noise at lag time t . We also define incoherent noise:

$$\overline{A^t}(\tau) = A(\tau) - A^t(\tau) \quad (29)$$

and

$$\overline{B^t}(\tau) = B(\tau) - B^t(\tau), \quad (30)$$

with variances $\sigma_{A^t}^2$ and $\sigma_{B^t}^2$. From eqs (28) to (30), it follows that

$$\int \overline{A^t}(\tau)\overline{B^t}(t + \tau)d\tau + \int \overline{A^t}(\tau)B^t(t + \tau)d\tau + \int A^t(\tau)\overline{B^t}(t + \tau)d\tau = 0. \quad (31)$$

3.2 Comparing three different expressions of the raw noise correlation

Fig. 3(b) shows the correlation $C_{AB}(t)$ computed in three different ways.

(i) Using eqs (23), (24) and (27) and the Jacobian of the change of variable (22)

$$J = \left(\frac{\Delta}{2}\right)^2 (\cosh^2\phi - \sin^2\theta), \quad (32)$$

we can rewrite the exact solution (14) as

$$C_{AB}(t) = \frac{v\Delta}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{\phi_0} \sqrt{\cosh^2\phi - \sin^2\theta} \{ci^+ - ci^-\} d\phi d\theta, \quad (33)$$

where $ci^+ = ci[(\omega_0 + \frac{\Delta\omega}{2})(t + \frac{\Delta \sin \theta}{v})]$, $ci^- = ci[(\omega_0 - \frac{\Delta\omega}{2})(t + \frac{\Delta \sin \theta}{v})]$ and $\phi_0 \gg 1$ is a finite value we introduce to bound the integration over ϕ (such as D in eq. 9). A numerical calculation of the double integral gives the curve in blue.

(ii) Using again eqs (23), (24), (27) and (32) we can express $\Psi_A(t)$ and $\Psi_B(t)$ as a function of θ and ϕ_0 . Then we can rewrite (28).

$$\begin{aligned} C_{AB}(t) &= \sigma_{A'} \sigma_{B'} \\ &= \sqrt{\Psi_A(t)} \sqrt{\Psi_B(t)} \\ &= \sqrt{\frac{v\Delta}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\sinh\phi_0 - \phi_0 \sin\theta) \{ci^+ - ci^-\} d\theta} \sqrt{\frac{v\Delta}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\sinh\phi_0 + \phi_0 \sin\theta) \{ci^+ - ci^-\} d\theta}. \end{aligned} \quad (34)$$

The two integrations over θ are performed numerically. We obtain the dashed red curve in Fig. 3(b). We see that this curve is indistinguishable from the blue curve, which confirms expression (18). It is actually very easy to give another proof, different from Appendix A, of this expression. Indeed, ϕ_0 is large so both $\Psi_A(t)$ and $\Psi_B(t)$ reduce to

$$\Psi_A(t) = \Psi_B(t) = \frac{v\Delta e^{\phi_0}}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \{ci^+ - ci^-\} d\theta. \quad (35)$$

Eq. (33) also reduces to

$$C_{AB}(t) = \frac{v\Delta e^{\phi_0}}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \{ci^+ - ci^-\} d\theta \quad (36)$$

so

$$C_{AB}^2(t) = \Psi_A(t) \Psi_B(t). \quad (37)$$

(iii) The expression of the correlation obtained using the stationary phase approximation is also plotted in Fig. 3(b) (dotted green). It corresponds to the real part of the inverse Fourier transform of eq. (9).

$$C_{AB}(t) = 2Dv \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{1}{\omega} \left[\frac{\cos(\omega t + k\Delta - \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} + \frac{\cos(\omega t - k\Delta + \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} \right] d\omega. \quad (38)$$

This approximate solution is good but not excellent. This is because $\Delta\omega$ is large. Very low frequencies are therefore involved and the exponential in the integrand in eq. (7) does not oscillate much.

3.3 Studying the case of narrow band sources

In practice, $\Delta\omega$ is not large; most of the ambient noise energy is concentrated between 5 and 20 s (Longuet-Higgins 1950; Friedrich *et al.* 1998; Schulte-Pelkum *et al.* 2004; Stehly *et al.* 2006; Gerstoft & Tanimoto 2007; Pedersen *et al.* 2007; Kedar *et al.* 2008; Stutzmann *et al.* 2009), so $\Delta\omega \sim \omega_0$. In this case, $\rho_{AB}(\theta; t)$ can be negative (Fig. 4a) and $\Psi_A(t)$ and $\Psi_B(t)$ as expressed in (19) and (20) are no longer sums of variances. Nevertheless, the stationary phase approximation can be used to rewrite these equations.

$$\begin{aligned} \Psi_A(t) &= \int_{\Delta}^{D+\frac{\Delta}{2}} \sigma_A^2(x, y=0) \rho_{AB}(x, y=0; t) dx \{1 + o[(\Delta/D)^0]\} \\ &\quad + \int_{\frac{\Delta}{2}-D}^0 \sigma_A^2(x, y=0) \rho_{AB}(x, y=0; t) dx \{1 + o[(\Delta/D)^0]\} \end{aligned} \quad (39)$$

and

$$\begin{aligned} \Psi_B(t) &= \int_{\Delta}^{D+\frac{\Delta}{2}} \sigma_B^2(x, y=0) \rho_{AB}(x, y=0; t) dx \{1 + o[(\Delta/D)^0]\} \\ &\quad + \int_{\frac{\Delta}{2}-D}^0 \sigma_B^2(x, y=0) \rho_{AB}(x, y=0; t) dx \{1 + o[(\Delta/D)^0]\}, \end{aligned} \quad (40)$$

with

$$\rho_{AB}(x, y=0; t) = v\sqrt{x(x-\Delta)} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{1}{\omega} \left[\frac{\cos(\omega t + k\Delta - \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} + \frac{\cos(\omega t - k\Delta + \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} \right] d\omega. \quad (41)$$

These expressions are easily obtained using eqs (A5)–(A8). Eq. (41) shows that the sign of ρ_{AB} is a constant over x for a given lag time t , so we can define coherent noise $A^t(\tau)$ and $B^t(\tau)$ with variances $\sigma_{A^t}^2$ and $\sigma_{B^t}^2$, respectively, as we did in the previous paragraph. We have

$$\sigma_{A^t}^2 = \text{sgn}[\rho_{AB}(x, y=0; t)] \Psi_A(t) \quad (42)$$

$$\begin{aligned} &= \int_{\Delta}^{D+\frac{\Delta}{2}} \sigma_A^2(x, y=0) |\rho_{AB}(x, y=0; t)| dx \{1 + o[(\Delta/D)^0]\} \\ &\quad + \int_{\frac{\Delta}{2}-D}^0 \sigma_A^2(x, y=0) |\rho_{AB}(x, y=0; t)| dx \{1 + o[(\Delta/D)^0]\}, \end{aligned} \quad (43)$$

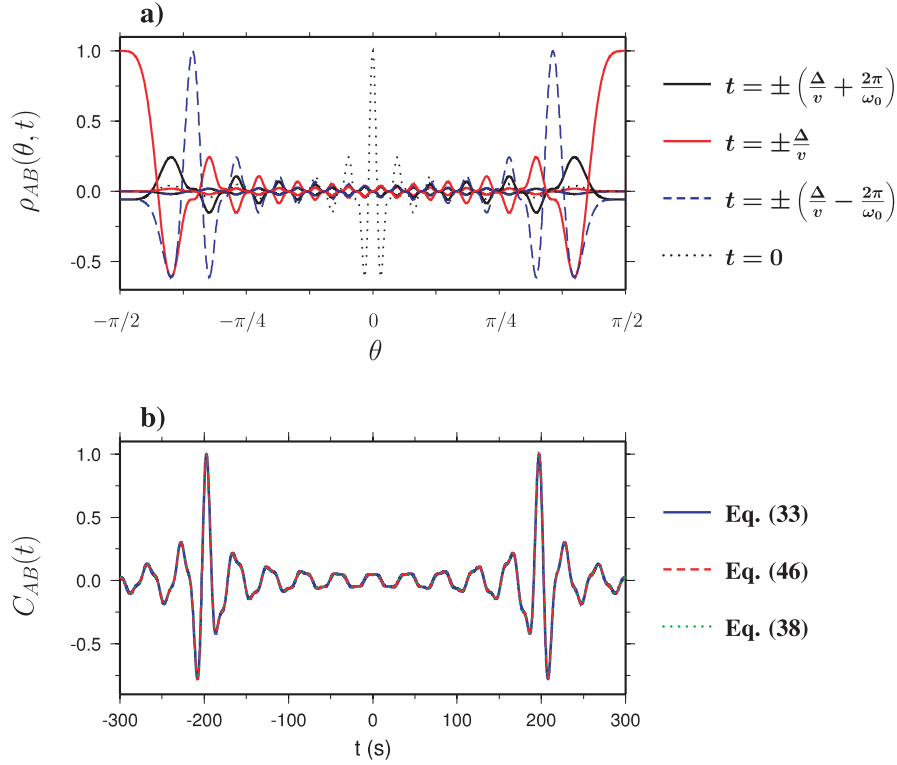
8 *P. Cupillard, L. Stehly and B. Romanowicz*

Figure 4. ρ_{AB} as a function of θ for four different values of t in the case of a small $\Delta\omega = \omega_0$ (a). Unlike the case of a large $\Delta\omega$ shown in Fig. 3, the four curves can be positive or negative depending on θ . The correlation $C_{AB}(t)$ is also plotted using three different equations (b). The exact solution (eq. 33) is in blue, the solution involving the coherent noise (eq. 46) is in dashed red and the solution given by the stationary phase approximation (eq. 38) is in dotted green. For the calculation we took $f_0 = 0.05$ Hz, $v = 3$ km s $^{-1}$ and $\Delta = 10\lambda_0 = 600$ km.

$$\sigma_{B^t}^2 = \text{sgn}[\rho_{AB}(x, y = 0; t)] \Psi_B(t) \quad (44)$$

$$\begin{aligned} &= \int_{\Delta}^{D+\frac{\Delta}{2}} \sigma_B^2(x, y = 0) |\rho_{AB}(x, y = 0; t)| dx \{1 + o[(\Delta/D)^0]\} \\ &+ \int_{\frac{\Delta}{2}-D}^0 \sigma_B^2(x, y = 0) |\rho_{AB}(x, y = 0; t)| dx \{1 + o[(\Delta/D)^0]\} \end{aligned} \quad (45)$$

and

$$C_{AB}(t) = \int A^t(\tau) B^t(t + \tau) d\tau = \text{sgn}[\rho_{AB}(x, y = 0; t)] \sigma_{A^t} \sigma_{B^t}. \quad (46)$$

In the last equation we see that the sign of ρ_{AB} indicates if $A^t(\tau)$ and $B^t(t + \tau)$ are perfectly correlated or perfectly anticorrelated. Of course we can also define incoherent noise.

Again, we plot $C_{AB}(t)$ using three different equations (Fig. 4b). (i) The exact solution (33) is represented in blue. (ii) The result from the stationary phase approximation (eq. 38) is the dotted green curve. We see that it perfectly fits the blue curve, meaning that the approximation is very good when $\Delta\omega = \omega_0$. (iii) The dashed red curve corresponds to eq. (46). In this last equation, σ_{A^t} and σ_{B^t} are obtained putting expression (41) in (43) and (45). This gives

$$\sigma_{A^t} = \sigma_{B^t} = \sqrt{2Dv \left| \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{1}{\omega} \left[\frac{\cos(\omega t + k\Delta - \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} + \frac{\cos(\omega t - k\Delta + \frac{\pi}{4})}{\sqrt{\frac{\pi}{2}k\Delta}} \right] d\omega \right|} \quad (47)$$

$$= \sqrt{|C_{AB}(t)|} \quad (48)$$

The fact that both σ_{A^t} and σ_{B^t} are equal to the square root of the correlation is not surprising. Indeed, expressions (39) and (40) have been obtained using the zeroth-order version ($o[(\Delta/D)^0]$) of eqs (A6) and (A8). In this case, both $\Psi_A(\omega)$ and $\Psi_B(\omega)$ are equal to $C_{AB}(\omega)$. This explains why the blue and the dashed red curves are identical in the figure.

4 THE ONE-BIT NOISE CORRELATION

4.1 General expression of the one-bit noise correlation

One-bit normalization consists of retaining only the sign of the raw signal by replacing all positive amplitudes with a 1 and all negative amplitudes with a -1 . The one-bit noise correlation $C_{AB}^{ob}(t)$ between $A(t)$ and $B(t)$ can be written

$$C_{AB}^{ob}(t) = \int \text{sgn}[A(\tau)] \text{sgn}[B(t + \tau)] d\tau \quad (49)$$

$$= n_1(t) - n_{-1}(t), \quad (50)$$

where $n_1(t)$ and $n_{-1}(t)$ are the number of samples for which $\text{sgn}[A(\tau)] = \text{sgn}[B(t + \tau)]$, and $\text{sgn}[A(\tau)] \neq \text{sgn}[B(t + \tau)]$, respectively.

For some samples τ , $|\overline{A^i}(\tau)| > |A^i(\tau)|$ or $|\overline{B^i}(t + \tau)| > |B^i(t + \tau)|$: at one of the two stations, the incoherent noise has a larger amplitude than the coherent noise and so controls the sign of the sample at this station. In this case, the two events $\text{sgn}[A(\tau)] = \text{sgn}[B(t + \tau)]$ and $\text{sgn}[A(\tau)] \neq \text{sgn}[B(t + \tau)]$ have the same probability. This is because of the incoherency of the random signals $\overline{A^i}(\tau)$ and $\overline{B^i}(t + \tau)$ (cf. eq. 31). Therefore we have $n_1(t) = n_{-1}(t)$ for this population of samples, which means that there is no contribution from this population to the value of $C_{AB}^{ob}(t)$.

For the other samples, $|\overline{A^i}(\tau)| < |A^i(\tau)|$ and $|\overline{B^i}(t + \tau)| < |B^i(t + \tau)|$: the coherent noise controls the sign of both $A(\tau)$ and $B(t + \tau)$, so $\text{sgn}[A(\tau)] = \text{sgn}[A^i(\tau)]$ and $\text{sgn}[B(t + \tau)] = \text{sgn}[B^i(t + \tau)]$. Because of the perfect correlation or anticorrelation between $A^i(\tau)$ and $B^i(t + \tau)$ (cf. eqs 28 and 46) we have $\text{sgn}[B^i(t + \tau)] = \text{sgn}[A^i(\tau)] \forall \tau$ (so $n_{-1}(t) = 0$) or $\text{sgn}[B^i(t + \tau)] \neq \text{sgn}[A^i(\tau)] \forall \tau$ (so $n_1(t) = 0$). Therefore, we can write

$$|C_{AB}^{ob}(t)| = n P_A^i P_B^i, \quad (51)$$

where n is the total number of samples in the correlation, P_A^i is the probability that $|A^i(\tau)| > |\overline{A^i}(\tau)|$ and P_B^i is the probability that $|B^i(\tau)| > |\overline{B^i}(\tau)|$.

Coherent and incoherent noise are both Gaussian so we are able to express P_A^i and P_B^i (Appendix C). Denoting by $\sigma_{\overline{A^i}}$ and $\sigma_{\overline{B^i}}$ the standard deviation of $\overline{A^i}(\tau)$ and $\overline{B^i}(\tau)$, respectively, we find

$$|C_{AB}^{ob}(t)| = n \left[1 - \frac{2}{\pi} \arctan \left(\frac{\sigma_{\overline{A^i}}}{\sigma_{A^i}} \right) \right] \left[1 - \frac{2}{\pi} \arctan \left(\frac{\sigma_{\overline{B^i}}}{\sigma_{B^i}} \right) \right]. \quad (52)$$

Eq. (52) is the most important result of this paper. It gives the expression of the one-bit noise correlation and shows how it is related to physical parameters. The involved physical parameters are the standard deviations of coherent and incoherent noise. More precisely, the ratio $w_{R^i} = \sigma_{\overline{R^i}}/\sigma_{R^i}$ at each receiver R is the argument of an inverse tangent function. At a given time t , if there is no coherent noise at one of the receivers, then no signal emerges from the correlation: $C_{AB}^{ob}(t) = 0$ because w_{R^i} tends to infinity so $1 - \frac{2}{\pi} \arctan(w_{R^i}) = 0$. On the contrary, if the coherent energy is large with respect to the incoherent energy at both receivers, then most of the n samples contribute to the correlation so $|C_{AB}^{ob}(t)|$ is large. As long as coherent and incoherent noise exist, eq. (52) is valid. In Section 3 we defined these two kinds of noise in the case of a uniform distribution of sources. One can reasonably think that they also exist in the case of other distributions.

4.2 The one-bit noise correlation in the case of a uniform distribution of sources

Now we study what (52) becomes in the case of a uniform distribution of noise sources. From eqs (29) and (30) we have

$$\sigma_{\overline{A^i}}^2 = \sigma_A^2 - \sigma_{A^i}^2, \quad (53)$$

$$\sigma_{\overline{B^i}}^2 = \sigma_B^2 - \sigma_{B^i}^2, \quad (54)$$

respectively, where σ_A^2 and σ_B^2 are the variance of $A(t)$ and $B(t)$, respectively. Then

$$|C_{AB}^{ob}(t)| = \left[1 - \frac{2}{\pi} \arctan \sqrt{\frac{\sigma_A^2}{\sigma_{A^i}^2} - 1} \right] \left[1 - \frac{2}{\pi} \arctan \sqrt{\frac{\sigma_B^2}{\sigma_{B^i}^2} - 1} \right]. \quad (55)$$

In this last expression, we normalized the correlation by the number of samples. The distribution of sources is uniform so $\sigma_A^2 = \sigma_B^2 = \sigma^2$, where $\sigma^2 = C_{UU}(t = 0)$ is the autocorrelation peak at any point U at the surface of the medium. Moreover, eq. (48) (and also eqs 35 and 36 in the case of a large $\Delta\omega$) shows that

$$\sigma_{A^i}^2 = \sigma_{B^i}^2 = |C_{AB}(t)|, \quad (56)$$

so (55) becomes

$$|C_{AB}^{ob}(t)| = \left[1 - \frac{2}{\pi} \arctan \sqrt{\frac{\sigma^2}{|C_{AB}(t)|} - 1} \right]^2. \quad (57)$$

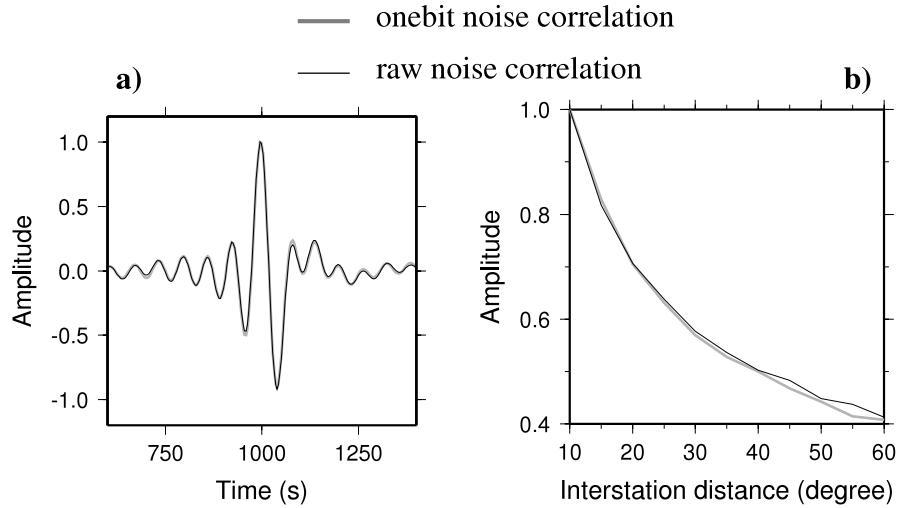


Figure 5. Comparison between raw noise correlations and one-bit noise correlations obtained by numerical simulation (*cf.* Fig. 1a). The time-derivatives of the two waveforms from the pair of stations 0 and 6 are very similar (a). The amplitude decays along the array of receivers also match very well (b).

Noting that $\sigma^2 \gg |C_{AB}(t)|$ and $\arctan \frac{1}{x} = \frac{\pi}{2} - \arctan x$, $\forall x \in \mathbb{R}^{+*}$, we obtain

$$|C_{AB}^{ob}(t)| = \left[\frac{2}{\pi} \arctan \sqrt{\frac{|C_{AB}(t)|}{\sigma^2}} \right]^2. \quad (58)$$

The argument of the arctan function is close to zero so the first term of its Taylor series ($\arctan x = x$) can be used as a good linear approximation. It follows that

$$|C_{AB}^{ob}(t)| = \left(\frac{2}{\pi\sigma} \right)^2 |C_{AB}(t)|. \quad (59)$$

This equation sets the equality between the raw and the one-bit noise correlations in the case of a uniform distribution of noise sources generating fundamental mode surface waves. Of course, the absolute amplitudes are not the same but the waveforms and the relative amplitudes are identical. Previous works (Derode *et al.* 1999; Larose *et al.* 2004) showed the emergence of a signal from a one-bit noise correlation and studied its signal-to-noise ratio, but it is the first time that the equality (59) is demonstrated. Correlations from the numerical simulation (*cf.* Fig. 1a) illustrate our result. One-bit noise correlations are compared to raw noise correlations (Fig. 5). We see that the waveforms and the amplitude decays are the same.

5 WHAT HAPPENS IN AN ANELASTIC MEDIUM

We now introduce intrinsic attenuation in the medium. The fundamental mode Rayleigh wave GF between two points U and V then is

$$G_{UV}(\omega) = \frac{\exp[-i(k d_{UV} + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}} k d_{UV}} \exp\left(-\frac{k d_{UV}}{2Q}\right), \quad (60)$$

where Q is the quality factor of the medium. Therefore, the raw noise correlation between $A(t)$ and $B(t)$ can be written in the frequency domain as

$$C_{AB}(\omega) = \frac{2}{\pi} \iint |S(\mathbf{r}; \omega)|^2 \frac{\exp[ik(d_A(\mathbf{r}) - d_B(\mathbf{r}))]}{k\sqrt{d_A(\mathbf{r})d_B(\mathbf{r})}} \exp\left[-\frac{k}{2Q}(d_A(\mathbf{r}) + d_B(\mathbf{r}))\right] d\mathbf{r}. \quad (61)$$

We assume that $|S(\mathbf{r}; \omega)|^2$ is a smooth function of \mathbf{r} so we can use the stationary phase approximation. For the sake of simplicity, we consider the signal emerging from the sources at $x < 0$ only. We obtain

$$C_{AB}(\omega) = -\frac{2}{ik} \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}} k \Delta} \exp\left(-\frac{k\Delta}{2Q}\right) \int_{-\infty}^0 |S(x, y=0; \omega)|^2 \exp\left(\frac{kx}{Q}\right) dx. \quad (62)$$

Expression (62) shows that the correlation contains the GF of the anelastic medium, including geometrical spreading as well as intrinsic attenuation. We assume a uniform distribution of noise so we get

$$C_{AB}(\omega) = -2L \frac{v(\omega)|S(\omega)|^2}{i\omega} \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}} k \Delta} \exp\left(-\frac{k\Delta}{2Q}\right), \quad (63)$$

where $L = \frac{Q}{k}$ is a factor also found by Snieder (2004). Eq. (63) shows that, again, the correlation has to be differentiated in time to match the GF. To check eq. (63), we carry out a new numerical experiment in which an anelastic PREM is considered. Results are shown in Fig. 6. We see that the time-derivative of the correlation fits both the waveform and the amplitude decay of the GF.

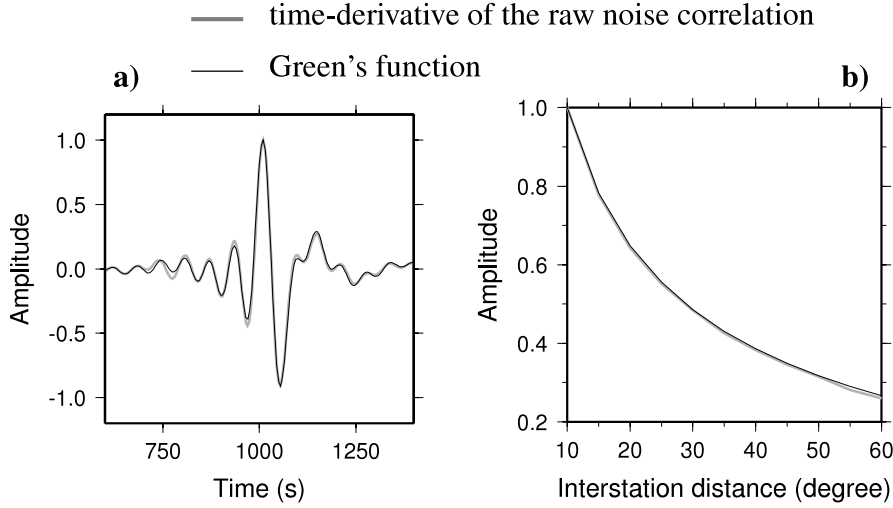


Figure 6. Comparison between the GF and the raw noise correlation in an anelastic earth. We use the pair of stations 0 and 6 to compare the waveforms (a). We see they are very similar. We also compare the amplitude decays (b). They are the same.

Following what we developed in Section 3, we introduce the functions $\sigma_A(\mathbf{r})$ and $\sigma_B(\mathbf{r})$ to write the correlation C_{AB} in the time domain as

$$C_{AB}(t) = \iint \sigma_A(\mathbf{r})\sigma_B(\mathbf{r})\rho_{AB}(\mathbf{r}; t) d\mathbf{r}. \quad (64)$$

In an anelastic medium, we have

$$\sigma_A(\mathbf{r}) = \frac{\exp\left(\frac{-kd_A(\mathbf{r})}{2Q}\right)}{\sqrt{d_A(\mathbf{r})}} \quad (65)$$

and

$$\sigma_B(\mathbf{r}) = \frac{\exp\left(\frac{-kd_B(\mathbf{r})}{2Q}\right)}{\sqrt{d_B(\mathbf{r})}}. \quad (66)$$

Using the inverse Fourier transform of eq. (61), we then find that the correlation coefficient ρ_{AB} is the same than in the pure elastic case (*cf.* eq. 17).

$$\rho_{AB}(\mathbf{r}; t) = \frac{2}{\pi} \int \frac{|S(\mathbf{r}; \omega)|^2}{k} \exp\left[i\omega\left(t - \frac{d_B(\mathbf{r}) - d_A(\mathbf{r})}{v(\omega)}\right)\right] d\omega. \quad (67)$$

The question now is to know if eq. (18) is still valid with the new definition of $\sigma_A(\mathbf{r})$ and $\sigma_B(\mathbf{r})$ (expressions 65 and 66). If so, then we will be able to define coherent and incoherent noise, and the general formula for the one-bit noise correlation (52) (and also 55) will be usable. Appendix D shows that

$$\begin{aligned} \Psi_A(t) &= \iint \sigma_A^2(\mathbf{r})\rho_{AB}(\mathbf{r}, t) d\mathbf{r} \\ &= -2 \Delta \Gamma\left(\frac{1}{2}\right) \int \frac{v(\omega)|S(\omega)|^2}{i\omega} U\left(\frac{1}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\exp\left[i\left(\omega t - k\Delta - \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} d\omega \end{aligned} \quad (68)$$

and

$$\begin{aligned} \Psi_B(t) &= \iint \sigma_B^2(\mathbf{r})\rho_{AB}(\mathbf{r}, t) d\mathbf{r} \\ &= -2 \Delta \Gamma\left(\frac{3}{2}\right) \int \frac{v(\omega)|S(\omega)|^2}{i\omega} U\left(\frac{3}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\exp\left[i\left(\omega t - k\Delta - \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} \exp\left(-\frac{k\Delta}{Q}\right) d\omega. \end{aligned} \quad (69)$$

Considering the real part of these equations and assuming that $v(\omega) = v$ and $|S(\omega)| = H(\omega_0 + \frac{\Delta\omega}{2}) - H(\omega_0 - \frac{\Delta\omega}{2})$, we get

$$\Psi_A(t) = 2 \Delta v \Gamma\left(\frac{1}{2}\right) \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} U\left(\frac{1}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\cos\left(\omega t - k\Delta + \frac{\pi}{4}\right)}{\omega\sqrt{\frac{\pi}{2}k\Delta}} d\omega \quad (70)$$

and

$$\Psi_B(t) = 2 \Delta v \Gamma\left(\frac{3}{2}\right) \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} U\left(\frac{3}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\cos\left(\omega t - k\Delta + \frac{\pi}{4}\right)}{\omega\sqrt{\frac{\pi}{2}k\Delta}} \exp\left(-\frac{k\Delta}{Q}\right) d\omega. \quad (71)$$

12 *P. Cupillard, L. Stehly and B. Romanowicz*

In Fig. 7, we plot the two normalized functions $\Psi_A(t)$ and $\Psi_B(t)$ for two different values of $\Delta\omega$ ($1.99\omega_0$ and ω_0) and two different values of Q (50 and 500). When Q is high, the two curves (plain blue for Ψ_A , red dashed for Ψ_B) are hardly distinguishable. This is not surprising: the intrinsic attenuation is small so we are close to the elastic case and we have $\Psi_A(t) \propto \Psi_B(t)$ (the absolute amplitudes are different so there is no equality). When Q is small, the effect of the intrinsic attenuation is significant and breaks the proportionality between the two functions. In the same figure, we also plot the product $\Psi_A(t)\Psi_B(t)$ and we compare the result with the square of the correlation (63) written in the time domain.

$$C_{AB}(t) = 2Qv^2 \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{\cos(\omega t - k\Delta + \frac{\pi}{4})}{\omega^2 \sqrt{\frac{\pi}{2}} k\Delta} \exp\left(-\frac{k\Delta}{2Q}\right) d\omega. \quad (72)$$

The two curves (plain blue for $\Psi_A(t)\Psi_B(t)$, red dashed for $C_{AB}^2(t)$) are normalized. In each case, they match very well. This is a nice result because it shows that eq. (18) still holds (up to a constant) in the anelastic case. We are not able to demonstrate this result analytically but the extreme values of Q (50 and 500 correspond to extremely low and high magnitudes in the Earth) and $\Delta\omega$ we use are good grounds to assess it, at least to first order. Finally, we plot $\Psi_A(t)\Psi_B(t)$ and $C_{AB}^2(t)$ at $t = t_0 = \frac{\Delta}{v} - \frac{\pi}{\omega_0^4}$ as a function of the interstation distance Δ ranging from $3\lambda_0$ to $20\lambda_0$. Again, the curves match very well. This means that the proportionality constant between $\Psi_A(t)\Psi_B(t)$ and $C_{AB}^2(t)$ is not a function of Δ . In other words, the amplitude decay of $\Psi_A(t)\Psi_B(t)$ is the same as the amplitude decay of $C_{AB}^2(t)$. We are now allowed to write

$$C_{AB}^2(t) \simeq \alpha \Psi_A(t)\Psi_B(t), \quad (73)$$

where α is a positive constant over t and Δ . As the expression of $\rho_{AB}(r;t)$ used in the definition of $\Psi_A(t)$ and $\Psi_B(t)$ is the same as the one in the elastic case (cf. eqs 17 and 67), we can take advantage of all the properties detailed in Section 3 and define coherent and incoherent noise. Therefore, formula (55) can be used to express the one-bit noise correlation. Putting the constant α into $\sigma_{A'}^2$, we find

$$|C_{AB}^{ob}(t)| = \left[1 - \frac{2}{\pi} \arctan \sqrt{\frac{\sigma_A^2}{\alpha |\Psi_A(t)|} - 1} \right] \left[1 - \frac{2}{\pi} \arctan \sqrt{\frac{\sigma_B^2}{|\Psi_B(t)|} - 1} \right]. \quad (74)$$

The distribution of noise sources is uniform so, again, we can use the fact that $\sigma_A^2 = \sigma_B^2 = \sigma^2 = C_{UU}(t=0) \gg \Psi_{A,B}(t)$ to reduce (74) to

$$|C_{AB}^{ob}(t)| = \left(\frac{2}{\pi\sigma} \right)^2 \sqrt{\alpha |\Psi_A(t)|} \sqrt{|\Psi_B(t)|}. \quad (75)$$

$\Psi_A(t)$ and $\Psi_B(t)$ have same sign so (75) becomes

$$|C_{AB}^{ob}(t)| = \left(\frac{2}{\pi\sigma} \right)^2 \sqrt{\alpha \Psi_A(t)\Psi_B(t)} \quad (76)$$

$$\simeq \left(\frac{2}{\pi\sigma} \right)^2 |C_{AB}(t)|. \quad (77)$$

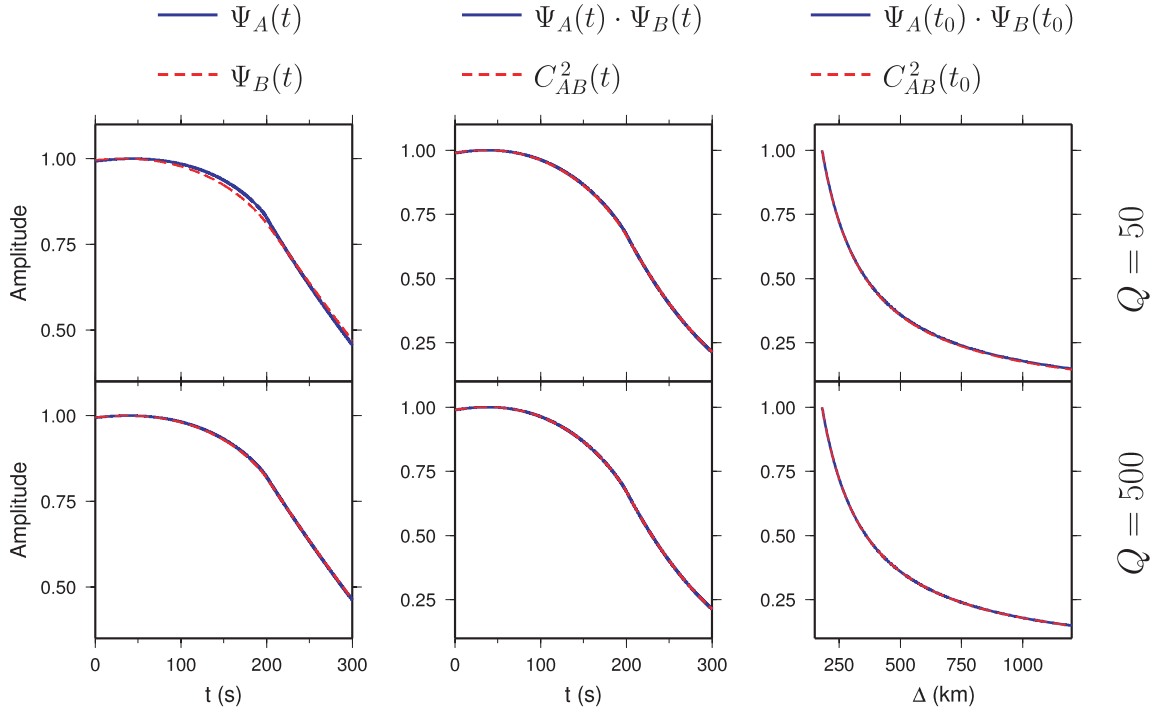
Similarly to the elastic case, we find that the one-bit noise correlation is equal to the raw noise correlation. This is confirmed by numerical results. In Fig. 8 we compare the waveforms and the amplitude decays of raw and one-bit noise correlations. We see they are the same.

6 DISCUSSION AND CONCLUSIONS

We provided an expression for the one-bit noise correlation. This expression involves the standard deviations of coherent and incoherent noise. For a given lag time t , the coherent noise $A^i(\tau)$ at a receiver A is a Gaussian signal that one can extract from the full noise record $A(\tau)$ and that perfectly correlates with the coherent noise $B^i(t+\tau)$ from another full noise record $B(t+\tau)$ at a receiver B . Then, the correlation of the coherent noises is exactly the correlation of the full noise records: $\int A(\tau)B(t+\tau)d\tau = \int A^i(\tau)B^i(t+\tau)d\tau$. The incoherent noise is the difference between the full noise record and the coherent noise. As long as you can define these two kinds of noise, our expression of the one-bit noise correlation is valid.

In this work, we detailed the coherent and incoherent standard deviations in the case of a uniform distribution of noise sources generating surface waves on a laterally homogeneous half-space. In this case, we showed that the one-bit noise correlation is equal to the raw noise correlation and so contains the GF. This property has been known for a long time. It is true in both elastic and anelastic media and it has been extensively used so far (e.g. Shapiro & Campillo 2004; Shapiro *et al.* 2005; Larose *et al.* 2007; Yao & van der Hilst 2009). Nevertheless, there was no theoretical proof of it. We here give one for the first time. An important result is that the equality does not only hold for the waveform but also for the relative amplitude: both geometrical spreading and intrinsic attenuation of the GF are retrieved by the one-bit noise correlation when the distribution of noise sources is uniform. This means that information can be extracted from the amplitude of one-bit noise correlations, which is not obvious because one-bit normalization is a very strong operation on the amplitude of noise recordings. In the case of a non-uniform distribution of sources, one has to study the coherent and incoherent standard deviations which correspond to the

a) $\Delta\omega = 1.99 \omega_0$:



b) $\Delta\omega = \omega_0$:

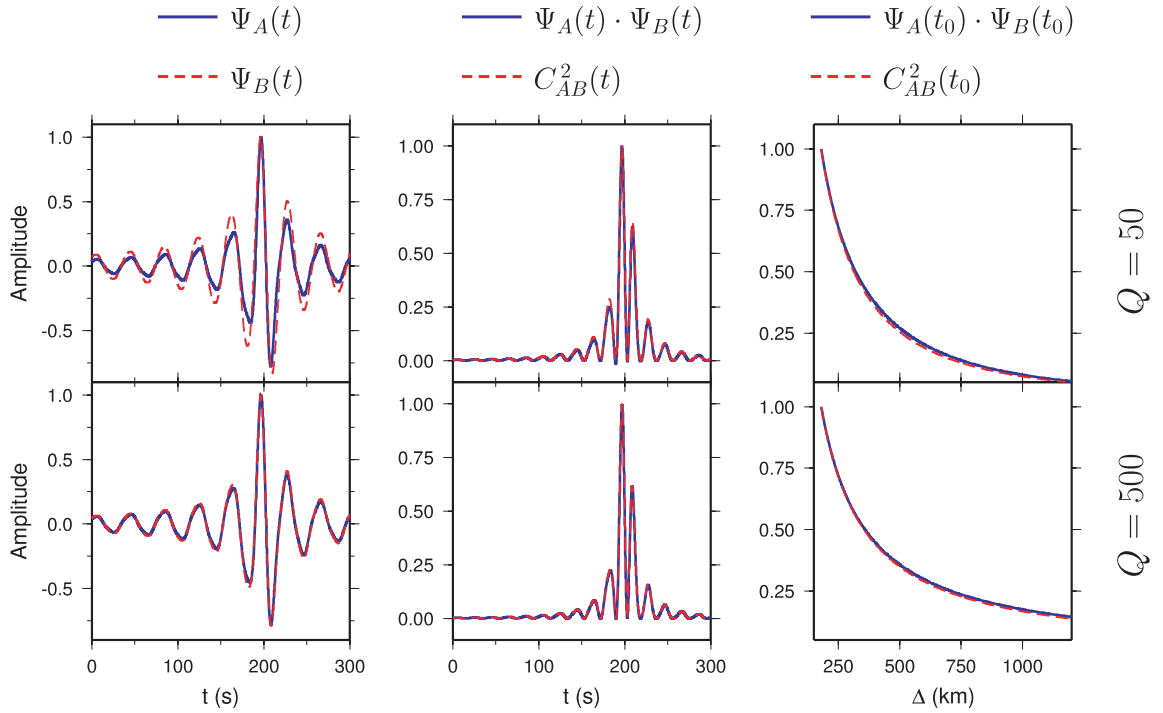


Figure 7. Comparison between $\Psi_A(t)$ and $\Psi_B(t)$ (left-hand side column), $\Psi_A(t)\Psi_B(t)$ and $C_{AB}^2(t)$ (middle column) and $\Psi_A(t_0)\Psi_B(t_0)$ and $C_{AB}^2(t_0)$ as a function of Δ (right-hand side column) for two different values of $\Delta\omega$: $1.99 \omega_0$ (a) and ω_0 (b). The medium is anelastic. Two different values of Q are tested: 50 and 500. The other parameters are $f_0 = 0.05$ Hz and $v = 3$ km s $^{-1}$. All the waveforms (functions of t) are computed using $\Delta = 10 \lambda_0 = 600$ km.

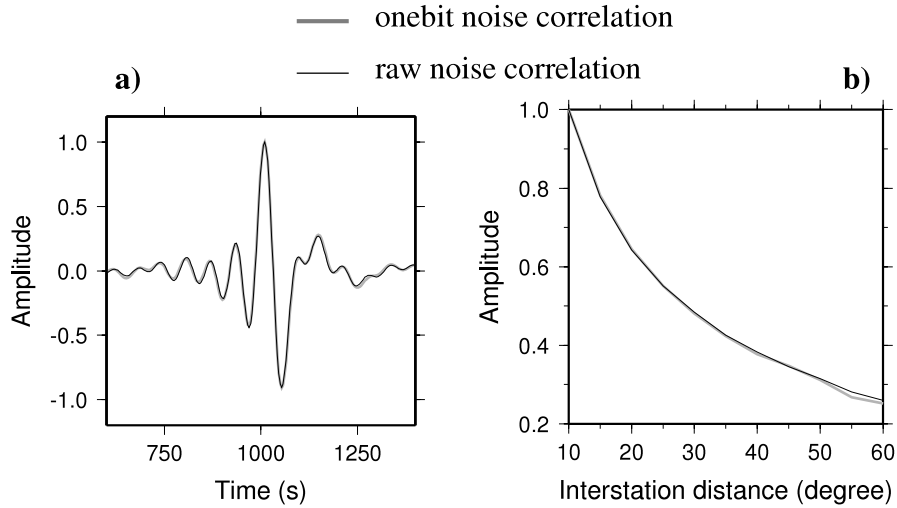


Figure 8. Same as Fig. 5 in an anelastic earth.

given distribution to estimate both the waveform and the relative amplitude of one-bit noise correlations. This is in agreement with numerical studies carried out by Cupillard & Capdeville (2010).

In pure elastic media, eq. (18) is rigorously demonstrated. This equation is at the basis of our theory. It relates the square of the raw noise correlation $C_{AB}^2(t)$ and the variances $\Psi_A(t)$ and $\Psi_B(t)$. In anelastic media, $C_{AB}^2(t)$ is plotted and compared with the product $\Psi_A(t) \times \Psi_B(t)$ to check if the expression is valid (Fig. 7, middle column) but no rigorous demonstration is provided. On the four plots, the curves are very similar but they are not exactly the same. It is particularly visible when $\Delta\omega = \omega_0$ and $Q = 50$. This means that eq. (18) is only true to first order when using the expressions we found for the coherent variances in the anelastic case. Further investigations are needed to check if exact expressions of the variances exist in this case. Here, we use approximate expressions and we assume that the small discrepancies in the amplitude decays (Fig. 7, right-hand side column) are due to these approximations. If it is not the case (i.e. if the small discrepancies are effective in practice), then measurements of intrinsic attenuation from one-bit noise correlations will give approximate values of Q .

The derivation shown in this paper involves surface waves. It should be easy to write it for body waves as well. The GF between two points U and V would be

$$G_{UV}(\omega) = \frac{\exp(-ikd_{UV})}{d_{UV}}, \quad (78)$$

and the coherent noise could be defined using the following correlation coefficient:

$$\rho_{AB}(\theta; t) = \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \cos\left[\omega\left(t + \frac{\Delta \sin \theta}{v}\right)\right] d\omega \quad (79)$$

$$= \Delta\omega \operatorname{sinc}\left[\frac{\Delta\omega}{2}\left(t + \frac{\Delta \sin \theta}{v}\right)\right] \cos\left[\omega_0\left(t + \frac{\Delta \sin \theta}{v}\right)\right]. \quad (80)$$

This last expression is very similar to the correlation coefficient introduced in the review paper by Larose (2006). This means that our definition of coherency is actually the same as previous definitions (Snieder 2004; Roux *et al.* 2005; Sabra *et al.* 2005). Here, we just go into the details of the concept to extract the properties we need to understand the one-bit noise correlation.

The analytical expressions of the standard deviations we provide in this work are obtained in the case of a 1-D layered medium. An extension of these expressions to a full 3-D case is not straightforward. It would require a description of complex GFs, and scattering should be taken into account (Halliday & Curtis 2009). Scatterers act as secondary noise sources and drastically change the size of the coherent and incoherent hyperbolas. When dealing with coda recordings, these hyperbolas depend on the part of the coda that is in use because such recordings are non-stationary. In this case, our statistical model would need standard deviations which depend on the time in the noise record, as suggested by Derode *et al.* (1999) and Larose *et al.* (2008). To study full 3-D cases, numerical simulations are actually necessary. Using numerical tools to evaluate standard deviations and understand what the one-bit noise correlation exactly contains in complex media will be the topic of future work.

ACKNOWLEDGMENTS

The authors would like to thank Francisco J. Sánchez-Sesma, Vedran Lekic, Yann Capdeville, Eric Larose and Michel Campillo for valuable discussions. Constructive comments from two anonymous reviewers and editor David Halliday also helped to improve the manuscript. This work was partially supported by NSF grant EAR-0738284. This is Berkeley Seismological Laboratory contribution 10-14.

REFERENCES

- Abramowitz, M. & Stegun, I.A., 1972. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing, Dover, New York.
- Aki, K. & Richards, P.G., 2002. *Quantitative Seismology*, 2nd edn, University Science Books, Sausalito.
- Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, **169**, 1239–1260, doi:10.1111/j.1365-246X.2007.03374.x.
- Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2008. Broad-band ambient noise surface wave tomography across the United States, *J. geophys. Res.*, **113**, B05306, doi:10.1029/2007JB005248.
- Bond, F.E. & Cahn, C.R., 1958. On sampling the zeros of bandwidth limited signals. *IRE Trans. Inform. Theory*, **4**, 110–113.
- Brooks, L. & Gerstoft, P., 2009. Green's function approximation from cross-correlations of 20–100 Hz noise during a tropical storm, *J. acoust. Soc. Am.*, **125**(2), 723–734.
- Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**, 547–549.
- Cho, K., Herrmann, R.B., Ammon, C.J. & Lee, K., 2007. Imaging the upper crust of the Korean peninsula by surface-wave tomography, *Bull. seism. Soc. Am.*, **67**, 198–207.
- Cupillard, P. & Capdeville, Y., 2010. On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach, *Geophys. J. Int.*, **181**(3), 1687–1700, doi:10.1111/j.1365-246X.2010.04586.x.
- Derode, A., Tourin, A. & Fink, M., 1999. Ultrasonic pulse compression with one-bit time reversal through multiple scattering, *J. appl. Phys.*, **85**, 6343–6352.
- Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M. & Fink, M., 2003. Recovering the Green's function from field-field correlations in an open scattering medium (L), *J. acoust. Soc. Am.*, **113**, 2973–2976.
- Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model. *Phys. Earth planet. Inter.*, **25**, 297–356.
- Friedrich, A., Krüger, F. & Klinge, K., 1998. Ocean-generated microseismic noise located with the Gräfenberg array, *J. Seismol.*, **2**, 47–64.
- Gerstoft, P. & Tanimoto, T., 2007. A year of microseisms in southern California, *Geophys. Res. Lett.*, **34**, L20304, doi:10.1029/2007GL031091.
- Gradshteyn, I.S. & Ryzhik, I.M., 2007. *Table of Integrals, Series and Products*, 7th edn, eds Jeffrey, A. & Zwillinger, D., Academic Press, Burlington.
- Halliday, D. & Curtis, A., 2008. Seismic interferometry, surface waves and source distribution, *Geophys. J. Int.*, **175**, 1067–1087, doi:10.1111/j.1365-246X.2008.03918.x.
- Halliday, D. & Curtis, A., 2009. Seismic interferometry of scattered surface waves in attenuative media, *Geophys. J. Int.*, **178**, 419–446, doi:10.1111/j.1365-246X.2009.04153.x.
- Kedar, S., Longuet-Higgins, M., Webb, F., Graham, N., Clayton, R. & Jones, C., 2008. The origin of deep ocean microseisms in the North Atlantic Ocean, *Proc. R. Soc. A*, **464**, 777–793.
- Kimman, W.P. & Trampert, J., 2010. Approximations in seismic interferometry and their effects on surface waves, *Geophys. J. Int.*, **182**, 461–476, doi:10.1111/j.1365-246X.2010.04632.x.
- Larose, E., 2006. Mesoscopics of ultrasound and seismic waves: application to passive imaging, *Ann. Phys. Fr.*, **31**(3), 1–126, doi:10.1051/anphys:2007001.
- Larose, E., Derode, A., Campillo, M. & Fink, M., 2004. Imaging from one-bit correlations of wide-band diffuse wavefields, *J. appl. Phys.*, **95**(12), 8393–8399.
- Larose, E., Roux, P. & Campillo, M., 2007. Reconstruction of Rayleigh-Lamb dispersion spectrum based on noise obtained from an air-jet forcing, *J. acoust. Soc. Am.*, **122**(6), 3437–3444.
- Larose, E., Roux, P., Campillo, M. & Derode, A., 2008. Fluctuations of correlations and Green's function reconstruction: role of scattering, *J. appl. Phys.*, **103**, 114907, doi:10.1063/1.2939267.
- Lin, F., Ritzwoller, M.H., Townend, J., Savage, M. & Bannister, S., 2007. Ambient noise Rayleigh wave tomography of New Zealand, *Geophys. J. Int.*, **170**, 649–666, doi:10.1111/j.1365-246X.2007.03414.x.
- Lin, F., Moschetti, M.P. & Ritzwoller, M.H., 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.*, **173**, 281–298, doi:10.1111/j.1365-246X.2008.03720.x.
- Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlation of a diffuse field, *J. acoust. Soc. Am.*, **110**, 3011–3017.
- Longuet-Higgins, M.S., 1950. A theory on the origin of microseisms, *Phil. Trans. R. Soc. Lond. A.*, **243**, 1–35.
- Matzel, E., 2008. Attenuation tomography using ambient noise correlation, *Seism. Res. Lett.*, **79**(2), 358.
- Pedersen, H.A., Krüger, F. & the SVEKALAPKO Seismic Tomography Working Group, 2007. Influence of the seismic noise characteristics on noise correlations in the Baltic shield, *Geophys. J. Int.*, **168**, 197–210, doi:10.1111/j.1365-246X.2006.03177.x.
- Prieto, G.A., Lawrence, J.F. & Beroza, G.C., 2009. Anelastic Earth structure from the coherency of the ambient seismic field, *J. geophys. Res.*, **144**, B07303, doi:10.1029/2008JB006067.
- Roux, P., Sabra, K.G. & Kuperman, W.A., 2005. Ambient noise cross-correlation in free space: theoretical approach, *J. acoust. Soc. Am.*, **117**(1), 79–84, doi:10.1121/1.1830673.
- Sabra, K.G., Roux, P. & Kuperman, W.A., 2005. Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide, *J. acoust. Soc. Am.*, **117**(1), 164–174, doi:10.1121/1.1835507.
- Schulte-Pelkum, V., Earle, P.S. & Vernon, F.L., 2004. Strong directivity of ocean-generated seismic noise, *Geochem. Geophys. Geosyst.*, **5**, Q03004, doi:10.1029/2003GC000520.
- Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, **31**, L07614, doi:10.1029/2004GL019491.
- Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface wave tomography from ambient seismic noise, *Science*, **307**, 1615–1618.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E*, **69**, 046610, doi:10.1103/PhysRevE.69.046610.
- Stehly, L., Campillo, M. & Shapiro, N., 2006. A study of the seismic noise from its long-range correlation properties, *J. geophys. Res.*, **111**, B10306, doi:10.1029/2005JB004237.
- Stehly, L., Fry, B., Campillo, M., Shapiro, N., Guilbert, J., Boschi, L. & Giardini, D., 2009. Tomography of the alpine region from observations of seismic ambient noise, *Geophys. J. Int.*, **178**, 338–350, doi:10.1111/j.1365-246X.2009.04132.x.
- Stutzmann, E., Schimmel, M., Patau, G. & Maggi, A., 2009. Global climate imprint on seismic noise, *Geochem. Geophys. Geosyst.*, **10**, Q11004, doi:10.1029/2009GC002619.
- van Tiggelen, B.A., 2003. Green function retrieval and time-reversal in a disordered world, *Phys. Rev. Lett.*, **91**(24), 243904, doi:10.1103/PhysRevLett.91.243904.
- Voelcker, H.B., 1966a. Toward a unified theory of modulation. Part I: phase-envelope relationships, *Proc. IEEE*, **54**(3), 340–353.
- Voelcker, H.B., 1966b. Toward a unified theory of modulation. Part II: zero manipulation, *Proc. IEEE*, **54**(5), 735–755.
- Voelcker, H.B. & Requicha, A.A.G., 1973. Clipping and signal determinism: two algorithms requiring validation, *IEEE Trans. Commun.*, COM-21(6), 738–744.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross-correlation, *Phys. Rev. Lett.*, **93**, 254–301, doi:10.1103/PhysRevLett.93.254301.
- Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized earth model, *Geophys. J. R. astr. Soc.*, **78**, 641–660.
- Yang, Y. & Ritzwoller, M.H., 2008. Characteristics of ambient seismic noise as a source for surface wave tomography, *Geochem. Geophys. Geosyst.*, **9**, Q02008, doi:10.1029/2007GC001814.

16 *P. Cupillard, L. Stehly and B. Romanowicz*

- Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.*, **168**, 259–274.
- Yao, H. & van der Hilst, R.D., 2009. Analysis of ambient noise energy distribution and phase velocity bias in ambient noise tomography, with application to SE Tibet, *Geophys. J. Int.*, **179**, 1113–1132.
- Yao, H., van der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave ar-

- ray tomography in SE Tibet from ambient seismic noise and two-station analysis, I: phase velocity maps, *Geophys. J. Int.*, **166**, 732–744, doi:10.1111/j.1365-246X.2006.03028.x.
- Yao, H., van der Hilst, R.D. & de Hoop, M.V., 2008. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis, II: crustal and upper-mantle structure, *Geophys. J. Int.*, **173**, 205–219, doi:10.1111/j.1365-246X.2007.03696.x.

APPENDIX A: PROOF OF EXPRESSION (18)

The goal of this appendix is to demonstrate that, in the case of a uniform distribution of noise sources, we have

$$\Psi_A(t)\Psi_B(t) = C_{AB}^2(t), \quad (\text{A1})$$

with

$$\Psi_A(t) = \iint \sigma_A^2(\mathbf{r})\rho_{AB}(\mathbf{r}, t) d\mathbf{r} \quad (\text{A2})$$

and

$$\Psi_B(t) = \iint \sigma_B^2(\mathbf{r})\rho_{AB}(\mathbf{r}, t) d\mathbf{r}. \quad (\text{A3})$$

We start from the Fourier transform of $\Psi_A(t)$:

$$\Psi_A(\omega) = \frac{2}{\pi k} \iint \frac{|S(\mathbf{r}; \omega)|^2}{d_A(\mathbf{r})} \exp[ik(d_A(\mathbf{r}) - d_B(\mathbf{r}))] d\mathbf{r}. \quad (\text{A4})$$

When $|S(\mathbf{r}; \omega)|^2$ is a smooth function of \mathbf{r} , this integral can be evaluated using the stationary phase approximation:

$$\begin{aligned} \Psi_A(\omega) &= \frac{2}{ik} \frac{\exp[i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \int_{\Delta}^{\infty} |S(x, y=0; \omega)|^2 \sqrt{\frac{x-\Delta}{x}} dx \\ &\quad - \frac{2}{ik} \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \int_{-\infty}^0 |S(x, y=0; \omega)|^2 \sqrt{\frac{x-\Delta}{x}} dx. \end{aligned} \quad (\text{A5})$$

Assuming a uniform distribution of noise sources ($|S(x, y=0; \omega)|^2 = |S(\omega)|^2$), it follows that

$$\begin{aligned} \Psi_A(\omega) &= 2D \frac{v(\omega)|S(\omega)|^2}{i\omega} \frac{\exp[i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \left[1 - \frac{\Delta}{2D} \ln\left(\frac{4D}{\Delta}\right) + o\left(\frac{\Delta}{D}\right) \right] \\ &\quad - 2D \frac{v(\omega)|S(\omega)|^2}{i\omega} \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \left[1 + \frac{\Delta}{2D} \ln\left(\frac{4D}{\Delta}\right) + o\left(\frac{\Delta}{D}\right) \right], \end{aligned} \quad (\text{A6})$$

where $D \gg \Delta$ is a distance we introduce to perform the integration over x and prevent this integration to diverge.

In the same way we find

$$\begin{aligned} \Psi_B(\omega) &= \frac{2}{ik} \frac{\exp[i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \int_{\Delta}^{\infty} |S(x, y=0; \omega)|^2 \sqrt{\frac{x}{x-\Delta}} dx \\ &\quad - \frac{2}{ik} \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \int_{-\infty}^0 |S(x, y=0; \omega)|^2 \sqrt{\frac{x}{x-\Delta}} dx \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} &= 2D \frac{v(\omega)|S(\omega)|^2}{i\omega} \frac{\exp[i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \left[1 + \frac{\Delta}{2D} \ln\left(\frac{4D}{\Delta}\right) + o\left(\frac{\Delta}{D}\right) \right] \\ &\quad - 2D \frac{v(\omega)|S(\omega)|^2}{i\omega} \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \left[1 - \frac{\Delta}{2D} \ln\left(\frac{4D}{\Delta}\right) + o\left(\frac{\Delta}{D}\right) \right]. \end{aligned} \quad (\text{A8})$$

Writing (A6) and (A8) in the time domain, we finally get

$$\Psi_A(t)\Psi_B(t) = 4D^2 \left\{ \int \frac{v(\omega)|S(\omega)|^2}{i\omega} \left[\frac{\exp[i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} - \frac{\exp[-i(k\Delta + \frac{\pi}{4})]}{\sqrt{\frac{\pi}{2}k\Delta}} \right] e^{i\omega t} d\omega \right\}^2 \left[1 + o\left(\frac{\Delta}{D}\right) \right]. \quad (\text{A9})$$

We recognize the first term on the right-hand side to be the square of the inverse Fourier transform of eq. (9), that is to say the square of the correlation between $A(t)$ and $B(t)$.

APPENDIX B: PRECISIONS ON EQ. (27)

The cosine-integral special function ci is defined in \mathbb{R}^{+*} by (Abramowitz & Stegun 1972).

$$\text{ci}(u) = - \int_u^\infty \frac{\cos x}{x} dx. \quad (\text{B1})$$

The extension of the definition of ci to \mathbb{R}^{-*} is straightforward. Indeed, the change of variable $y = -x$ in eq. (B1) yields

$$\text{ci}(u) = - \int_{-u}^{-\infty} \frac{\cos y}{y} dy, \quad (\text{B2})$$

so it is natural to pose $\text{ci}(u) = \text{ci}(-u) \forall u \in \mathbb{R}^{-*}$.

To fully describe expression (27), we now have to define the function $\text{ci}[(\omega_0 + \frac{\Delta\omega}{2})(t + \alpha)] - \text{ci}[(\omega_0 - \frac{\Delta\omega}{2})(t + \alpha)]$ in 0. To do so, we evaluate the limit of this function when $t \rightarrow -\alpha$. By definition we have

$$\text{ci} \left[\left(\omega_0 + \frac{\Delta\omega}{2} \right) (t + \alpha) \right] - \text{ci} \left[\left(\omega_0 - \frac{\Delta\omega}{2} \right) (t + \alpha) \right] = \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{\cos[\omega(t + \alpha)]}{\omega} d\omega, \quad (\text{B3})$$

so

$$\lim_{t \rightarrow -\alpha} \text{ci} \left[\left(\omega_0 + \frac{\Delta\omega}{2} \right) (t + \alpha) \right] - \text{ci} \left[\left(\omega_0 - \frac{\Delta\omega}{2} \right) (t + \alpha) \right] = \lim_{t \rightarrow -\alpha} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{d\omega}{\omega} \quad (\text{B4})$$

$$= \ln \left(\frac{\omega_0 + \frac{\Delta\omega}{2}}{\omega_0 - \frac{\Delta\omega}{2}} \right). \quad (\text{B5})$$

The limit exists and is finite, so eq. (27) is defined and continuous in 0.

APPENDIX C: PROOF OF EQ. (52)

The probability density function of the random variables $A^t(\tau)$ and $\overline{A^t}(\tau)$ is

$$f(x) = \frac{1}{\sigma_{A^t} \sqrt{2\pi}} \exp \left(\frac{-x^2}{2\sigma_{A^t}^2} \right) \quad (\text{C1})$$

and

$$g(x) = \frac{1}{\sigma_{\overline{A^t}} \sqrt{2\pi}} \exp \left(\frac{-x^2}{2\sigma_{\overline{A^t}}^2} \right), \quad (\text{C2})$$

respectively.

P_A^t is the probability that $|A^t(\tau)| > |\overline{A^t}(\tau)|$ so

$$P_A^t = \int_0^\infty 2g(x) \int_x^\infty 2f(y) dy dx \quad (\text{C3})$$

$$= 1 - \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-x^2} \text{erf} \left(\frac{\sigma_{\overline{A^t}}}{\sigma_{A^t}} x \right) dx. \quad (\text{C4})$$

This last integral can be calculated following Gradshteyn & Ryzhik (2007). Then

$$P_A^t = 1 - \frac{2}{\pi} \arctan \left(\frac{\sigma_{\overline{A^t}}}{\sigma_{A^t}} \right). \quad (\text{C5})$$

In the same way we find the probability that $|B^t(\tau)| > |\overline{B^t}(\tau)|$:

$$P_B^t = 1 - \frac{2}{\pi} \arctan \left(\frac{\sigma_{\overline{B^t}}}{\sigma_{B^t}} \right). \quad (\text{C6})$$

Finally, we obtain

$$|C_{AB}^{ob}(t)| = n \left[1 - \frac{2}{\pi} \arctan \left(\frac{\sigma_{\overline{A^t}}}{\sigma_{A^t}} \right) \right] \left[1 - \frac{2}{\pi} \arctan \left(\frac{\sigma_{\overline{B^t}}}{\sigma_{B^t}} \right) \right]. \quad (\text{C7})$$

APPENDIX D: Ψ_A AND Ψ_B IN THE ANELASTIC CASE

$$\Psi_A(\omega) = \frac{2}{\pi k} \iint \frac{|S(\mathbf{r}; \omega)|^2 \exp \left(\frac{-k d_A(\mathbf{r})}{Q} \right)}{d_A(\mathbf{r})} \exp [i k (d_A(\mathbf{r}) - d_B(\mathbf{r}))] d\mathbf{r}. \quad (\text{D1})$$

18 *P. Cupillard, L. Stehly and B. Romanowicz*

Using the stationary phase approximation in the Cartesian coordinate system shown in Fig. 2, we can reduce this expression to

$$\Psi_A(\omega) = -\frac{2}{ik} \frac{\exp\left[-i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} \int_{-\infty}^0 |S(x, y=0; \omega)|^2 \exp\left(\frac{kx}{Q}\right) \sqrt{\frac{x-\Delta}{x}} dx. \quad (\text{D2})$$

To get this expression, we considered the signal emerging from the sources at $x < 0$ only. Assuming a uniform distribution of noise sources ($|S(x, y=0; \omega)|^2 = |S(\omega)|^2$), we are able to perform the integral over x following Gradshteyn & Ryzhik (2007). Expression (D2) then becomes

$$\Psi_A(\omega) = -\frac{2\Delta}{ik} \Gamma\left(\frac{1}{2}\right) |S(\omega)|^2 U\left(\frac{1}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\exp\left[-i\left(k\Delta + \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}}, \quad (\text{D3})$$

where Γ is the gamma function and U is the confluent hypergeometric function of the second kind (Abramowitz & Stegun 1972). Writing eq. (D3) in the time domain, we get

$$\Psi_A(t) = -2\Delta \Gamma\left(\frac{1}{2}\right) \int \frac{|S(\omega)|^2}{ik} U\left(\frac{1}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\exp\left[i\left(\omega t - k\Delta - \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} d\omega. \quad (\text{D4})$$

In the same way we find

$$\Psi_B(t) = -2\Delta \Gamma\left(\frac{3}{2}\right) \int \frac{|S(\omega)|^2}{ik} U\left(\frac{3}{2}, 2; \frac{k\Delta}{Q}\right) \frac{\exp\left[i\left(\omega t - k\Delta - \frac{\pi}{4}\right)\right]}{\sqrt{\frac{\pi}{2}k\Delta}} \exp\left(-\frac{k\Delta}{Q}\right) d\omega. \quad (\text{D5})$$

RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale

Paul Cupillard, Elise Delavaud,* Gaël Burgos, Geatano Festa,† Jean-Pierre Vilotte, Yann Capdeville‡ and Jean-Paul Montagner

Equipe de sismologie, IPGP, 1 rue Jussieu, 75238 Paris cédex 05, France. E-mail: paulcup@ipgp.jussieu.fr

Accepted 2011 November 20. Received 2011 November 17; in original form 2011 April 29

SUMMARY

The spectral element method, which provides an accurate solution of the elastodynamic problem in heterogeneous media, is implemented in a code, called RegSEM, to compute seismic wave propagation at the regional scale. By regional scale we here mean distances ranging from about 1 km (local scale) to 90° (continental scale). The advantage of RegSEM resides in its ability to accurately take into account 3-D discontinuities such as the sediment-rock interface and the Moho. For this purpose, one version of the code handles local unstructured meshes and another version manages continental structured meshes. The wave equation can be solved in any velocity model, including anisotropy and intrinsic attenuation in the continental version. To validate the code, results from RegSEM are compared to analytical and semi-analytical solutions available in simple cases (e.g. explosion in PREM, plane wave in a hemispherical basin). In addition, realistic simulations of an earthquake in different tomographic models of Europe are performed. All these simulations show the great flexibility of the code and point out the large influence of the shallow layers on the propagation of seismic waves at the regional scale.

RegSEM is written in Fortran 90 but it also contains a couple of C routines. It is an open-source software which runs on distributed memory architectures. It can give rise to interesting applications, such as testing regional tomographic models, developing tomography using either passive (i.e. noise correlations) or active (i.e. earthquakes) data, or improving our knowledge on effects linked with sedimentary basins.

Key words: Site effects; Computational seismology; Wave propagation; Europe.

1 INTRODUCTION

Solving the wave equation in realistic geological media is a crucial issue to properly model and study the propagation of seismic waves. At large scale, the effect of both crust and upper mantle 3-D structures on seismograms has been known for a long time (e.g. Montagner & Tanimoto 1991). At local scale, geological site conditions are now recognized as one of the dominant factors controlling the variations in ground motion (e.g. Olsen 2000). The accurate incorporation of geological structures in wave propagation modelling would therefore greatly improve the knowledge in fields such as tomography and site effects estimation.

Numerous techniques, like ray tracing or normal mode summation, have been developed for decades to compute synthetic

seismograms. Ray tracing assumes that the seismic wavefield can be modelled as a large number of very narrow beams (e.g. Cervený 2001). For many purposes, this technique is very convenient, but it relies on a high frequency approximation, which means that it is only valid when seismic wavelength is much smaller than the scale of heterogeneity. In the case of low-frequency waves with large Fresnel zones, the ray theory no longer holds. The normal mode summation technique (e.g. Gilbert 1971), in addition with a high-order perturbation theory (Lognonné & Romanowicz 1990; Lognonné 1991; Clévéde & Lognonné 1996), is able to model waves with large Fresnel zones in 3-D Earth models, but the computation cost of such a technique is quickly prohibitive as the number of modes to couple increases with frequency. Moreover, normal mode perturbation methods are limited to weak lateral heterogeneities.

Direct numerical solutions have also been investigated to solve the wave equation. Finite differences have been applied (e.g. Alterman & Kara 1968; Boore 1972; Kelly *et al.* 1976; Virieux 1984, 1986; Olsen & Archuleta 1996; Moczo *et al.* 2007) but they present intrinsic problems in dealing with strong and deformed interfaces like basin edges, the Moho, the free surface and solid-fluid

*Now at: ETH, Institut für Geophysik, 8092 Zürich, Switzerland.

†Now at: RISSC-Lab, Department of Physics, University Federico II, Naples, Italy.

‡Now at: LPGN, UFR sciences et techniques, Université de Nantes, France.

2 P. Cupillard et al.

discontinuities. Such limitations do not exist in finite element methods, but the low polynomial order classically used in this kind of approaches make them inaccurate and dispersive when applied to elastodynamic problems (Lysmer & Drake 1972; Dupond 1973; Backer 1976; Marfurt 1984; Toshinawa & Ohmachi 1992). Since the 1990s, efforts have been focused on developing higher-order numerical modelling of seismic wave propagation. An important result from these efforts is the discontinuous Galerkin method (e.g. Dumbser & Käser 2006). Another major result, which is used all along this paper, is the spectral element method (SEM). Initially introduced in fluid mechanics (Patera 1984; Maday & Patera 1989), this method has been successfully applied to elastodynamics with the increasing concern of developing numerical techniques ensuring both a great precision and a reasonable numerical cost (Seriani & Priolo 1994; Faccioli *et al.* 1997; Komatitsch & Vilotte 1998; Seriani 1998; Komatitsch & Tromp 1999). The SEM was first applied at the global scale (Chaljub 2000; Komatitsch & Tromp 2002a,b; Komatitsch *et al.* 2002; Capdeville *et al.* 2003; Chaljub *et al.* 2003; Chaljub & Valette 2004). In the more recent years, applications to local (Komatitsch *et al.* 2004; Delavaud *et al.* 2006; Lee *et al.* 2008; Stupazzini *et al.* 2009; Chaljub *et al.* 2010; Peter *et al.* 2011) and continental scales (Chen *et al.* 2007a; Fichtner *et al.* 2009a) appeared. Such applications of the SEM proved that a great precision and a weak numerical dispersion can be obtained.

The SEM is a major contribution to seismology because it allows to compute the whole seismic wavefield propagating in a 3-D Earth model with no approximation on the wave equation (except minor numerical approximations). In the recent years, it was used to solve the inverse problem and get images of the Earth interior: Tape *et al.* (2009) and Tape *et al.* (2010) developed a model in southern California by inverting traveltimes measurements down to 2 s period; Fichtner *et al.* (2009b) and Fichtner *et al.* (2010) obtained a model of the Australian region using a full waveform inversion down to 30 s period. To avoid the large number of simulations classically required to compute all the Fréchet derivatives, these authors implemented the adjoint technique (Tarantola 1984; Tromp *et al.* 2005; Fichtner *et al.* 2006a,b). This makes the computation of the gradient of the misfit function independent of the number of stations and parameters of the model. When dealing with a large set of sources, an alternative to the adjoint technique is the scattering-integral approach (Chen *et al.* 2007b). In any case, with the current computational power, one cannot solve the inverse problem using the SEM with a classic procedure.

In this paper, we focus our attention on the forward modelling. When using the SEM to compute 3-D seismic wave propagation, the main practical issue consists in meshing the medium using hexahedra. Indeed, the mesh has to honour the discontinuities of the geological model under study to fully benefit from the accuracy of the method and properly model effects associated with these discontinuities such as wave diffraction. Because realistic models of the Earth often have complex geometry, lots of efforts and time are usually needed to build an appropriate mesh. Casarotti *et al.* (2008b) developed automatic procedures to create 3-D unstructured hexahedral meshes, but it is still not possible to generate in a fully automatic way meshes that would honour detailed geological discontinuities such as realistic sediment-rock interfaces. The goal of this paper is to promote a code, RegSEM, that can accurately take into account 3-D discontinuities in regional meshes and then compute seismic waves within them using the SEM. More precisely, RegSEM has two versions: a continental version which is able to generate structured meshes of crustal and mantle structures separated by a 3-D Moho in spherical geometry, and a local version which uses an

external mesh generator, CUBIT (<http://cubit.sandia.gov>), to produce 3-D unstructured meshes essentially designed to study the seismic response of sedimentary basins.

In a first part, RegSEM's features are described, including the different kinds of meshes the code can provide and/or handle. In a second part, waveforms computed with RegSEM in simple elastic models are compared to analytical or semi-analytical solutions. These comparisons allow to validate the code. The last part presents simulations of a real earthquake in different tomographic models of Europe. A comparison between synthetic seismograms and real data enables to investigate the improvement of the fit when using 3-D models and point out the great influence of the shallow layers. The large number of simulations shown in this work serves as a set of examples to put in evidence the capability and the great versatility of RegSEM. The code sources, in addition with a manual and several examples, can be downloaded at www.ipgp.fr/~paulcup/RegSEM.html.

2 REGSEM'S FEATURES

2.1 The spectral element method

The SEM was developed in fluid dynamics in the 1980s (Patera 1984; Maday & Patera 1989) and was adapted to elastodynamics in the 1990s (Seriani & Priolo 1994; Faccioli *et al.* 1997; Komatitsch & Vilotte 1998; Seriani 1998; Komatitsch & Tromp 1999). The SEM is similar to a finite element method. It is based on a primal variational formulation of the equations of motion. This formulation enables to naturally take into account both interface and free boundary surface conditions, allowing a good resolution of evanescent interface and surface waves.

The discretization process implies the decomposition of the spatial domain into non-overlapping elements. Classical implementations of the SEM in computational seismology are based on hexahedral elements to benefit from advantageous properties of tensorization. Although hexahedra are less favourable than tetrahedra for meshing geometrically complex structures, a certain flexibility is ensured by a local geometrical transformation from a reference element (unit cube) to any deformed element, as detailed in Section 2.2.1. Unstructured meshes offer additional possibilities, as seen in Section 2.2.2.

Associated with the domain decomposition, the functional discretization is based on a piecewise high-order polynomial approximation. The specificity of the SEM holds in the choice of basis functions intimately related to the Gauss-Lobatto-Legendre (GLL) quadrature used to evaluate the integrals in the variational formulation. The basis is obtained from the orthogonal Lagrange polynomials associated with $(N + 1)$ interpolation nodes (where N is the polynomial order). These nodes are chosen to be the nodes of the GLL quadrature. Such nodes define a tensor product grid where the displacement, its spatial derivatives and products encountered in the variational formulation are evaluated. The choice of a Lagrangian interpolation associated with the GLL nodes gives the SEM a very interesting convergence property: an increase of the polynomial order leads to an exponential diminution of the aliasing error. This property, called spectral precision, gives its name to the method.

Inserting the polynomial interpolation and quadrature rules into the variational form of the equations of elastodynamics leads to a system of ordinary differential equations governing the evolution at the global nodal position, which can be written as follows:

$$\mathbb{M}\dot{\mathbf{V}} = \mathbb{F}^{\text{ext}} - \mathbb{F}^{\text{int}}(\mathbf{U}) + \mathbb{F}^{\text{trac}}(\mathcal{T}) \quad (1)$$

$$\dot{\mathbf{U}} = \mathbf{V}, \quad (2)$$

where \mathbf{U} , \mathbf{V} and \mathcal{T} are vectors containing the components of the displacement, velocity and traction at the global nodes, respectively. \mathbb{M} is the mass matrix. The vectors \mathbb{F}^{ext} and \mathbb{F}^{int} contain the external and internal forces, respectively, and \mathbb{F}^{trac} corresponds to the traction forces. The use of an orthogonal basis defined as the Lagrangian functions associated with the GLL nodes leads to a second interesting characteristic of the hexahedral version of the SEM: the mass matrix is diagonal. This property enables to use an explicit time stepping in which the inverse mass matrix \mathbb{M}^{-1} can be exactly computed. In RegSEM, like in most of the SEM implementations, this time stepping is a second-order finite difference scheme.

2.2 Meshing a chunk of the Earth with hexahedra

When using the SEM for 3-D complex geological models, the main difficulty consists in meshing the model under study. Indeed, to benefit from the high accuracy of the method and properly model the effects linked with the geology, the mesh has to adapt the velocity structure of the model, in particular the zeroth-order discontinuities such as the Moho and the sediment-rock interface. This task often requires significant efforts and time.

In this section, we describe two kinds of meshes in which RegSEM can simulate seismic wave propagation. The first kind corresponds to regular meshes of crustal and mantle structures in spherical geometry. These meshes are suitable to compute seismic waves for source–receiver distances ranging from 1 to 90°. RegSEM not only handles such meshes but it can also create them in a versatile way: the size and the location of the chunk, the spherical discontinuities and possible 3-D Moho and 3-D free surface can be defined by the user. The second kind of meshes requires more efforts from the user because it has to be generated externally. It corresponds to unstructured meshes. Such meshes can deal with more complex geometries. In the following, they are used to study the seismic response of sedimentary basins, but they could be designed for other applications.

Before giving more details on the capabilities of RegSEM, it must be pointed out two important conditions that all meshes have to fulfill to make the SEM accurate and stable:

(i) In classical applications of the SEM, $4 \leq N \leq 8$. For such values of N , at least five GLL nodes per wavelength are needed everywhere in the region to properly describe the seismic wavefield (e.g. Komatitsch & Vilotte 1998). This means that the size of the elements d and the polynomial order N are both constrained by the shortest wavelength λ_{min} propagated in the medium. This condition can be summarized by the following relation:

$$d \leq \frac{N}{5} \lambda_{\text{min}}. \quad (3)$$

(ii) To ensure the stability of the time-marching, the time step Δt of the finite difference scheme has to verify the Courant–Friedrichs–Lewy (CFL) condition:

$$\Delta t \leq C \left[\frac{\Delta x}{\alpha} \right]_{\text{min}}, \quad (4)$$

where C denotes the Courant number, usually chosen between 0.3 and 0.4, and $\left[\frac{\Delta x}{\alpha} \right]_{\text{min}}$ the minimum ratio of grid spacing Δx (distance between two GLL nodes) and P -wave speed α .

2.2.1 Regular meshes of crustal and mantle structures

RegSEM can provide a regular mesh of any chunk of the Earth whose lateral size is smaller than 90°. To do so, the code uses the so-called cubed sphere mapping (Sadourny 1972; Ronchi *et al.* 1996). For each element, this mapping allows to define the Cartesian coordinates of 27 control points. Using the Lagrange polynomials of degree 2 associated with these control points, the unit cube can be deformed and the shape and position of each element in the chunk can be defined. Such a classical procedure enables to easily design a structured and conformal mesh for any section of the Earth (Fig. 1). Moreover, ellipticity can be taken into account using the Clairaut's equation (Dahlen & Tromp 1998).

As mentioned above, the seismic discontinuities in the velocity model have to be honoured by the mesh. If all the discontinuities are spherical, then RegSEM is quite versatile: one just needs to introduce the radius of each discontinuity, and then the code fills the seismic layers with the appropriate number of elements. Of course, this number depends on the vertical size of the elements. This size is first equal to the horizontal size d_h introduced by the user, and then it is adjusted in each seismic layer to fit the thickness of the layer. In the case of PREM (Dziewonski & Anderson 1981), there is one more level of sophistication because not only the thickness of the layers but also the seismic velocities in the layers are used by the code to constrain and optimize the vertical size of the elements. Fig. 2 shows an example of a mesh of PREM. Some elements within this mesh appear to have a large aspect ratio (up to 5). In the context of the SEM, this is not a problem: accuracy is preserved thanks to the high spatial degree of the method (Oliveira & Seriani 2011) and stability is kept up because the shear deformation of the elements is small. Examples of simulations in PREM are shown in Section 3.2.

RegSEM can also mesh any surface and Moho topography (Fig. 3). This is an important feature because the crust has significant effects on surface waves (Montagner & Tanimoto 1991; Curtis *et al.* 1998; Komatitsch *et al.* 2002; Shapiro & Ritzwoller 2002; Marone & Romanowicz 2007; Ferreira *et al.* 2010), even at relatively long period (up to about 60 s). The capability to consider any model with a realistic crust is therefore a major benefit. To do so, the code uses only one layer of elements in the crust. This means that discontinuities within the crust, such as the sediment-rock interface and the upper-lower crust interface, cannot be taken into account. Moreover, the fact that only one layer of elements is used to mesh the crust limits the frequency content that the simulation can handle. For example, when performing a simulation in Tibet (where the crust is more than 70 km thick) with a polynomial order $N = 8$, the highest frequency to be propagated will be approximately 0.1 Hz. Examples of simulations in 3-D crustal models are shown in Section 4.

Realistic models of the Earth all have thin shallow layers. This is the case in Figs 2 and 3. Because of the CFL condition 4, these thin layers imply a very small time step, which makes the computation cost high:

(i) The mesh shown in Fig. 2 is designed to propagate a wavefield with a minimum wavelength $\lambda_{\text{min}} = 60$ km using a polynomial order $N = 4$. The Earth model is PREM. In this case, ratio $\frac{\Delta x}{\alpha}$ is minimum in a thin layer defined by two discontinuities at 15 and 24.4 km depth. Indeed, the elements used to mesh this layer have a vertical size of 9.4 km, which is very small compared to the size of the other elements. The layer has a P -wave speed $\alpha = 6.8$ km s⁻¹,

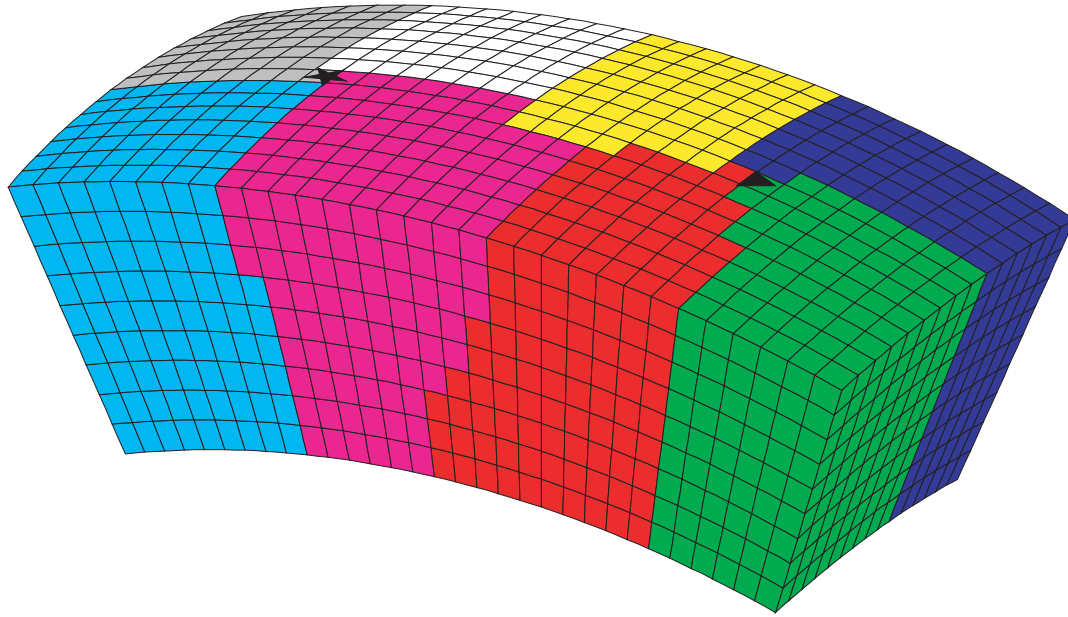
4 *P. Cupillard et al.*

Figure 1. A chunk of the Earth meshed by hexahedral elements. It is $20^\circ \times 40^\circ$ large and 1400 km thick. The elements are gathered under different colours. Each colour represents a subdomain. Here, the chunk is divided in eight subdomains. The star and the triangle correspond to the source and the station used in Section 3.1.

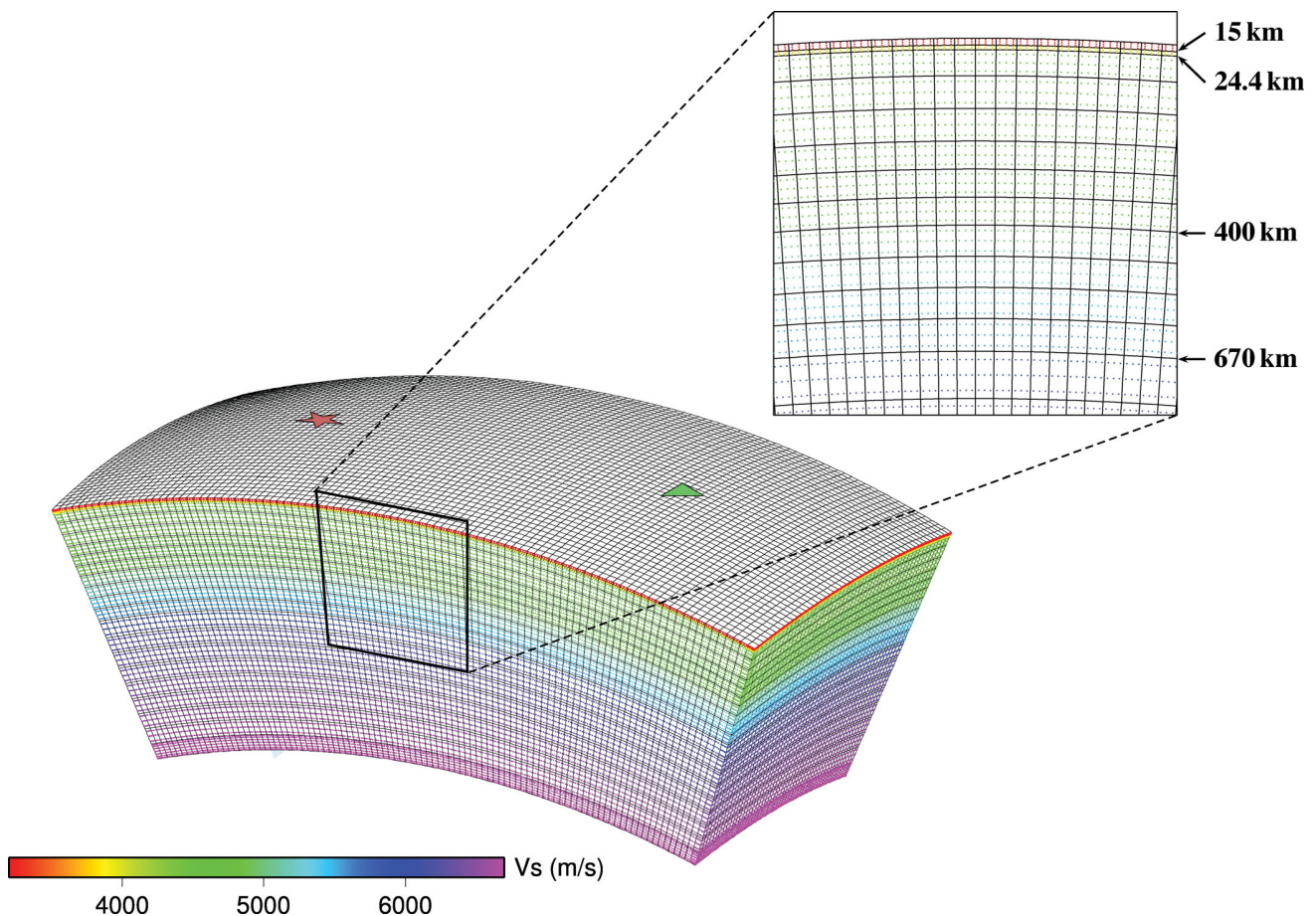


Figure 2. A mesh for PREM. The shallowest fluid layer has been replaced by the underlying solid. The chunk is $20^\circ \times 40^\circ$ large and 1400 km thick. The horizontal size of the elements is $d_h = 0.44^\circ$. This allows the propagation of a 20 s period wavefield using a polynomial order $N = 4$. The S -wave speed at each GLL node on the vertical border of the domain is plotted. A zoom into the upper part of the model shows the discontinuities. The star and the triangle correspond to the source and the station used in Section 3.2.

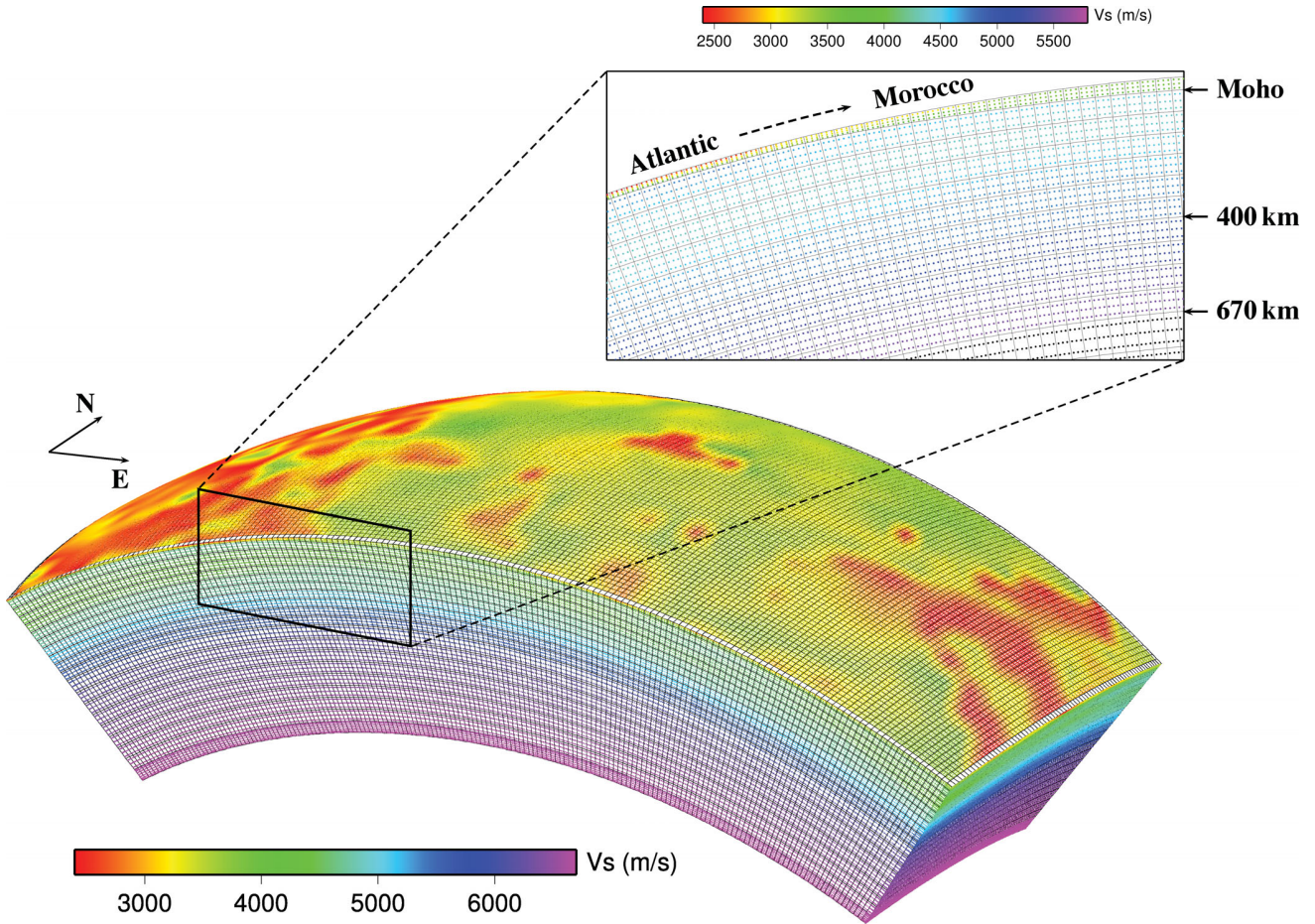


Figure 3. A mesh for CUB in the Atlantic–European region. The chunk is $30^\circ \times 70^\circ$ large and 1500 km thick. The horizontal size of the elements is $d_h = 0.35^\circ$. This allows the propagation of a 20 s period wavefield using a polynomial order $N = 4$. The S -wave speed at each GLL node on the border of the domain is plotted. A zoom into the upper part of the southern side of the chunk shows the discontinuities. The transition between a thin oceanic crust and a thick continental crust is visible. A map of the Moho corresponding to the present chunk is shown in Fig. 12.

so we find

$$\left[\frac{\Delta x}{\alpha} \right]_{\min} = \frac{9.4/\gamma_N}{6.8} s, \quad (5)$$

where γ_N is a coefficient which depends on N and which comes from the fact that the GLL nodes are non-evenly spaced in an element. For $N = 4$, $\gamma_N \simeq 6$. Combining eqs 4 and 5 and taking $C = 0.35$, we obtain $\Delta t \leq 0.080637$ s, which is small and makes the computation cost high. Note that in practice, RegSEM computes the time step automatically: it first finds $\left[\frac{\Delta x}{\alpha} \right]_{\min}$ by a grid search and then determines Δt using 4 with $C = 0.35$.

(ii) The mesh shown in Fig. 3 is designed to propagate a wavefield with a minimum wavelength $\lambda_{\min} = 50$ km using a polynomial order $N = 4$. The region under study is the Atlantic ocean and Europe. The velocity model is CUB (Shapiro & Ritzwoller 2002), which has a realistic crust. Such a crust is thin below the ocean: the minimum thickness is about 7 km, so there are some elements that have a very small vertical size compared to the size of the other elements. Ratio $\frac{\Delta x}{\alpha}$ is minimum in these elements. Assuming a P -wave velocity $\alpha = 5 \text{ km s}^{-1}$ in the oceanic crust, we find the following condition on the time step: $\Delta t \leq 0.081667$ s.

To avoid the small time step induced by a thin layer, recent works developed techniques to replace the layer by a thicker effective medium (Capdeville & Marigo 2008; Fichtner & Igel 2008;

Lekić *et al.* 2010). These techniques yield a new layer with a large and constant thickness at the top of the Earth. Although this kind of effective layers is easy to mesh, we do not show examples of computation in such media in this work.

2.2.2 Unstructured meshes of sedimentary basins

Basin effects are characterized by scattering, focusing and basin-edge induced surface waves which are closely associated with the geometry of the basin. These effects are recognized to be responsible for a long duration of the seismic signal in the basin and especially for large local amplifications. To accurately model and study these effects, especially at high frequencies, it is important that the sediment-rock interface is honoured. Considering the complexity of most of these discontinuities, 3-D unstructured meshes are necessary to achieve this goal. The local version of RegSEM has the ability to handle such unstructured meshes where topology is totally arbitrary. To do so, the code is written according to a strategy of independence against the Cartesian coordinates. This enables to handle the random orientation of the four different objects (elements, faces, edges and vertices) which compose the mesh. Defining such objects allows to assign specific actions to each of them, such as Neumann conditions (Section 2.5).

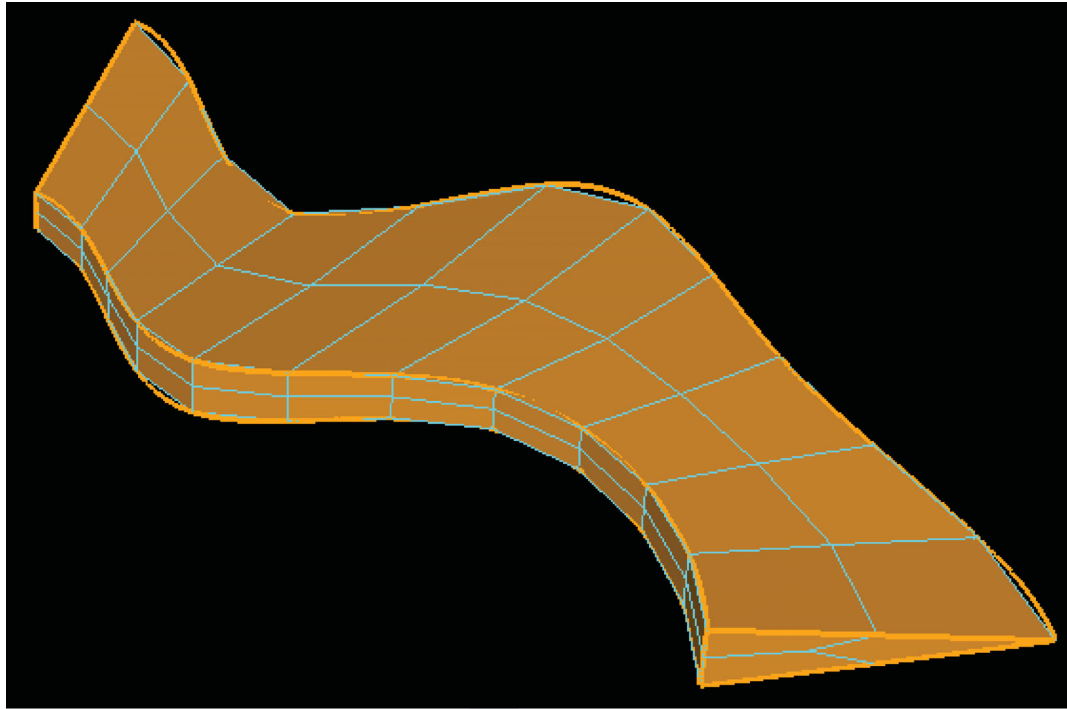


Figure 4. Detail of a subdomain as part of the inner outline of the Caracas basin. The volume was meshed from the projection of the 2-D triangular front mesh.

The creation of the mesh is not performed by RegSEM; it is done externally using the CUBIT mesh generation tool kit (<http://cubit.sandia.gov>). Input mesh files for RegSEM are then created from export CUBIT mesh files. Considering the limited choice of commercial and non-commercial codes dealing with hexahedra compared to the case of tetrahedra, we think that this tool kit offers the best alternative. However, the CUBIT mesh generation for 3-D complex structures is not totally automatic and requires many steps and user interventions. Automatic procedures have been developed by Casarotti *et al.* (2008b) to generate meshes according to different strategies. Embedded in a parallel Message Passing Interface (MPI) environment, they can fast create simple 3-D unstructured meshes. However, in the case of a complex basin, the outcrop is generally not honoured when this one exhibits too many variations. In this case, a robust fully 3-D unstructured algorithm for hexahedra is still not available.

The meshing technique we use in this paper has been successfully applied to the valley of Grenoble by Stupazzini *et al.* (2009). It consists in building a conform mesh of the model from separately meshed subdomains, to better control the size and shape of the elements. This work requires a pre-process with a CAO software which can handle NURBS (Non-Uniform Rational B-Splines) geometries. From the digital terrain model data, this software creates NURBS curves which define the interfaces of the 3-D model, the topography, the basin basement and the numerical boundaries. The total volume is then partitioned into subdomains which are also defined by a group of curves and exported in an Initial Graphics Exchange Specification format readable by CUBIT. These subdomains are independently reconstructed by CUBIT which assembles them in a conform way to form the whole domain. Each subdomain can then be individually meshed. A more detailed description of this meshing process can be found in Delavaud (2007) and in the manual of the code. The advantage of this procedure holds in the possibility of associating to each subdomain different element sizes

and types of meshing while the mesh remains conform. The number of subdomains can be substantial, depending on the complexity of the structures which have to be meshed. In the presence of a basin, this partitioning is mainly controlled by its outline in surface which needs to be isolated to correctly mesh the bend and the shape of the edge. Fig. 4 shows one of the subdomains which defines the inner outline belt of the Caracas basin. One of the triangular surface at the ends of the subdomain is first meshed with a meshing scheme called 'triprimitive' which applies to three side surfaces. The size of the elements is inherited from an interval size assigned to the edges. The total volume is then meshed by a projection (sweep) of the two dimension mesh along the edges towards the opposite triangular surface. As one can see, the element at the edge is particularly deformed and introduces a very small minimum distance between the GLL nodes. The variation of the free surface topography also influences the cutting into subdomains needed to ensure a homogeneous mesh in depth. The meshing strategy remains the same as the one described for the outline of the basin: the free surface is first meshed with an unstructured meshing scheme, then a sweep is applied in the vertical direction. Therefore, the mesh is unstructured only in two directions (in depth for the interior outline belt of the basin, and horizontally for the other parts). As an example, the case of a simple hemispherical basin is presented in Fig. 5.

Meshes which honour geological discontinuities might present highly deformed elements, especially at the edges. To assess the quality of a mesh and identify such elements, CUBIT offers different metrics. Skewness, distortion or shear associated with the Jacobian of each element are possible quality measurements. In the case of the SEM, the effect of deformed elements on precision and efficiency is fortunately limited by the high degree of the method (Oliveira & Seriani 2011).

A recent study by Pelties *et al.* (2010) compares the method that interpolates the outline of a basin by Lagrange polynomials (e.g. Komatitsch *et al.* 2004; Casarotti *et al.* 2008b; Lee *et al.* 2008)

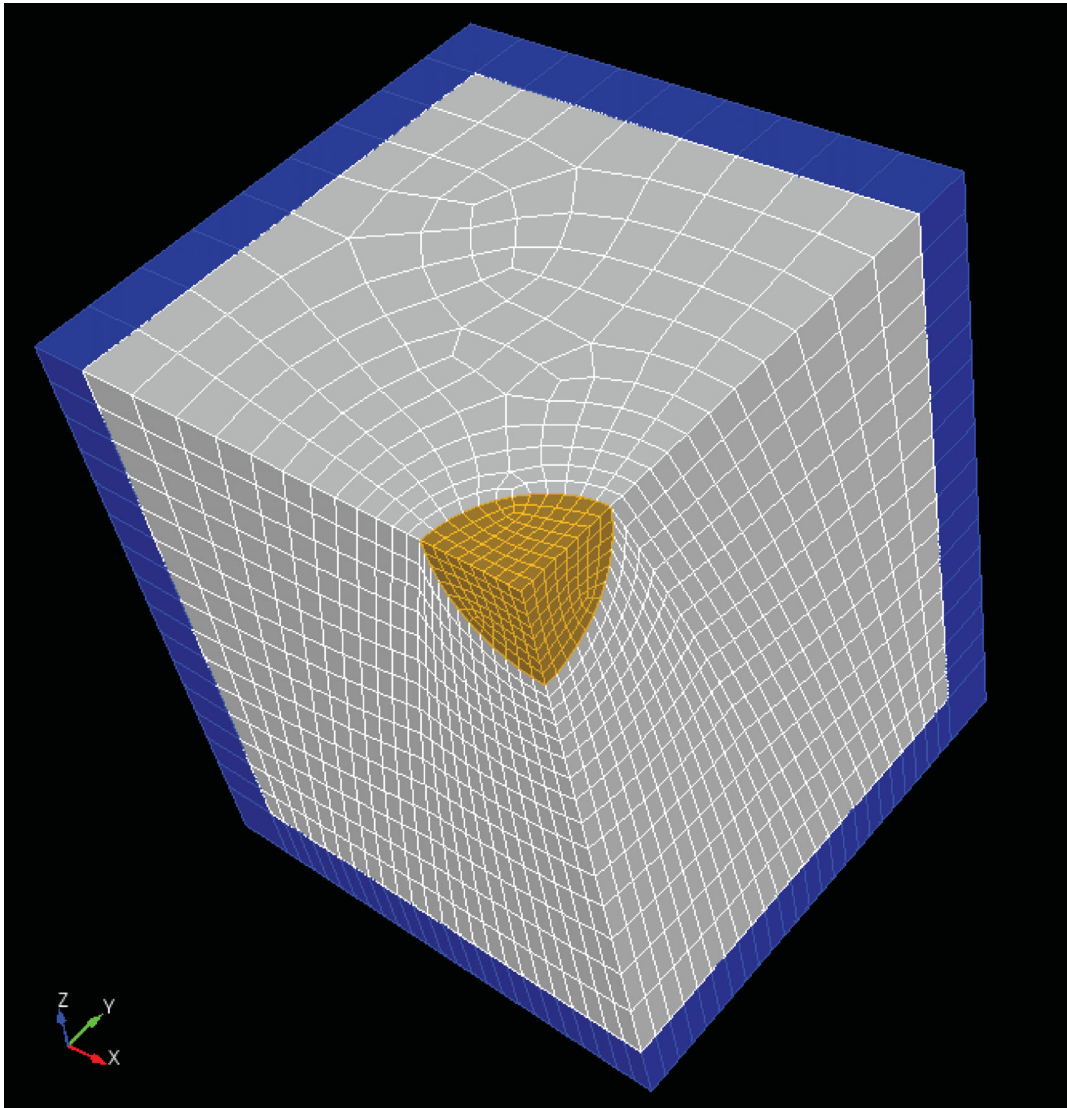


Figure 5. Mesh of a quarter of a half space containing a hemispherical basin. The model is cut into four quarters independently meshed using CUBIT and then reassembled. The basin (yellow) is meshed first, then the free surface, and finally the volume. The blue elements correspond to the PML.

with a meshing that fully honours it (e.g. Stupazzini *et al.* 2009). From tests performed at different frequencies in different velocity contrasts, this study provides empirical rules to ensure the reliability of Lagrange interpolation. In the case of 3-D simulations in the Grenoble valley, Casarotti *et al.* (2008a) observed differences in amplitude and phase of the order of 15 per cent between the two methods above 1 Hz. At lower frequencies, the detail of the basin shape has less, or even no, influence. Although it is not the scope of this paper to present a comprehensive study about the differences between honouring and not honouring interfaces, we briefly show some comparisons based on a 2-D profile of the Caracas basin. The effects of three meshing strategies on spectral ratios are presented. Spectral ratios are amplification factors with respect to the spectrum of the incident plane wave. We compute them for frequencies up to 5 Hz and for each receiver along the free surface. In Fig. 6, on the left panel, the ratios are shown for a mesh that respects the corners of the basin. These corners are critical because they generate diffracted surface waves responsible for large amplifications. Bi-dimensional effects are characterized by amplifications

at frequencies higher than the fundamental frequency (0.5 Hz), especially above the thicker part of the basin at 1.5 Hz, 2.5 Hz and 3.5 Hz. We also consider a regular mesh in which the discontinuity is interpolated at the GLL nodes (Fig. 6, mid panel). In this case, we show the ratio in logarithmic scale of the spectra recorded at the free surface of the regular mesh and those recorded at the free surface of the honouring mesh. Up to 2 Hz, the spectra are almost similar. At higher frequencies, large discrepancies appear, with ratio of the order of 2. A third meshing strategy consists in respecting the interface until a depth of 60 m, about 50 per cent of the minimum propagated wavelength, and to interpolate the discontinuity in the elements at the edges of the basin (Fig. 6, right panel). The spectra obtained in this case are very similar to the fully honouring case up to 2.5 Hz. Above this frequency, large but localized discrepancies appear, also of the order of 2. This example put in evidence that the entire rock-sediment interface, especially the edge of the basin, should be respected for high frequency simulations (>2.5 Hz). For shorter frequencies, the direct discretization of the interface at the GLL nodes down to a reasonable depth seems appropriate.

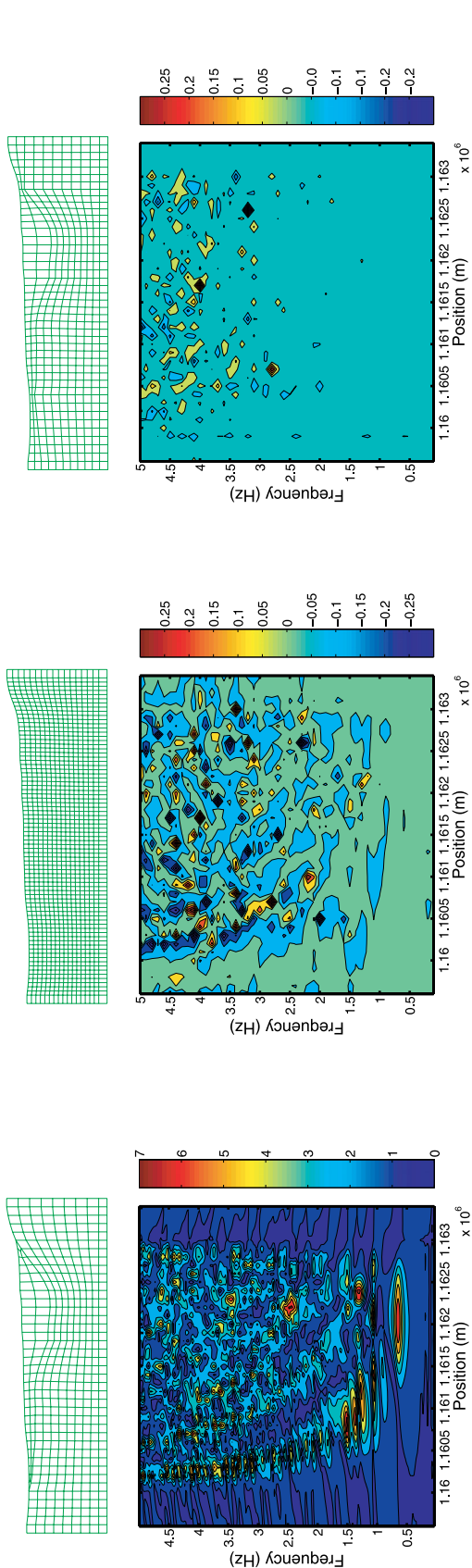


Figure 6. Effects of three different meshing strategies for the Caracas basin. Left panel: the sediment-rock interface is fully honoured. The spectral ratios computed at the free surface are shown. Mid panel: the interface is interpolated by a thin GLL grid. The ratio (in logarithmic scale) of the spectra recorded at the surface and the spectra computed in the honouring mesh is plotted. Beyond 2 Hz, significant differences appear. Right panel: the whole interface is honoured except its edges. In this case, comparing the spectra with those computed in the fully honouring mesh shows large but localized discrepancies at high frequency.

Advantages and drawbacks of these strategies in terms of precision and implementation complexity could be discussed in more detail based on the analysis conducted by Maday & Ronquist (1990).

2.3 Introduction of the elastic parameters

As RegSEM is able to handle a large set of meshes, it has to be versatile in introducing elastic models as well. In the continental version of the code, this is achieved thanks to a Fortran module that the user can change by himself and which is conceived to provide the elastic parameters at any location (radius, latitude and longitude) in the Earth. Both radial and azimuthal anisotropies are implemented. Moreover, the anelastic structure can be taken into account using a series of standard linear solids, as suggested by Emmerich & Korn (1987) and Komatitsch & Tromp (1999). The unstructured version of the code is limited to lossless isotropic media.

Rotation and self-gravitation, which involve non-neglectable effects at the global scale and very long periods (Komatitsch & Tromp 2002b; Chaljub & Valette 2004), are not included. Propagation in fluid has not been implemented, so the waves either from the outer core or from the oceans cannot be simulated. Nevertheless, following Komatitsch & Tromp (2002b), the mass of the oceans can be taken into account when a bathymetry is used at the surface of the Earth.

2.4 Absorbing boundary conditions

To avoid artificial reflections at the border of the chunk, it is necessary to implement efficient absorbing boundary conditions. RegSEM uses the velocity-stress formulation of the so-called Perfectly Matched Layers (hereafter PML; see Festa & Vilotte 2005). This formulation requires an unphysical splitting of the field variables along the directions of normal and parallel derivatives with respect to the interface PML volume. This means that in practice, the splitting directions have to be known at every GLL node belonging to the PML, which is not obvious when working with Cartesian coordinates in a deformed layer (such as the lowermost layer of the chunks presented in Figs 1–3). Therefore, we here make an assumption: for all the GLL nodes of a given element, the splitting directions defined at the centre of the element are used. The effect of such an assumption on the stability and accuracy of the PML is not clear. Furthermore, our PML are isotropic, so spurious reflections can be created when considering anisotropic media. Examples in the following parts of the paper will show that our implementation of the PML however provides satisfactory results.

Note that the SEM does not require to use the same polynomial order for all directions. Our code takes advantage of this flexibility: in the PML, it is possible to use a different polynomial order in the damping directions. In the following, we will always use $N = 8$ in these directions.

2.5 Sources

Force-vector and moment-tensor point-sources can be placed at any location in the chunk. Four different functions are provided to describe the time signal at the sources: a Gaussian, its derivative (i.e. a Ricker wavelet), its antiderivative (i.e. an error function) or the Fourier transform of a frequency band defined by a cosine taper.

More interestingly, in particular for basin response modelling, an incident wavefield can be introduced in the unstructured version of RegSEM. The method developed for that purpose is based

on a decomposition technique and exploits the natural presence of the traction in the SEM formulation, that is the \mathbb{I}^{trac} term in eq. 1. The wavefield is introduced on an interface in the domain, for example, the sediment-rock interface, by its action on the traction forces (Neumann condition). Similar ideas had been introduced by Bielak & Christiano (1984) in the context of finite elements for the problem of soil–structure interaction. The main interest of this method consists in avoiding the propagation of the incident field, which is known analytically or numerically, as long as it has not reached any discontinuity with which it will interact by reflection, transmission or diffraction. This type of introduction is compatible with any boundary conditions, including PML. Moreover, diffraction problems for non-vertical incidences are prevented. Finally, the computational domain does not have to be large to hold the incident wave. We refer to the manual of the code for more details about the implementation of this method.

2.6 Parallel implementation

The SEM can be easily implemented on distributed memory architectures. Given a number n of CPUs, the computational domain has to be divided into n subdomains. To do so, we use the software library METIS (<http://glaros.dtc.umn.edu/gkhome/views/metis>) that ensures an efficient partitioning which minimizes the communications between the subdomains. These communications occur at every time step of the time-marching scheme. To perform them, we use the MPI. Fig. 1 shows an example of a chunk partitioned by METIS. Fig. 7 shows the good scalability of our parallel implementation in the case of the experiment described in Section 3.1.

3 VALIDATION

In this part, a series of numerical experiments are carried out to validate our code. We start with the simple case of a homogeneous medium, then we consider a layered medium (PREM) and we finally study the case of a hemispherical basin. For all these experiments, a reference solution is known.

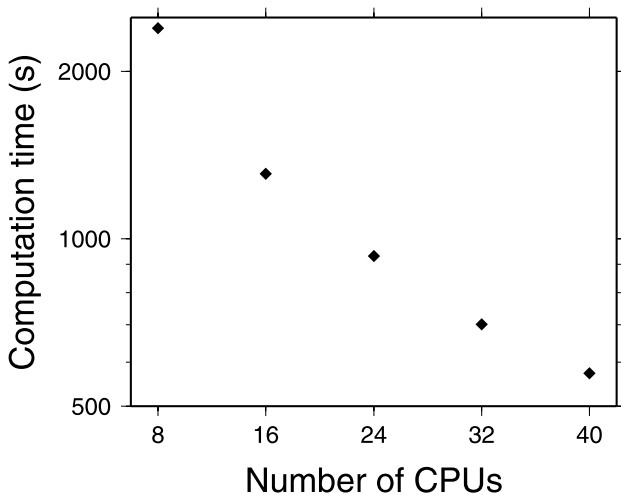


Figure 7. Scalability of the RegSEM parallel implementation. The simulation used to do this test is the one described in Section 3.1. The points are almost aligned, showing that the computation time goes like the inverse of the number of CPUs.

3.1 Simulation in a homogeneous medium

We first consider a homogeneous medium in spherical geometry. In such a context, the normal mode summation technique provides a quasi-analytical solution (Capdeville 2000). The P -wave speed, S -wave speed and density of the medium are $\alpha = 8 \text{ km s}^{-1}$, $\beta = 5 \text{ km s}^{-1}$ and $\rho = 3000 \text{ kg m}^{-3}$, respectively. The chunk used in the SEM simulation is shown in Fig. 1. The elements are 1.3° large. This enables to use a 50 mHz cut-off frequency with a polynomial order $N = 8$. The source is an explosion located at 10 km depth. The receiver is on the free surface. The epicentral distance is 20° . On three Intel Xeon 2.5 GHz quad-core dual-processor nodes (i.e. 24 CPUs), it takes about 16 min to compute 1200 s. In Fig. 7, the computation time for other numbers of CPUs is shown. The scalability of our parallel implementation is seen to be good.

Fig. 8 shows the comparison between the SEM result and the normal mode solution. All the seismograms are normalized with respect to the amplitude of the vertical component obtained with the SEM, so the relative amplitudes are preserved. Because the

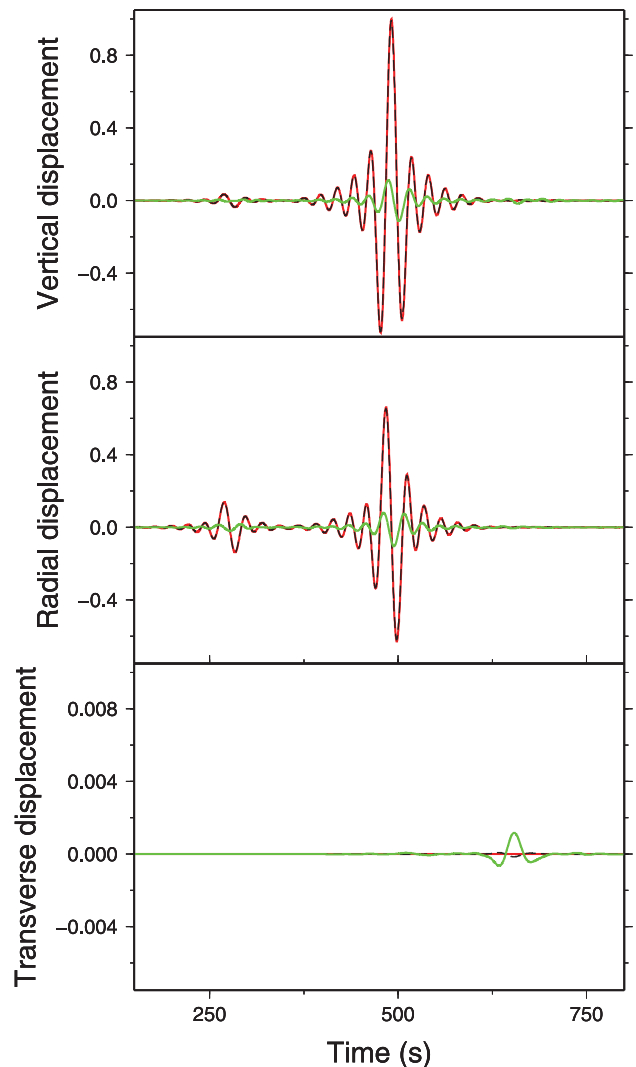


Figure 8. Comparison of the SEM solution (dashed black) with the normal mode solution (red) obtained in a homogeneous medium ($\alpha = 8 \text{ km s}^{-1}$, $\beta = 5 \text{ km s}^{-1}$, $\rho = 3000 \text{ kg m}^{-3}$). The epicentral distance is 20° . The residual multiplied by 10 is plotted in green.

10 *P. Cupillard et al.*

source is an explosion, there is no SH energy and the waves only lie on the vertical and radial components. On these components, the traces obtained from the two methods are indistinguishable. The plot of the residual multiplied by 10 shows that the maximum error is around 2 per cent. This small error is essentially due to the finite difference time-scheme whose order is only 2. On the transverse component (whose y -axis is 100 times larger than the other components), the SEM solution is not exactly zero: around 650 s, the residual multiplied by 10 reveals a spurious reflection coming from the bottom of the chunk. This signal is extremely small and is hardly seen on the radial and vertical components, meaning that our PML are good.

3.2 Simulation in PREM

We use PREM to perform a second validation test. The thin fluid layer which lies on the top of this model is replaced by the underlying solid. The chunk used in the SEM simulation is shown in Fig. 2. The elements are 0.44° large. This enables to use a 50 mHz cut-off frequency with a polynomial order $N = 4$. The source-receiver configuration is the same as in the previous test.

3.2.1 With no attenuation

We first do not take into account the anelastic structure of PREM. In this case, the normal mode summation gives a good reference solution. Fig. 9 shows a comparison between this solution and our SEM result. Again, the two waveforms are indistinguishable on both vertical and radial components. Nevertheless, on the radial component, the residual shows significant amplitudes after the main phases, which was not the case in the test performed in the homogeneous medium. This is explained by two reasons. First, our chunk is cut at depth, so the phases reflected at the core–mantle boundary, such as the *PcS* and *ScS* phases, are missing. Second, a spurious reflection from the PML on the vertical sides of the chunk is detected. This signal appears here because PREM is anisotropic. The magnitude of this spurious reflection is similar to the error due to the finite difference time-scheme, so even though they are not perfect, our PML are satisfactory. On the transverse component, a tiny reflection coming from the bottom of the chunk is observed, as it was the case in the homogeneous medium. This reflection arrives earlier here because PREM velocities increase with depth.

3.2.2 With attenuation

When introducing the anelastic structure, the comparison between the SEM and the normal mode solutions (Fig. 10) does not change a lot. The main difference with the lossless case is a large residual value (5–10 per cent) for the body waves and the beginning part of surface wavetrain. This large value is probably due to the fact that the normal mode result in an attenuating medium relies on a first-order approximation. Nevertheless, we can conclude that the attenuation is well taken into account in RegSEM because the SEM and normal mode summation waveforms are very similar to each other and are both very different from the lossless case.

3.3 Plane wave on a hemispherical basin

To validate our code with an unstructured mesh, we now consider the half space containing the hemispherical basin of radius R (Fig. 5). For such a simple shape, the boundary element method developed

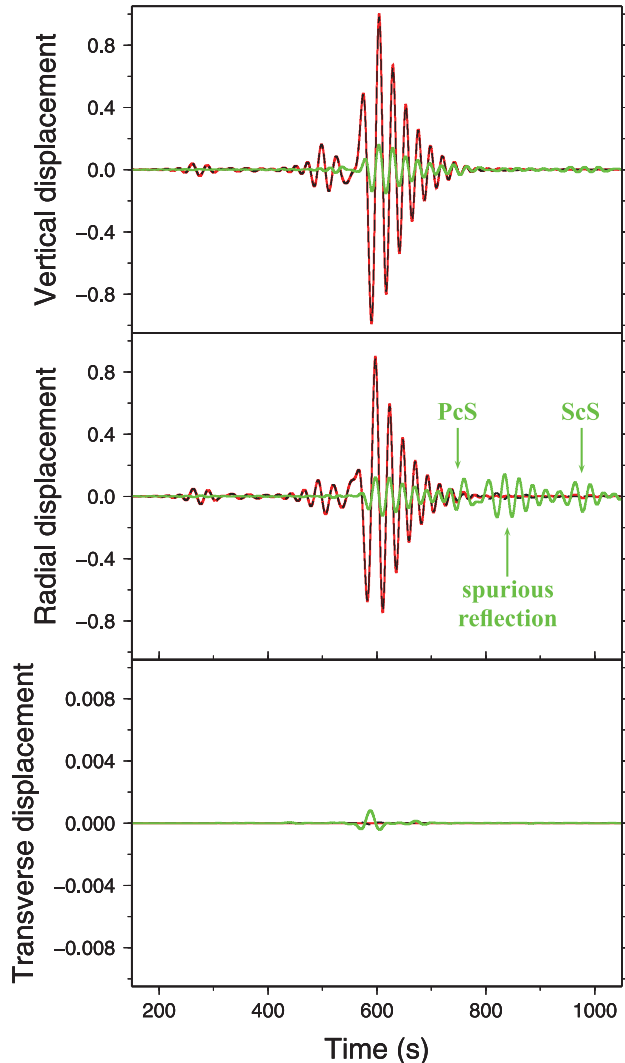


Figure 9. Comparison of the SEM solution (dashed black) with the normal mode solution (red) obtained in PREM with no attenuation. The epicentral distance is 20° . The residual multiplied by 10 is plotted in green.

by Sánchez-Sesma (1983) provides a semi-analytical solution of the diffracted wavefield in the frequency domain. The parameters of the material are summarized in Table 1. They respect the conditions on the shear modulus μ , density ρ and Poisson coefficient ν described by Sánchez-Sesma (1983): $\mu^R/\mu^E = 0.3$, $\rho^R/\rho^E = 0.6$, $\nu^R = 0.3$ and $\nu^E = 0.25$, where the exponents R and E correspond to the basin and the rest of the half-space, respectively. The model is excited by a plane P -wave with a vertical incidence. The results are presented for a normalized frequency $\eta_P = \frac{2R}{\Lambda_P}$, where Λ_P is the wavelength of the incident plane P wave. This normalized frequency determines the central frequency f_P of the Ricker pulse which defines the time function of the incident plane wave. For the normalized frequency $\eta_P = 0.5$ considered by Sánchez-Sesma (1983), $f_P = 5.76$ Hz.

In the SEM simulation, a polynomial order $N = 4$ is chosen in the elastic medium. The signal is recorded at the free surface, along the line $[Ox]$, where O is the centre of the basin and x a horizontal axis. We are interested in the transfer function, which is the ratio between the spectrum of the signal recorded at each receiver along this line and the spectrum of the incident plane wave. Fig. 11 shows the

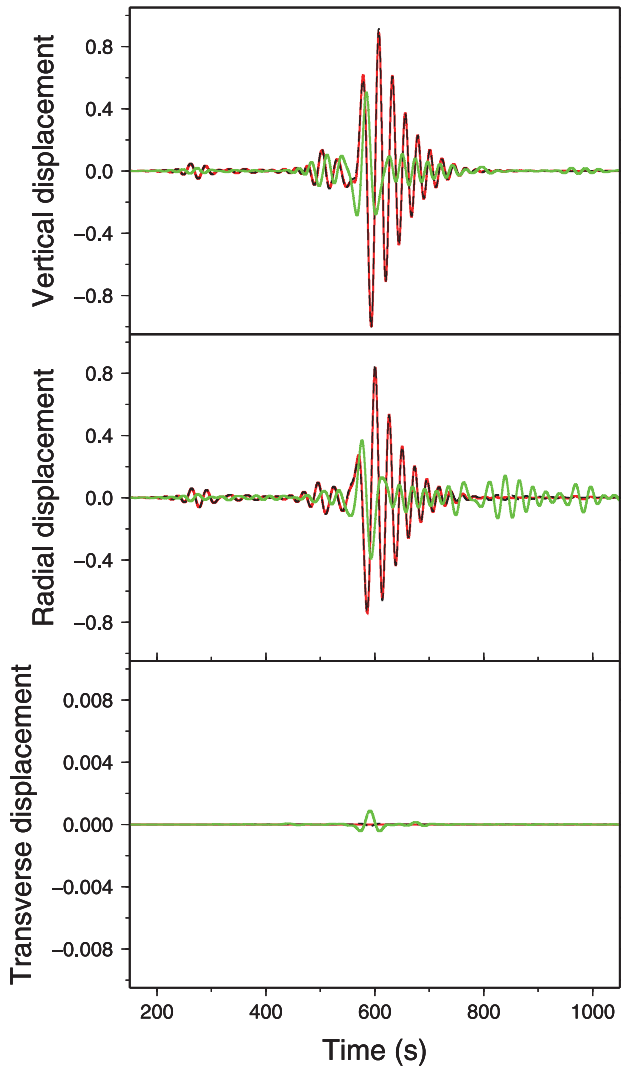


Figure 10. Comparison of the SEM solution (dashed black) with the normal mode solution (red) obtained in PREM with attenuation. The epicentral distance is 20° . The residual multiplied by 10 is plotted in green.

transfer functions, for the SEM and the semi-analytical solution, of the vertical and horizontal components of the displacement at the normalized frequency $\eta_P = 0.5$ as a function of the normalized distance x/R at $t = 5$ s. These transfer functions are also represented at the normalized frequency $\eta_P = 0.7$ which corresponds to the frequency $f_P = 8$ Hz. As a result, the two methods provides the same behaviour. Small differences of 2 per cent can be observed, which is comparable to the numerical dispersion of the SEM and the magnitude of the PML reflections.

A variability of the behaviour according to the frequency is observed, knowing that the 1-D resonance frequency of the basin

equals $\alpha^R/4R = 4.4$ Hz. At 5.76 Hz, the curve is close to the 1-D case, with a maximum amplification at the centre of the basin, about 175 per cent of the amplitude obtained in the case of a homogeneous medium. The amplitude decreases when approaching the basin border and then keeps the state of a homogeneous medium. The horizontal component, on the contrary, exhibits a null amplitude at the centre of the basin, due to the symmetry, and reaches its maximum at $x/R = 0.5$. Then it converges to a stable amplitude. At 8 Hz, an important wave conversion appears around $x/R = 0.35$ where the horizontal component reaches its maximum and the vertical component its minimum. Outside the sphere, the two components tend to the same stable state as at 5.76 Hz. The maximum amplification at the centre of the basin is reached for a normalized frequency $\eta_P = 0.61$ ($f_P = 7$ Hz), 1.6 times the 1-D resonance frequency. These results underline the influence of the basin effects, even in a simple and symmetric case.

4 SIMULATIONS IN 3-D MODELS OF EUROPE

In this part, we point out the influence of the shallow structures (crust and uppermost mantle) on a regional wavefield. RegSEM's flexibility is used to generate structured meshes and simulate an event in two 3-D models of the European region. The models are Crust2.0 (Bassin *et al.* 2000) over the 1-D model PREM, and the crustal and upper mantle model CUB (Shapiro & Ritzwoller 2002). The corresponding synthetic seismograms are compared with the normal mode solution in PREM, and above all with the real data seismograms.

4.1 Models

Crust2.0 (Bassin *et al.* 2000) was obtained by compiling seismic studies and tectonic, geological settings. The model is defined as cells on a $2 \times 2^\circ$ geographical grid. In each cell, we average over the different layers (there are up to seven layers, including ice or water, sedimentary layers and layers in the bedrock) to get homogeneous seismic velocities and density. Then, a 2-D Gaussian filter is applied horizontally to set up the seismic parameters at every GLL nodes. Below the 3-D Moho, PREM is used to represent the deeper Earth.

CUB (Shapiro & Ritzwoller 2002) is a radially anisotropic model obtained by surface wave tomography. The model is derived from the inversion of group and phase dispersion data of the fundamental mode of both Rayleigh and Love waves in the 16–200 s period range. It is defined on a $2 \times 2^\circ$ geographical grid. It is characterized by a vertical block parametrization for the crust, a 3-D Moho and a spline parametrization for the upper mantle. We actually use a smoothed version of the CUB crust: the fluid and low-velocity ($\beta \leq 2.4 \text{ km s}^{-1}$) shallow layers are replaced by the underlying material, and a vertical smoothing is applied within the crust using the intrinsic interpolation law of the SEM. Moreover, because the resolution of CUB is poor deeper than 250 km, a linear transition

Table 1. Parameters of the half-space containing the hemispherical basin used in Section 3.3. α , β and ρ correspond to the P -wave velocity, S -wave velocity and density, respectively. Exponent R refers to the basin and exponent E refers to the rest of the half-space. η_P is the normalized frequency.

α^E	β^E	ρ^E	α^R	β^R	ρ^R	R	η_P
1730 m s^{-1}	1000 m s^{-1}	2000 kg m^{-3}	1320 m s^{-1}	710 m s^{-1}	1200 kg m^{-3}	75 m	0.5

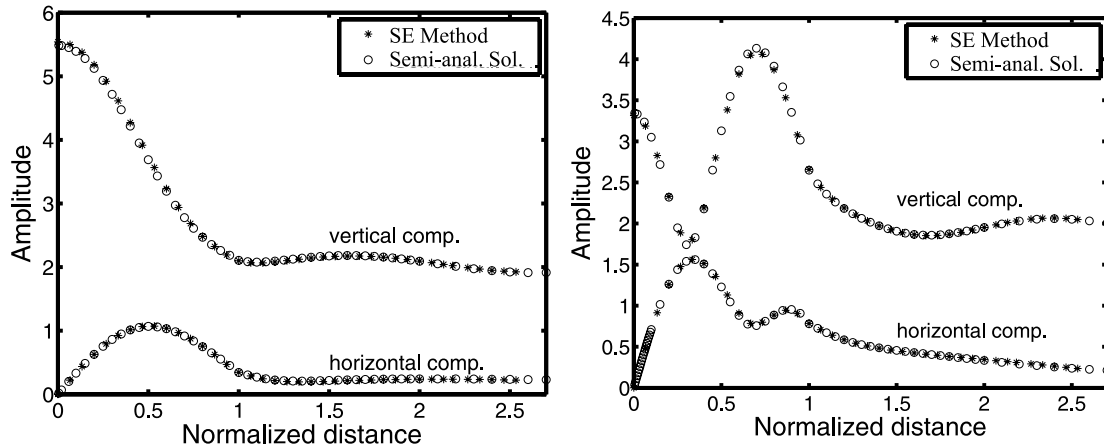
12 *P. Cupillard et al.*

Figure 11. Transfer functions of the vertical and horizontal components of the displacement as a function of the normalized distance x/R at $t = 5$ s. We compare the SEM (stars) and a semi-analytical solution (circles) based on an expansion of Bessel functions (Sánchez-Sesma 1983) in the case of a hemispherical basin, for two frequencies: 5.76 Hz ($\eta_P = 0.5$) on the left and 8 Hz ($\eta_P = 0.7$) on the right.

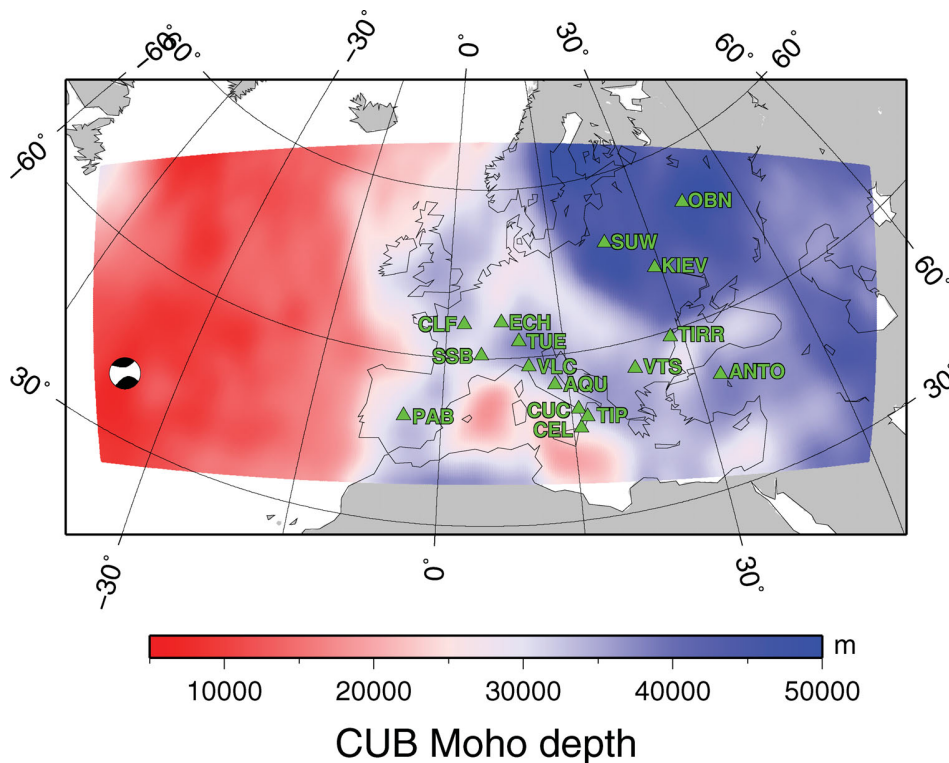


Figure 12. Map of the Moho of model CUB (Shapiro & Ritzwoller 2002) in the Atlantic–European region. The source–receiver configuration used in Section 4.3 is also shown.

towards PREM is imposed down to the transition zone. Again, we use a Gaussian filter in the horizontal directions. The resulting model in the Atlantic–European region is shown in Fig. 3. A map of the Moho is shown in Fig. 12.

4.2 Data

The earthquake we consider has been recorded at 16 receivers on the continental Europe. All data come from broad-band stations operated by the GSN, GEOSCOPE, GEOFON and MEDNET networks. Both vertical and horizontal components on the BH channels are

selected. To be compared to the data, the synthetic seismograms are convolved with the instrumental transfer function. This avoids the deconvolution of the response on the data.

The comparison between data and synthetics is performed in four period bands: 100–200 s, 50–100 s, 30–50 s and 20–30 s. To obtain a quantitative estimate of the misfit between data and synthetics, we apply a systematic cross-correlation between the two using a group velocity criterion. Two parameters of the cross-correlograms are extracted: the delay δt of the main peak and the amplitude ratio R_A between the main peak and the autocorrelation of the data. A perfect matching therefore yields $\delta t = 0$ s and $R_A = 1$. Our comparison is

mainly based on the vertical component. The signal-to-noise ratio in the radial and transverse components is often poor, so a more severe data selection would actually be necessary to include more horizontal components in our study. Note that the seismograms in the 100–200 s period band are not presented because they show the same behaviour as those in the 50–100 s period band. Furthermore, in the shortest period band (20–30 s), waveforms are complex and some correlation parameters therefore do not make sense.

4.3 Results

The event we investigate occurred along the Mid-Atlantic Ridge on 2010 May 25, at latitude 35.41°N and longitude 35.93°W. Its depth has been estimated to 10 km and its USGS CMT solution provides a moment magnitude of 6.3. The regional chunk used for the SEM simulations of this event is shown in Fig. 3. A map view of the chunk can be found in Fig. 12, representing also the source–receiver configuration. The seismograms obtained at four relevant stations are shown in Fig. 13. The correlation parameters for these stations

and the mean of the parameters over all the stations are presented in Table 2.

In the long-period bands (100–200 s, 50–100 s), the synthetic waveforms are similar to the data with a correct amplitude. However, at every station, a systematic positive delay is observed between the data and the synthetics computed in PREM or in Crust2.0. This delay increases with the epicentral distance. CUB shows a much better fit: for example, at station KIEV, models PREM and Crust 2.0 present a large delay ($\delta t = 20$ s and $\delta t = 34$ s, respectively) whereas CUB shows a short delay ($\delta t = -6$ s). In the medium band (30–50 s), the waveforms are still similar and the amplitude is still correct. The CUB model improves the fitting to the data in comparison with what PREM and Crust2.0 do. However, when increasing the epicentral distance, the trends of the last two models differ. On stations SSB and ECH (short paths), PREM and Crust2.0 both show the same large positive delay. On stations AQU and KIEV (longer paths), the Crust2.0 model presents an increasing delay ($\delta t = 95$ s) while PREM model presents a decreasing delay ($\delta t = 7.8$ s). In the short-period band (20–30 s), the waveforms significantly differ,

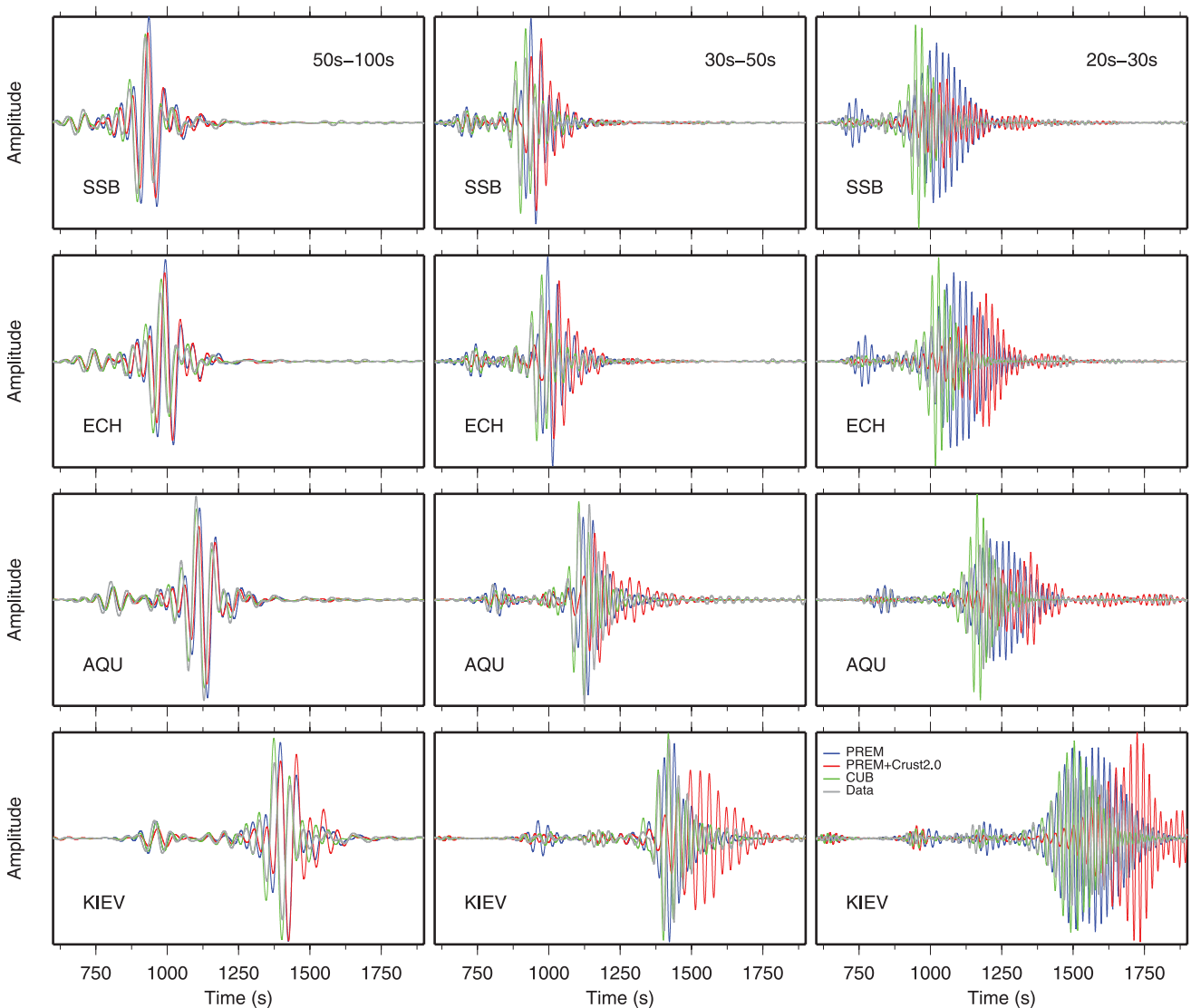


Figure 13. Waveforms induced by the mid-Atlantic ridge earthquake at four stations in Europe for three different period bands. We compare the real data (grey) with the waveforms obtained in three different Earth models (PREM in blue, PREM+Crust2.0 in red and CUB in green).

Table 2. Parameters (time delay δt in s and amplitude ratio R_A) of cross-correlations between data and waveforms obtained from a mid-Atlantic ridge earthquake in three different Earth models (PREM, Crust2.0 and CUB) at four different stations for three different period bands. The average μ^{all} of the two parameters over all the stations is also shown.

50–100 s	Δ	R_A [PREM]	δt [PREM]	R_A [Crust2.0]	δt [Crust2.0]	R_A [CUB]	δt [CUB]
SSB	3563.69	1.2104	19.9	1.0324	18.5	1.0454	−1.4
ECH	3791.79	1.4728	21.0	1.3104	23.3	1.2127	−2.2
AQU	4280.63	0.8988	19.1	0.7514	20.7	0.8565	1.9
KIEV	5378.05	1.1814	20.0	1.1647	34.0	1.2336	−6.0
μ^{all}	4498.27	0.4574	14.7	0.9828	19.7	1.2894	−6.0
30–50 s	Δ	R_A [PREM]	δt [PREM]	R_A [Crust2.0]	δt [Crust2.0]	R_A [CUB]	δt [CUB]
SSB	3563.69	1.4931	33.5	1.3292	51.8	1.3599	0.4
ECH	3791.79	1.5459	37.0	1.2138	62.9	1.2559	1.7
AQU	4280.63	0.8002	7.4	0.6084	38.8	0.8187	−18.0
KIEV	5378.05	0.9071	7.8	0.8694	94.7	0.9530	−6.5
μ^{all}	4498.27	0.6797	12.5	1.0752	60.0	1.4994	−9.3
20–30 s	Δ	R_A [PREM]	δt [PREM]	R_A [Crust2.0]	δt [Crust2.0]	R_A [CUB]	δt [CUB]
SSB	3563.69	2.0489	63.1	0.9551	96.2	1.8675	0.3
ECH	3791.79	2.4523	57.0	1.5337	135.1	2.1946	−3.7
AQU	4280.63	1.0182	30.7	0.6072	125.8	1.2822	−28.5
KIEV	5378.05	1.3895	−3.8	1.1147	165.1	1.0200	−27.6
μ^{all}	4498.27	0.9749	12.8	0.9433	124.6	1.8047	−33.2

in particular when increasing the epicentral distance. The Crust2.0 model always shows the worst waveforms and sometimes have non-sense delays ($\delta t = 165$ s at KIEV) due to the strong coda wavetrain. At short epicentral distance (SSB for example), the CUB model presents very good fits ($\delta t = 0.3$ s) and is better than PREM ($\delta t = 63$ s). At longer epicentral distance (KIEV for example), situation is reversed: PREM is better ($\delta t = -3.8$ s vs. $\delta t = -28$ s).

In the short-period range, the effect of the crust on delay times is predominant. For oceanic paths, PREM has a too thick crust (slow velocity compared to fast velocity in the underlying lithosphere) which tends to induce positive delay times. The situation is reversed in continents, which are characterized by a thick crust and fast lithospheric velocities. Consequently, when only the effect of the crust is taken into account in the numerical simulations, the delay time tends to increase compared to what PREM does. That is exactly what we observe: for Crust2.0, the fit to the data is always poor when the path is dominantly oceanic or continental. The success of the CUB model in almost all numerical simulations is due to the fact that both crust and upper mantle 3-D structures are taken into account. For some specific paths, the right balance in the mixing of oceanic and continental paths can provide very good fits with PREM.

4.4 Conclusion

The comparison of seismograms computed in different Earth models clearly shows that the effect of the crust is large and non-linear. For pure oceanic paths or pure continental paths, the account of the crust tends to increase the residual delay times computed in PREM. This effect is well known (Montagner & Tanimoto 1991). When incorporating both crust and upper mantle 3-D structures, the time residuals are significantly improved, particularly at periods larger than 30 s. At short periods, the strong scattering effect of the crust gives rise to long coda waves which often makes our simple cross-correlation technique inefficient.

Our brief analysis puts in evidence how important the account for the shallow structures is to correctly model the seismic wave propagation at the regional scale. Previous studies (Komatitsch *et al.*

2002; Bozdag & Trampert 2008) used the SEM at the global scale and already noted the great influence of the shallow parts, but they did not honour the topography of the Moho. Because the computation cost is less important at the regional scale and because of the ability of RegSEM to take into account any Moho topography, this is overcome here. To exhibit the importance of honouring the Moho, we compare two seismograms computed at the same station (ANTO) in two different meshes of the CUB model (Fig. 14). In the first mesh, the Moho is fully honoured. In the second mesh, the elements at the top of the chunk are all 50 km thick, so the Moho is not honoured at all. The differences between the two traces are clear: a shift of half a period appears in the high frequency part of the signals (20–30 s), in addition with large discrepancies in amplitude (up to a factor of 2). At the global scale, similar differences are observed by Capdeville & Marigo (2008). These results illustrate the influence of the Moho on surface wavetrains and point out how crucial meshing its topography is to get accurate seismograms.

Thanks to its flexibility, RegSEM is a suitable tool to assess the quality of regional tomographic models, as done at the global scale by Qin *et al.* (2009) and Bozdag & Trampert (2010). Nevertheless, the simulations presented here do not give precise conclusions on the quality of the CUB and Crust2.0 models. First, Crust2.0 is used with PREM although they might not be compatible as these two models have been obtained from two different data set and techniques. Second, the brutal smoothing applied to Crust2.0 and the CUB crust probably changes the effective elastic properties of the two models. A more rigorous smoothing technique, such as the one suggested by Capdeville *et al.* (2010) and Guillot *et al.* (2010), should be used to precisely assess the models.

5 DISCUSSION

We showed that RegSEM is an efficient tool to compute seismic wavefields in geological media with possible complex geometries. Two particular cases are preferentially investigated. At the local scale, unstructured meshes of sedimentary basins (externally generated by CUBIT following the procedure suggested by Stupazzini

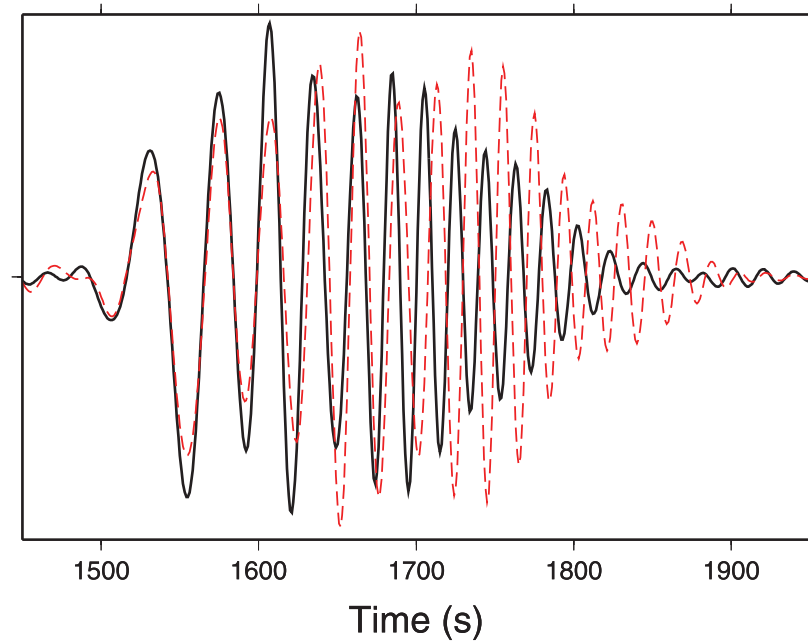


Figure 14. Rayleigh wave (vertical component) computed at station ANTO from two different simulations of a mid-Atlantic ridge earthquake in model CUB. In one simulation, the Moho is fully honoured (black line). In the other simulation, the Moho is interpolated by Lagrange polynomials of degree 4 defined in 50 km thick elements (red dashed line). The period range of the signals is 20–200 s.

et al. 2009) can be handled to study site effects. At longer scales (up to 90°), the influence on wave propagation of crustal and mantle 3-D structures with a possible Moho topography can be simulated using structured meshes generated by an internal routine. In both cases, the shallow structures yield important non-linear effects, even at wavelengths larger than the size of the heterogeneities. We showed that meshing these shallow structures is a crucial issue for a proper modelling of realistic wavefields.

In this work, the use of the SEM is limited to forward modelling. As mentioned in Section 1, the method can also be used to solve the inverse problem. A now classical and popular technique to achieve this goal consists in computing sensitivity kernels using the adjoint method (Tarantola 1984; Tromp *et al.* 2005; Fichtner *et al.* 2006a,b). This can be done with RegSEM, which provides kernels by simultaneously computing the adjoint wavefield and reconstructing the regular wavefield from time-reversed seismograms recorded at the boundaries (Gauthier *et al.* 1986). This process prevents from storing the whole regular wavefield, but it is rigorously valid in non-dissipative media only. Nevertheless, an artificial amplification can be introduced when reconstructing the regular wavefield in dissipative media (Tarantola 1988). Performing tomographic inversions using RegSEM is kept for future papers.

Another interesting application of our code is the computation of synthetic microseismic noise correlations. These data were introduced in seismology by Shapiro & Campillo (2004). Because most of the microseismic energy is contained in the 5–20 s period range and propagates as surface waves (Longuet-Higgins 1950), noise correlations are very sensitive to the shallow structure. Therefore, RegSEM is a suitable tool to perform realistic simulations (i.e. including the full complexity of wave propagation in 3-D media) of correlation waveforms. As discussed by Tromp *et al.* (2010), this is crucial to go beyond the ray theory classically used when inverting noise-correlation data. To mimic noise sources, the code imposes a random traction at the free surface of the chunk using the \mathbb{I}^{trac} term

in eq. 1. First correlations computed with RegSEM can be found in Stehly *et al.* (2011).

Eventhough RegSEM is a well advanced code, improvements are considered. First, there are simple options that could be easily added, such as introducing two layers of elements in thick crusts and allowing for external source time functions. Second, the creation of a realistic 3-D unstructured hexahedral mesh still remains a long and little flexible process. Moreover, the computation cost can be significantly high due to small or deformed elements resulting from the meshing of the interfaces. A de-refinement in depth could be set up to decrease the computation cost. Indeed, seismic velocities increase with depth so the wavelengths in the deeper part of the chunk are larger than in the shallow part. Therefore, keeping the same horizontal size of the elements everywhere in the medium yields an oversampling of the wavefield in the deeper part. This is particularly obvious in PREM (Fig. 2) when going into the lower mantle (i.e. crossing the 670 km discontinuity). To partially solve the problems associated with the mesh generation, we could think of a performant coupling between a tetrahedral SEM (Mercerat *et al.* 2006) for complex geometries and a classical hexahedral SEM for the rest of the domain. Refinement and de-refinement not only in space but also in time could also greatly help for the reduction in computation cost. More promising alternatives for the modelling of complex heterogeneities can be found in the field of mechanics. Homogeneization techniques, originally developed in material mechanics for the static case, have been recently applied to the propagation of seismic waves (Capdeville *et al.* 2010; Guillot *et al.* 2010). We plan to adapt RegSEM to this new technique in the future. Third, the PML could be improved. Figs 9 and 10 show the presence of a spurious wave in the medium. The amplitude of this wave is small but it could be even smaller if using filtering PML (Festa *et al.* 2005) or implementing the unsplit convolutional PML suggested by Martin & Komatitsch (2009). Finally, fluid could be incorporated in the code, using either a normal mode coupling

16 P. Cupillard et al.

(Capdeville *et al.* 2003) or the acoustic version of the SEM. All these improvements are in progress and will be available in future versions.

ACKNOWLEDGMENTS

The authors want to thank Andreas Fichtner and Carl Tape for their insightful reviews which greatly helped in improving the manuscript. Interesting comments also emerged from discussions during different QUEST meetings. Remarks from Barbara Romanowicz, Diego Mercerat, Laurent Guillot, Ana Ferreira, Nikolai Shapiro and Mark Panning have been appreciated. The authors thank Dimitri Komatitsch for providing the routines which compute internal forces and attenuation, as well as Marco Stupazzini for his precious help in the use of the CUBIT mesh generator. Many thanks also to Huaiyu Yuan, Laurent Stehly, Yilong Qin, Aimin Cao and Murat Erduran for using RegSEM and reporting some bugs. The experience of Geneviève Moguilny and Patrick Stoclet from the SCP (*Service de Calcul Parallèle*) of the IGP has also been helpful for developing the code. The simulations presented in this paper have been performed using the Berkeley Seismological Laboratory cluster and the SCP cluster. ANR mémé partially supported this work.

REFERENCES

- Alterman, Z.S. & Karal, F.C., 1968. Propagation of elastic waves in layered media by finite-difference method, *Bull. seism. Soc. Am.*, **58**, 367–398.
- Backer, G., 1976. Error estimates for the finite element method for second order hyperbolic equations, *SIAM J. Numer. Anal.*, **13**, 564–575.
- Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS, Trans. Am. geophys. Un.*, **81**, F897.
- Bielak, J. & Christiano, P., 1984. On the effective seismic input for non-linear soil-structure interaction systems, *Earthq. Eng. Struct. Dyn.*, **12**, 107–119.
- Boore, D.M., 1972. Finite-difference method for seismic wave propagation in heterogeneous materials, in *Methods in Computational Physics*, Vol. 11, pp. 1–37, ed. Bolt, B.A., Academic Press, New York, NY.
- Bozdag, E. & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophys. J. Int.*, **172**, 1066–1082, doi:10.1111/j.1365-246X.2007.03690.x.
- Bozdag, E. & Trampert, J., 2010. Assessment of tomographic mantle models using spectral element seismograms, *Geophys. J. Int.*, **180**, 1187–1199, doi:10.1111/j.1365-246X.2009.04468.x.
- Capdeville, Y., 2000. Méthode couplée éléments spectraux – solution modale pour la propagation d’ondes dans la Terre à l’échelle globale. *PhD thesis*, Université Paris 7.
- Capdeville, Y. & Marigo, J., 2008. Shallow layer correction for spectral element like methods, *Geophys. J. Int.*, **172**, 1135–1150.
- Capdeville, Y., Chaljub, E., Vilotte, J.P. & Montagner, J.P., 2003. Coupling the Spectral Element Method with a modal solution for elastic wave propagation in global Earth models, *Geophys. J. Int.*, **152**, 34–66.
- Capdeville, Y., Guillot, L. & Marigo, J., 2010. 2-D nonperiodic homogenization to upscale elastic media for P-SV waves, *Geophys. J. Int.*, **182**, 903–922, doi:10.1111/j.1365-246X.2010.04636.x.
- Casarotti, E. et al., 2008a. Mesh generation for short-period seismic wave propagation based upon the spectral-element method: southern California, *Eos, Trans. Am. geophys. Un.*, **89**(53), S61B–1119 (abstract), Fall meet. suppl.
- Casarotti, E., Stupazzini, M., Lee, S., Komatitsch, D., Piersanti, A. & Tromp, J., 2008b. CUBIT and seismic wave propagation based upon the spectral-element method: an advanced unstructured mesher for complex 3D geological media, in *Proceedings of the 16th International Meshing Roundtable*, pp. 579–597, eds Brewer, M.L. & Marcum, D., Springer, Berlin.
- Cervený, V., 2001. *Seismic Ray Theory*, Cambridge University Press, Cambridge.
- Chaljub, E., 2000. Modélisation numérique de la propagation d’ondes sismiques à l’échelle du globe. *PhD thesis*, Université Paris 7.
- Chaljub, E., Capdeville, Y. & Vilotte, J., 2003. Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids, *J. Comp. Physics*, **183**, 457–491.
- Chaljub, E., Moczo, P., Tsuno, S., Bard, P.-Y., Kristek, J., Käser, M., Stupazzini, M. & Kristekova, M., 2010. Quantitative comparison of four numerical predictions of 3D ground motion in the Grenoble valley, France, *Bull. seism. Soc. Am.*, **100**(4), 1427–1455.
- Chaljub, E. & Valette, B., 2004. Spectral element modelling of three-dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core, *Geophys. J. Int.*, **183**, 131–141.
- Chen, M., Tromp, J., Helmberger, D. & Kanamori, H., 2007a. Waveform modeling of the slab beneath Japan, *J. geophys. Res.*, **112**, B02305, doi:10.1029/2006JB004394.
- Chen, P., Jordan, T.H. & Zhao, L., 2007b. Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods, *Geophys. J. Int.*, **170**, 175–181.
- Clévédy, E. & Lognonné, P., 1996. Fréchet derivatives of coupled seismograms with to an anelastic rotating earth, *Geophys. J. Int.*, **124**, 456–482.
- Curtis, A., Dost, B., Trampert, J. & Snieder, R., 1998. Eurasian fundamental mode surface wave phase velocities and their relationship to tectonic structures, *J. geophys. Res.*, **103**, 26 919–26 947.
- Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*. Princeton University Press, NJ.
- Delavaud, E., 2007. Simulation numérique de la propagation d’ondes en milieu géologique complexe: application à l’évaluation de la réponse sismique du bassin de Caracas (Vénézuéla). *PhD thesis*, Université Paris 7.
- Delavaud, E., Cupillard, P., Festa, G. & Vilotte, J.P., 2006. 3D Spectral Element Method simulations of the seismic response in the Caracas basin, in *Proceedings of the Third International Symposium on the Effects of Surface Geology on Seismic Motion*, Grenoble, France, pp. 515–522.
- Dumbser, M. & Käser, M., 2006. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - II. The three-dimensional isotropic case, *Geophys. J. Int.*, **167**(1), 319–336.
- Dupond, T., 1973. A L^2 estimate of Galerkin methods for second order hyperbolic equations, *SIAM J. Numer. Anal.*, **10**, 880–891.
- Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**, 297–356.
- Emmerich, H. & Korn, M., 1987. Incorporation of attenuation into time-domain computations of seismic wave fields, *Geophysics*, **52**, 1252–1264.
- Faccioli, E., Maggio, F., Paolucci, R. & Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudospectral domain decomposition method, *J. seism.*, **1**, 237–251.
- Ferreira, A.M.G., Woodhouse, J.H., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography, *J. geophys. Res.*, **115**, B04313, doi:10.1029/2009JB006716.
- Festa, G. & Vilotte, J.P., 2005. The Newmark scheme as velocity-stress time-staggering: an efficient PML for spectral element simulations of elastodynamics, *Geophys. J. Int.*, **161**(3), 789–812.
- Festa, G., Delavaud, E. & Vilotte, J.P., 2005. Interaction between surface waves and absorbing boundaries for wave propagation in geological basins: 2D numerical simulations, *Geophys. Res. Lett.*, **32**, L20306, doi:10.1029/2005GL024091.
- Fichtner, A. & Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models - a technique based on non-linear dispersion curve matching (dcm), *Geophys. J. Int.*, **173**(2), 519–533.
- Fichtner, A., Bunge, H.-P. & Igel, H., 2006a. The adjoint method in seismology: I - Theory, *Phys. Earth planet. Inter.*, **157**, 86–104.
- Fichtner, A., Bunge, P. & Igel, H., 2006b. The adjoint method in seismology: II - Applications: traveltimes and sensitivity functionals, *Phys. Earth planet. Inter.*, **157**, 105–123.

- Fichtner, A., Igel, H., Bunge, H.P. & Kennett, B.L.N., 2009a. Simulation and inversion of seismic wave propagation on continental scales based on a Spectral-Element Method, *J. Numer. Anal. Ind. appl. Math.*, **4**, 11–22.
- Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2009b. Full waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.*, **179**(3), 1703–1725.
- Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2010. Full waveform tomography for radially anisotropic structure: new insights into present and past states of the Australasian upper mantle, *Earth planet. Sci. Lett.*, **290**, 270–280.
- Gauthier, O., Virieux, J. & Tarantola, A., 1986. Two-dimensional non-linear inversion of seismic waveforms: numerical results, *Geophysics*, **51**(7):1387–1403.
- Gilbert, F., 1971. Excitation of normal modes of the Earth by earthquake sources, *Geophys. J. R. astr. Soc.*, **22**, 223–226.
- Guillot, L., Capdeville, Y. & Marigo, J., 2010. 2-D nonperiodic homogenization of the elastic wave equation: SH case, *Geophys. J. Int.*, **182**, 1438–1454, doi:10.1111/j.1365-246X.2010.04688.x.
- Kelly, K.R., Ward, R.W., Treitel, S. & Alford, R.M., 1976. Synthetic seismograms, a finite difference approach, *Geophysics*, **41**, 2–27.
- Komatitsch, D. & Tromp, J., 1999. Introduction to the Spectral Element Method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139**, 806–822.
- Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation, part I: validation, *Geophys. J. Int.*, **149**, 390–412.
- Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation, part II: 3-D models, oceans, rotation, and gravity, *Geophys. J. Int.*, **150**, 303–318.
- Komatitsch, D. & Vilotte, J.P., 1998. The Spectral Element Method: an efficient tool to simulate the seismic response of 2d and 3d geological structures, *Bull. seism. Soc. Am.*, **88**, 368–392.
- Komatitsch, D., Liu, Q., Tromp, J., Süss, P., Stidham, C. & Shaw, J.H., 2004. Simulations of ground motion in the Los Angeles basin based upon the Spectral-Element Method, *Bull. seism. Soc. Am.*, **94**(1), 187–206.
- Komatitsch, D., Ritsema, J. & Tromp, J., 2002. The Spectral-Element Method, Beowulf computing, and global seismology, *Science*, **298**, 1737–1742.
- Lee, S.-J., Chen, H.-W., Liu, Q., Komatitsch, D., Huang, B.-S. & Tromp, J., 2008. Three-dimensional simulations of seismic-wave propagation in the Taipei basin with realistic topography based upon the Spectral Element Method, *Bull. seism. Soc. Am.*, **98**(1), 253–264.
- Lekić, V., Panning, M. & Romanowicz, B., 2010. A simple method for improving crustal corrections in waveform tomography, *Geophys. J. Int.*, **182**, 265–278.
- Lognonné, P., 1991. Normal modes and seismograms in an anelastic rotating earth, *J. geophys. Res.*, **96**, 20 309–20 319.
- Lognonné, P. & Romanowicz, B., 1990. Modelling of coupled normal modes of the Earth: the spectral method, *Geophys. J. Int.*, **102**, 365–395.
- Longuet-Higgins, M.S., 1950. A theory on the origin of microseisms, *Philos. Trans. R. Soc. Lond. A.*, **243**, 1–35.
- Lysmer, J. & Drake, L.A., 1972. A finite element method for seismology, in *Methods in Computational Physics*, Vol. 11, Academic Press, New York, NY.
- Maday, Y. & Patera, A., 1989. Spectral Element Methods for the incompressible Navier-Stokes equations, in *State of the Art Survey in Computational Mechanics*, pp. 71–143, eds Noor, A. & Oden, J., ASME, New-York.
- Maday, Y. & Ronquist, E.M., 1990. Optimal error analysis of spectral methods with emphasis on non-constant coefficients and deformed geometries, *Comput. Methods Appl. Mech. Eng.*, **80**, 91–115.
- Marfurt, K.J., 1984. Accuracy of finite-difference and finite-element modeling of the scalar wave equation, *Geophysics*, **49**, 533–549.
- Marone, F. & Romanowicz, B., 2007. Non-linear crustal corrections in high-resolution regional waveform seismic tomography, *Geophys. J. Int.*, **170**, 460–467.
- Martin, R. & Komatitsch, D., 2009. An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation, *Geophys. J. Int.*, **179**, 333–344, doi:10.1111/j.1365-246X.2009.04278.x.
- Mercerat, E.D., Vilotte, J.P. & Sánchez-Sesma, F.J., 2006. Triangular spectral element simulation of two-dimensional elastic wave propagation using unstructured triangular grids, *Geophys. J. Int.*, **166**, 679–698.
- Moczo, P., Robertsson, J.O.A. & Eisner, L., 2007. The finite-difference time-domain method for modeling of seismic wave propagation, in *Advances in Wave Propagation in Heterogeneous Earth*, eds Wu, R., Maupin, V. & Dmowska, R., Vol. 48, pp. 421–516, Elsevier/Academic Press, San Diego.
- Montagner, J.P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.*, **96**, 20 337–20 351.
- Oliveira, S. & Seriani, G., 2011. Effect of element distortion on the numerical dispersion of Spectral Element Methods, *Commun. Comput. Phys.*, **9**, 937–958.
- Olsen, K.B., 2000. Site amplification in the Los Angeles basin from 3D modeling of ground motion, *Bull. seism. Soc. Am.*, **90**, S77–S94.
- Olsen, K.B. & Archuleta, R.J., 1996. 3-D simulation of earthquakes on the Los Angeles fault system, *Bull. seism. Soc. Am.*, **86**, 575–596.
- Patera, A.T., 1984. A Spectral Element Method for fluid dynamics: laminar flow in a channel expansion, *J. Comput. Phys.*, **54**, 468–488.
- Pelties, C., Käser, M., Hermann, V. & Castro, C.E., 2010. Regular versus irregular meshing for complicated models and their effect on synthetic seismograms, *Geophys. J. Int.*, **183**, 1031–1051, doi:10.1111/j.1365-246X.2010.04777.x.
- Peter, D. et al., 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**, 721–739, doi:10.1111/j.1365-246X.2011.05044.x.
- Qin, Y., Capdeville, Y., Montagner, J., Boschi, L. & Becker, T.W., 2009. Reliability of mantle tomography models assessed by spectral-element simulation, *Geophys. J. Int.*, **175**, 598–616.
- Ronchi, C., Iano, R. & Paolucci, P.S., 1996. The ‘Cubed Sphere’: a new method for the solution of partial differential equations in spherical geometry, *J. Comput. Phys.*, **124**, 93–114.
- Sadourny, R., 1972. Conservative finite-difference approximations of the primitive equation on quasi-uniform spherical grids, *Mon. Wea. Rev.*, **100**, 136–144.
- Sánchez-Sesma, F.J., 1983. Diffraction of elastic waves by three-dimensional surface irregularities, *Bull. seism. Soc. Am.*, **73**(6), 1621–1636.
- Seriani, G., 1998. 3D large-scale wave propagation modeling by a Spectral Element Method on a Cray T3E multiprocessor, *Comput. Methods appl. Mech. Eng.*, **164**(1), 235–247.
- Seriani, G. & Priolo, E., 1994. Spectral Element Method for acoustic wave simulation in heterogeneous media, *Finite Elem. Anal. Des.*, **16**, 337–348.
- Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, **31**, L07614, doi:10.1029/2004GL019491.
- Shapiro, N.M. & Ritzwoller, M.H., 2002. Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int.*, **151**, 88–105.
- Stehly, L., Cupillard, P. & Romanowicz, B., 2011. Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise, *C. R. Geoscience*, **343**, 591–599, doi:10.106/j.crte.2011.03.005.
- Stupazzini, M., Paolucci, R. & Igel, H., 2009. Near-fault earthquake ground-motion simulation in the Grenoble valley by a high-performance spectral element code, *Bull. seism. Soc. Am.*, **99**(1), 286–301.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science*, **325**, 988–992.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, **180**, 433–462.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259–1266.
- Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure appl. Geophys.*, **128**(1/2), 365–399.
- Toshinawa, T. & Ohmachi, T., 1992. Love wave propagation in three-dimensional sedimentary basin, *Bull. seism. Soc. Am.*, **82**, 1661–1667.

18 *P. Cupillard et al.*

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.

Tromp, J., Luo, Y., Hanasoge, S. & Peter, D., 2010. Noise cross-correlation sensitivity kernels, *Geophys. J. Int.*, **183**, 791–819, doi:10.1111/j.1365-246X.2010.04721.x.

Virieux, J., 1984. SH wave propagation in heterogeneous media: velocity-stress finite difference method, *Geophysics*, **49**, 1933–1942.

Virieux, J., 1986. P-SV wave propagation in heterogeneous media: velocity-stress finite difference method, *Geophysics*, **51**, 889–901.

Non-periodic homogenization of 3-D elastic media for the seismic wave equation

Paul Cupillard¹ and Yann Capdeville²

¹*GeoResources, Université de Lorraine/CNRS, F-54500 Vandœuvre-lès-Nancy, France. E-mail: paul.cupillard@univ-lorraine.fr*

²*LPGN, Université de Nantes/CNRS, F-44000 Nantes, France*

Accepted 2018 January 30. Received 2018 January 11; in original form 2017 July 28

SUMMARY

Because seismic waves have a limited frequency spectrum, the velocity structure of the Earth that can be extracted from seismic records has a limited resolution. As a consequence, one obtains smooth images from waveform inversion, although the Earth holds discontinuities and small scales of various natures. Within the last decade, the non-periodic homogenization method shed light on how seismic waves interact with small geological heterogeneities and ‘see’ upscaled properties. This theory enables us to compute long-wave equivalent density and elastic coefficients of any media, with no constraint on the size, the shape and the contrast of the heterogeneities. In particular, the homogenization leads to the apparent, structure-induced anisotropy. In this paper, we implement this method in 3-D and show 3-D tests for the very first time. The non-periodic homogenization relies on an asymptotic expansion of the displacement and the stress involved in the elastic wave equation. Limiting ourselves to the order 0, we show that the practical computation of an upscaled elastic tensor basically requires (i) to solve an elastostatic problem and (ii) to low-pass filter the strain and the stress associated with the obtained solution. The elastostatic problem consists in finding the displacements due to local unit strains acting in all directions within the medium to upscale. This is solved using a parallel, highly optimized finite-element code. As for the filtering, we rely on the finite-element quadrature to perform the convolution in the space domain. We end up with an efficient numerical tool that we apply on various 3-D models to test the accuracy and the benefit of the homogenization. In the case of a finely layered model, our method agrees with results derived from Backus. In a more challenging model composed by a million of small cubes, waveforms computed in the homogenized medium fit reference waveforms very well. Both direct phases and complex diffracted waves are accurately retrieved in the upscaled model, although it is smooth. Finally, our upscaling method is applied to a realistic geological model. The obtained homogenized medium holds structure-induced anisotropy. Moreover, full seismic wavefields in this medium can be simulated with a coarse mesh (no matter what the numerical solver is), which significantly reduces computation costs usually associated with discontinuities and small heterogeneities. These three tests show that the non-periodic homogenization is both accurate and tractable in large 3-D cases, which opens the path to the correct account of the effect of small-scale features on seismic wave propagation for various applications and to a deeper understanding of the apparent anisotropy.

Key words: Numerical solutions; Computational seismology; Seismic anisotropy; Theoretical seismology; Wave propagation; Wave scattering and diffraction.

1 INTRODUCTION

Seismic waves are one of the most powerful tools to image the Earth’s interior. Giving access to the geometry of geological structures and to the distribution of mechanical properties within our planet, they lead to a better understanding of geodynamic processes

and resource potentials. In the last decades, the seismic tomography and imaging community took advantage of the increasing computational power and the development of efficient numerical methods to improve its techniques and results. Accurate solutions to the forward modelling (see Virieux *et al.* 2011, for a review of the various numerical methods available to model the seismic wave propagation)

and the inverse problem (e.g. Pratt *et al.* 1998; Tromp *et al.* 2005; Plessix 2006; Fichtner *et al.* 2008; Métivier *et al.* 2013; Brossier *et al.* 2015; Warner & Guasch 2016) now allow accounting for full seismic waveforms and getting robust and well-resolved models of the Earth at different scales (from the subsurface in exploration geophysics to the entire globe in seismology).

One of the important remaining challenges in seismic wave simulation and inversion is the understanding of and the correct account for the effect of small heterogeneities on wave propagation. By small heterogeneities we here mean structures which are smaller than the minimum seismic wavelength propagating in the medium. Seismic waves indeed have a finite frequency band which implies the existence of a minimum wavelength, whereas heterogeneities within the Earth can occur at all scales. When propagating through small heterogeneities, seismic waves naturally average the elastic properties of the medium. A deeper understanding of this averaging process could significantly improve the interpretation of seismic inversion results, pointing out what small-scale features can be ‘hidden’ behind the smooth images one usually obtains. A proper knowledge of the averaging is of major interest for the forward modelling as well, because replacing a given discontinuous and highly heterogeneous model by an effective (or, equivalently, ‘upscaled’ or ‘long-wave equivalent’) medium greatly eases the numerical simulation of wave propagation. When present, small-scale features indeed control the spatial sampling of the model and, consequently, the time sampling too. This is because of a stability condition that all the wave simulators have to ensure (Courant *et al.* 1928):

$$\Delta t \leq C \left[\frac{\Delta x}{V_P} \right]_{\min}, \quad (1)$$

Δt denoting the time-step of the simulation, C a constant and $\left[\frac{\Delta x}{V_P} \right]_{\min}$ the minimum ratio of grid spacing Δx and P -wave speed V_P . Small heterogeneities therefore induce massive, possibly prohibitive, computation costs. In addition, an accurate simulation usually requires all the physical discontinuities to be honoured by the mesh, which can lead to enormous meshing efforts, especially when dealing with hexahedra (e.g. Casarotti *et al.* 2008; Peter *et al.* 2011). For these reasons, working with effective media is much preferable.

Theoretical studies for understanding the effective properties of heterogeneous elastic media was initiated in the sixties by Hashin & Shtrikman (1963) and Hill (1965). These studies both use the pioneering ideas of Voigt (1889) and Reuss (1929) to derive upper and lower-bound effective constants of periodic materials. After these works, many analytical derivations followed (e.g. Kutsenko *et al.* 2013, and the references within), handling more and more complex periodic units or improving accuracy and efficiency of the effective properties calculation. In the case of waves propagating in finely layered media, Backus (1962) derived formula for long-wave equivalent elastic coefficients which are still widely used in seismic exploration. Fruitfully, these formulae quantify the seismic anisotropy produced by fine layering, showing that a stack of isotropic layers can explain the anisotropy observed on wave measurements. Schoenberg & Muir (1989) extended the Backus theory by including any kind of anisotropy within each layer, which enables the account for several sets of fractures in the layers. Further studies on the upscaling of fractured media then came out, such as Sayers & Kachanov (1991); Mauge & Kachanov (1994); Sayers & Kachanov (1995); Schoenberg & Sayers (1995); Schoenberg & Helbig (1997); Sayers (1998); Grechka & Kachanov (2006); Grechka (2007); Carcione *et al.* (2012). At larger scale, particular efforts have been focused on smoothing the Earth’s crust to ease the simulation of

long-period surface waves (Capdeville & Marigo 2008; Fichtner & Igel 2008; Lekić *et al.* 2010). More recently, the understanding of the seismic anisotropy in the upper-mantle drew the attention (Fichtner *et al.* 2013a; Wang *et al.* 2013; Bodin *et al.* 2015). In this context, Jordan (2015) developed an elegant effective medium theory for random media in which statistics of the local anisotropy (produced by lattice-preferred orientation of mineral grains for instance) are separated from those of ellipsoidal geometric heterogeneities (i.e. shape-preferred orientation of simple geological structures).

In this paper, the upscaling method in consideration is the two-scale homogenization technique. This method emerged in the seventies from research in micromechanics for predicting the macroscopic response of composite and random materials to either static or dynamic excitations (Bensoussan *et al.* 1978; Sanchez-Palencia 1980; Papanicolaou & Varadhan 1981). Since then, the technique has been successfully applied to many physical processes, such as heat transfer (e.g. Allaire & Habibi 2013), Stokes flow (e.g. Hornung 1997), neutronic diffusion (e.g. Allaire & Capdeboscq 2000), magnetization (e.g. Santugini-Repique 2007) and elastic wave propagation (e.g. Boutin & Auriault 1993; Fish & Chen 2001, 2004; Parnell & Abrahams 2008; Bacigalupo & Gambarotta 2014). In this last field, the two-scale homogenization was adapted to non-periodic media within the last decade (Capdeville & Marigo 2007; Capdeville *et al.* 2010a,b; Guillot *et al.* 2010; Cance & Capdeville 2015; Capdeville *et al.* 2015), which opened the path to the upscaling of general elastic media, with no constraint on the shape and size of the heterogeneities. The method has been tested in 1-D (Capdeville & Marigo 2007; Capdeville *et al.* 2010a) and 2-D (Capdeville *et al.* 2010b; Guillot *et al.* 2010; Capdeville *et al.* 2015); the goal of the present paper is to show its efficiency (in terms of accuracy and computation speed) on 3-D cases. To solve the homogenization problem, we will rely on a finite-element method, but it is worth noting here that the algorithm proposed by Capdeville *et al.* (2015) could be used as well. This latter is up and running on continuous or pixel-based 3-D media.

The non-periodic two-scale homogenization actually is different to what some authors called numerical homogenization (e.g. Zijl *et al.* 2002; Grechka 2003; Gao *et al.* 2015) or heterogeneous multi-scale methods (e.g. Engquist *et al.* 2009; Abdulle & Grote 2011). In these classes of techniques, the effective elastic tensor is computed by solving the wave equation in the static regime (i.e. at zero frequency so that a simple elastostatic problem arises) with a set of unit stresses applied on a representative volume. Such a volume implicitly delineates the small and the large scales. It fully depends on the mesh and the numerical technique that will then be used for simulating seismic waves. Even though our non-periodic homogenization method also includes the numerical resolution of a static problem, this latter does not involve a representative volume. Our static problem is actually defined on the whole medium at once with volume forces equal to the divergence of the elastic tensor (eq. 33). Low-pass filtering the strain and the stress which result from these forces then leads to the effective properties (eq. 30). Our filter only depends on the minimum wavelength to be propagated in the medium and a precision factor, so that the scales are separated by the waves rather than imposed by a numerical method. This makes non-periodic homogenization results physically meaningful and independent from wave equation solvers. Homogenized properties actually reveal what the waves ‘see’. In particular, Capdeville *et al.* (2013) showed that models obtained by full-waveform inversion are homogenized models.

Applying the non-periodic homogenization method in 3-D does not require further theoretical derivations than those developed in

2-D by Capdeville *et al.* (2010b) and Guillot *et al.* (2010). In a first part, the present paper recalls the ideas and concepts of the method. From this theoretical part emerge the practical issues which have to be tackled to get homogenized properties, namely (i) the resolution of an elastostatic problem and (ii) a filtering process. The second part of the article describes the implementation of a parallel finite-element method for solving the 3-D elastostatic problem and provides some details on the 3-D filtering. In a third part, the accuracy and performance of the resulting homogenization code are challenged on various elastic models: (i) a finely layered medium for which analytical expressions of the upscaled properties are available, (ii) a large and highly heterogeneous medium in which reference seismograms can be computed and (iii) a realistic geological model made of multiple folded and faulted horizons. Finally, we discuss the possible improvements of our code and the numerous perspectives opened by the 3-D homogenization method in terms of solving forward and inverse problems.

2 THE NON-PERIODIC HOMOGENIZATION THEORY IN A NUTSHELL

We here attempt to give a synoptic overview of the non-periodic homogenization theory, starting from basics introduced in micromechanics for 1-D periodic materials (e.g. Bensoussan *et al.* 1978), then extending these basics to non-periodic media (Capdeville *et al.* 2010a), and finally moving to the general 2-D/3-D non-periodic case (Capdeville *et al.* 2010b; Guillot *et al.* 2010). Our goal is to provide a digest of the theory to make the non-periodic homogenization understandable by a large number of geophysicists, so we intentionally skip some technical details of the whole Capdeville *et al.* (2010a,b) derivation to focus on the main ideas and concepts of the method.

2.1 1-D periodic homogenization

The homogenization theory relies on an ansatz for the solution of the physical problem in consideration. In elastodynamics, the displacement $u(x; t)$ and the stress $\sigma(x; t)$ involved in the 1-D case (x being the space variable and t being the time) are postulated to be two-scale asymptotic expansions:

$$u(x; t) = \sum_{i=0}^{+\infty} \varepsilon^i u_i \left(x, \frac{x}{\varepsilon}; t \right), \quad (2)$$

$$\sigma(x; t) = \sum_{i=0}^{+\infty} \varepsilon^i \sigma_i \left(x, \frac{x}{\varepsilon}; t \right), \quad (3)$$

where $\varepsilon = \frac{l}{\lambda_m}$ is the ratio of the size of the periodic cell which constitutes the 1-D medium to the minimum wavelength propagating in this medium (Fig. 1). By definition, l is microscopic and λ_m is macroscopic. ε , which is called scaling parameter, is therefore much smaller than 1. It enables to explicitly separate the scales within coefficients u_i and σ_i of series (2) and (3), x capturing the large-scale variations and $y = \frac{x}{\varepsilon}$ handling the small-scale variations. Because any change in y induces a very slight change in x , the two variables can be treated independently and the spatial derivative operator can be written

$$\nabla = \varepsilon^{-1} \nabla_y + \nabla_x. \quad (4)$$

The introduction of the small-scale variable $y = \frac{x}{\varepsilon}$ also allows us to rewrite the physical parameters $E(x)$ (Young modulus) and $\rho(x)$ (density) of the medium as λ_m -periodic quantities $E(y)$ and

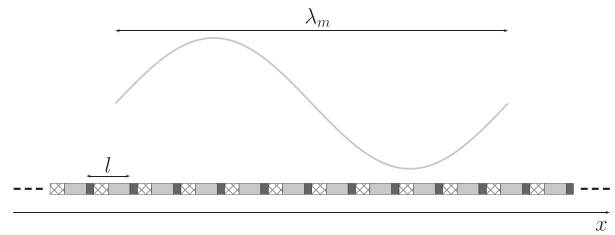


Figure 1. 1-D periodic homogenization framework: a wavefield having a minimum wavelength λ_m propagates in an infinite 1-D medium made of periodic cells whose size l is much smaller than λ_m .

$\rho(y)$. In other words, the l -periodic bar in the initial one-variable problem is now seen as a medium containing only small scales which are repeated in space with a λ_m -periodicity. Furthermore, coefficients u_i and σ_i are assumed to be λ_m -periodic in y as are E and ρ . This assumption imposes that small-scale variations of the displacement and stress fields are due to local small-scale structures of the medium. In seismology, such a phenomenon is commonly called site effect.

Plugging eqs (2)–(4) into the elastodynamic problem (i.e. the wave equation and Hooke's law) yields a cascade of coupled equations which can be solved for each i using the average over the periodic cell

$$\langle f \rangle(x) = \frac{1}{\lambda_m} \int_{-\frac{\lambda_m}{2}}^{\frac{\lambda_m}{2}} f(x, y) dy, \quad \forall f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (5)$$

and the periodicity in y of the problem. Doing so, it turns out that the zeroth-order terms u_0 and σ_0 do not depend on the microscopic variable y . This result formalizes the poor sensitivity of the wavefield to small heterogeneities. Going further, one shows that u_0 and σ_0 are the solution of the so-called homogenized problem, which is a classical elastodynamic problem with homogeneous effective properties E^* and ρ^* such that

$$\rho^* = \langle \rho(y) \rangle \quad (6)$$

and

$$E^* = \langle E(y)[1 + \nabla_y \chi(y)] \rangle. \quad (7)$$

In this last equation, χ is the so-called first-order corrector. It is the solution of the so-called cell problem, which is

$$\nabla_y [E(1 + \nabla_y \chi)] = 0 \quad (8)$$

defined on the cell with periodic boundary conditions. Because an analytical solution exists for this problem, eq. (7) finally reduces to

$$E^* = \left\langle \frac{1}{E(y)} \right\rangle^{-1}. \quad (9)$$

In conclusion, the quantity to average for obtaining the zeroth-order long-wave equivalent Young modulus E^* is the inverse of the initial Young modulus E (eq. 9), whereas the zeroth-order long-wave equivalent density ρ^* simply is the average of the initial density ρ (eq. 6).

Contrary to u_0 and σ_0 , the higher-order terms of series (2) and (3) contain small-scale variations. For instance, the first-order displacement term is y -dependent through the first-order corrector:

$$u_1(x, y) = \chi(y) \nabla_x u_0(x) + \langle u_1 \rangle(x). \quad (10)$$

Examples of the contribution of this last term to the whole wavefield can be found in Capdeville *et al.* (2010a). In the present work, we

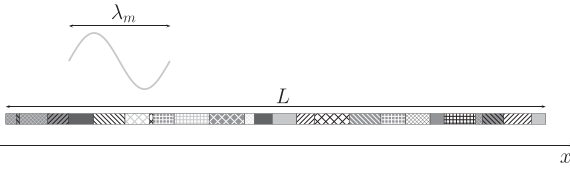


Figure 2. 1-D non-periodic homogenization framework: a wavefield having a minimum wavelength λ_m propagates in a finite 1-D medium of size L . Contrary to the periodic case, no assumption is made on the distribution of the mechanical properties $E(x)$ and $\rho(x)$ within the medium.

focus on the order 0 so we do not provide any further details on the higher-order terms.

A crucial point in the homogenization theory is that the asymptotic convergence can be proved mathematically. One actually does not directly show the convergence of series (2) and (3) towards the exact wavefield; one uses the so-called Γ -convergence instead (Dal Maso 1993). Rather than studying a single problem for the physically relevant value of ε , one considers a sequence of problems indexed by ε which is now regarded as a small parameter going to zero. In other words, one builds a fictitious sequence of problems in which the periodic cell becomes smaller and smaller. In this context, one demonstrates that the exact solution converges to the solution of the homogenized problem u_0 and σ_0 (Nguetseng 1989; Allaire 1992). Such a demonstration provides rigorous mathematical foundations for the homogenization theory.

2.2 1-D non-periodic homogenization

Let us now consider any given distribution of the properties E and ρ within a finite 1-D medium. The length of the medium is denoted by L . Again, our goal here is to find long-wave equivalent properties. We denote by λ_m the minimum wavelength of the wavefield propagating in the bar (Fig. 2). Inspired by the periodic homogenization, we will represent the displacement $u(x; t)$ and the stress $\sigma(x; t)$ as power series in ε (to be redefined) with coefficients u_i and σ_i as functions of x and $y = \frac{x}{\varepsilon}$. Plugging these series into the wave equation and Hooke's law will yield a cascade of equations. To get a solution for these equations, let us bring a periodicity into the problem by imposing periodic boundary conditions at the border of the domain. We end up with an infinite medium made of a periodic cell, so we can certainly benefit from the mathematical results of the previous section (eqs 6 and 9). Nevertheless, two big issues remain:

(i) Because the boundary conditions are forced to be periodic, the effective properties computed near the border of the domain are not meaningful for any kind of physically interesting conditions such as Dirichlet or Neumann conditions. In the present paper, we do not investigate this issue. Disregarding boundary effects, we focus on the computation of accurate effective properties in the interior of the domain. Precisions on what we exactly mean by the *interior of the domain* are given in Part 3.

(ii) Contrary to the periodic case, the size of the periodic cell is not microscopic. The cell can contain various sizes of heterogeneity, including macroscopic scales (i.e. sizes equal to or larger than λ_m), so ε defined as the ratio of the size L of the periodic cell to the minimum wavelength λ_m no longer is a relevant scaling parameter to separate the scales through variables x and $y = \frac{x}{\varepsilon}$. The goal of the present section is (i) to redefine ε and (ii) to separate the scales within the mechanical properties of the bar in a way that allows for a solution of our homogenization problem.

2.2.1 Redefining ε

Because there is no quantity for defining the small scales yet, we introduce $\lambda_0 < \lambda_m$: all the heterogeneities whose size is smaller than λ_0 are considered as small. λ_0 therefore is analogous to l in the periodic case. Using it, we can bring in a new scaling parameter $\varepsilon_0 = \frac{\lambda_0}{\lambda_m} < 1$ and we can choose $\varepsilon \leq \varepsilon_0$. Similarly to the periodic case, ε is a formal quantity which can go to zero to prove the Γ -convergence of the asymptotic solution. It can be seen as the ratio of a small length $\lambda \leq \lambda_0$ to the minimum wavelength of the wavefield to be propagated λ_m , λ becoming smaller and smaller when studying the convergence. In practice, the only physically relevant value of ε is ε_0 .

With this definition of ε , variables x and $y = \frac{x}{\varepsilon}$ can be treated independently. We denote by Ψ the set of functions $f(x, y): \mathbb{R}^2 \rightarrow \mathbb{R}$ such that the x part of f carries the large-scale variations while the y part of f handles the small-scale variations. As in the periodic case, we impose coefficients u_i and σ_i of the asymptotic expansions (2) and (3) to belong to Ψ . Moreover, we can write the spatial derivative operator ∇ as in eq. (4).

2.2.2 Separating the scales within the 1-D medium

In the non-periodic case, the main issue we face is that the medium contains both microscopic and macroscopic scales, so we cannot write E and ρ as a function of y only. Using a low-pass filter $\mathcal{F}^{\varepsilon_0}$ (Appendix A) to separate large-scale and small-scale variations, we will have to find a proper way to build $E^{\varepsilon_0}(x, y)$ and $\rho^{\varepsilon_0}(x, y)$ from $E(x)$ and $\rho(x)$. By *proper* we here mean *which allows for a solution to the cascade of equations that arise when plugging eqs (2)–(4) into the elastodynamic problem*.

Let us introduce $\eta = \frac{L}{\lambda}$ (which implies that $\varepsilon = \frac{L}{\eta\lambda_m}$) and assume that we properly built $E^{\varepsilon_0}(x, y)$ and $\rho^{\varepsilon_0}(x, y)$ from $E(x)$ and $\rho(x)$. Because E and ρ are L -periodic in x , E^{ε_0} and ρ^{ε_0} are $\eta\lambda_m$ -periodic in y . Assuming that coefficients $u_i(x, y)$ and $\sigma_i(x, y)$ are also $\eta\lambda_m$ -periodic in y and using the average over the periodic cell

$$\langle f \rangle(x) = \frac{1}{\eta\lambda_m} \int_{-\frac{\eta\lambda_m}{2}}^{\frac{\eta\lambda_m}{2}} f(x, y) dy, \quad \forall f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (11)$$

we can easily derive the cascade of equations. Similarly to the periodic case, it turns out that the zeroth-order displacement u_0 and stress σ_0 do not depend on y . Again, these fields are the solution of the so-called homogenized problem, which is a classical elastodynamic problem involving effective properties ρ^{ε_0*} and E^{ε_0*} such that

$$\rho^{\varepsilon_0*}(x) = \langle \rho^{\varepsilon_0}(x, y) \rangle \quad (12)$$

and

$$E^{\varepsilon_0*}(x) = \langle E^{\varepsilon_0}(x, y)[1 + \nabla_y \chi^{\varepsilon_0}(x, y)] \rangle. \quad (13)$$

Note that the effective properties no longer are homogeneous in the non-periodic case. Moreover, they depend on the choice of ε_0 . The first-order corrector $\chi^{\varepsilon_0}(x, y)$ also depends on ε_0 . It necessarily belongs to Ψ and it is $\eta\lambda_m$ -periodic in y . Similarly to the periodic case, it is the solution of the cell problem, that is,

$$\nabla_y [E^{\varepsilon_0}(1 + \nabla_y \chi^{\varepsilon_0})] = 0 \quad (14)$$

with periodic boundary conditions. As shown by Capdeville *et al.* (2010a), function $\nabla_y \chi^{\varepsilon_0}(x, y)$ therefore satisfies

$$\nabla_y \chi^{\varepsilon_0} = -1 + \left\langle \frac{1}{E^{\varepsilon_0}} \right\rangle^{-1} \frac{1}{E^{\varepsilon_0}}. \quad (15)$$

This last equation implies that

(i) Eq. (13) reduces to

$$E^{\varepsilon_0^*}(\mathbf{x}) = \left\langle \frac{1}{E^{\varepsilon_0}(\mathbf{x}, \mathbf{y})} \right\rangle^{-1}, \quad (16)$$

which along with eq. (12) tells the quantity to average to get the zeroth-order effective properties.

(ii) $\frac{1}{E^{\varepsilon_0}(\mathbf{x}, \mathbf{y})}$ belongs to Ψ , which is the condition to meet for building $E^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$ properly. Another equation from our cascade tells that $\rho^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$ has to lie in Ψ (Appendix B), so we also have a condition for building $\rho^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$ properly. Following these conditions, we easily form

$$E^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \left[\mathcal{F}^{\varepsilon_0} \left\{ \frac{1}{E} \right\}(\mathbf{x}) + \left[\frac{1}{E} - \mathcal{F}^{\varepsilon_0} \left\{ \frac{1}{E} \right\}(\mathbf{y}) \right] \right]^{-1} \quad (17)$$

and

$$\rho^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{x}) + [\rho - \mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{y})]. \quad (18)$$

Fig. 3 illustrates the mathematical construction. In this figure, symbol g represents either ρ or $\frac{1}{E}$, and FT stands for Fourier transform. From any distribution of g over a length L extended to \mathbb{R} by periodicity (Fig. 3a), large scales (Fig. 3b) and small scales (Fig. 3c) are extracted using $\mathcal{F}^{\varepsilon_0}$. Within the small scales, space variable x is replaced by $y = \frac{x}{\varepsilon}$ (Fig. 3d), thus changing the L -periodicity in x to a $\eta\lambda_m$ -periodicity in y . Then, the two-variable quantity $g^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$ is formed by simply adding the large scales (which are a function of x) to the small scales (which are a function of y). Fig. 3(e) represents g^{ε_0} for a particular value of $x = \bar{x}$; it shows the small scales oscillating around $\mathcal{F}^{\varepsilon_0} \{ g \}(\bar{x})$.

2.2.3 Final results

Including eq. (18) into (12) yields

$$\begin{aligned} \rho^{\varepsilon_0^*}(\mathbf{x}) &= (\mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{x}) + [\rho - \mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{y})]) \\ &= (\mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{x}) + \langle [\rho - \mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{y})] \rangle) \\ &= \mathcal{F}^{\varepsilon_0} \{ \rho \}(\mathbf{x}). \end{aligned} \quad (19)$$

Similarly, including eq. (17) into (16) leads to

$$E^{\varepsilon_0^*}(\mathbf{x}) = \left[\mathcal{F}^{\varepsilon_0} \left\{ \frac{1}{E} \right\}(\mathbf{x}) \right]^{-1}. \quad (20)$$

Eqs (19) and (20) tell that the zeroth-order effective properties of any 1-D medium can be computed by filtering the density and the inverse of the Young modulus. This result can be easily intuited from the solution inferred in the periodic case (eqs 6 and 9). We have here given a rigorous demonstration for it.

To derive a zeroth-order solution for our non-periodic homogenization problem, we introduced a new scaling parameter ε_0 . Because the obtained effective medium depends on this parameter (eqs 19 and 20), u_0 and σ_0 also depend on it. For sake of simplicity, we did not index these two fields by ε_0 as Capdeville *et al.* (2010a) did. Nevertheless, we will study the ε_0 -convergence of these fields in a specific case (Section 4.2).

2.3 3-D non-periodic homogenization

In the 3-D case, Hooke's law involves a fourth-order tensor \mathbf{C} to relate the stress and the strain, so the linear elastic behaviour of a given medium no longer is fully described by E only. Nevertheless, the homogenization theory developed for 1-D non-periodic media

is still valid up to the zeroth-order effective properties (12) and (13) which are now written

$$\rho^{\varepsilon_0^*}(\mathbf{x}) = \langle \rho^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) \rangle_3 \quad (21)$$

and

$$\mathbf{C}^{\varepsilon_0^*}(\mathbf{x}) = \left\langle \mathbf{C}^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) : \left[\mathbf{I} + \frac{1}{2} (\nabla_y \chi^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) + {}^t \nabla_y \chi^{\varepsilon_0}(\mathbf{x}, \mathbf{y})) \right] \right\rangle_3, \quad (22)$$

where \mathbf{I} is the fourth-order identity tensor, t is the transpose operator, $:$ is the tensor contraction $[\mathbf{A} : \mathbf{B}]_{ijkl} = A_{ijmn} B_{mnkl}$ and $\langle \cdot \rangle_3$ is the average on \mathbf{y} of any function $f(\mathbf{x}, \mathbf{y}) : \mathbb{R}^{3 \times 2} \rightarrow \mathbb{R}$ over the periodic cell. The first-order corrector χ^{ε_0} now is a third-order tensor. It is periodic in \mathbf{y} and it necessarily belongs to Ψ_3 (the extension of Ψ to 3-D). It is the solution of the following cell problem:

$$\nabla_y \cdot \left\{ \mathbf{C}^{\varepsilon_0} : \left[\mathbf{I} + \frac{1}{2} (\nabla_y \chi^{\varepsilon_0} + {}^t \nabla_y \chi^{\varepsilon_0}) \right] \right\} = \mathbf{0} \quad (23)$$

with periodic boundary conditions. Contrary to the 1-D case, there is no analytical solution for this problem here (unless the medium is layered transverse isotropic, as demonstrated by Guillot *et al.* (2010) and Lin *et al.* (2017)), so we are not able (i) to write $\mathbf{C}^{\varepsilon_0^*}$ as a function of $\mathbf{C}^{\varepsilon_0}$ only and (ii) to get a mathematical condition on $\mathbf{C}^{\varepsilon_0}$ for properly separating the scales within the elastic properties of the medium. To overcome this issue, a procedure inspired by Papanicolaou & Varadhan (1981) is proposed. Introducing $\mathbf{G}^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$ and $\mathbf{H}^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$ such that

$$\mathbf{H}^{\varepsilon_0} = \mathbf{C}^{\varepsilon_0} : \mathbf{G}^{\varepsilon_0} \quad (24)$$

$$= \mathbf{C}^{\varepsilon_0} : \left[\mathbf{I} + \frac{1}{2} (\nabla_y \chi^{\varepsilon_0} + {}^t \nabla_y \chi^{\varepsilon_0}) \right], \quad (25)$$

this procedure enables an implicit construction of $\mathbf{C}^{\varepsilon_0^*}$ such that both $\mathbf{H}^{\varepsilon_0}$ and χ^{ε_0} are periodic in \mathbf{y} and belong to Ψ_3 , which are the conditions to meet to get a proper solution to our homogenization problem. Here are the steps of the procedure:

(i) Eq. (23) with periodic boundary conditions is solved using $\mathbf{C}(\mathbf{y})$ instead of $\mathbf{C}^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$. $\mathbf{C}(\mathbf{y})$ is just the original elastic tensor in which the space variable \mathbf{x} has been changed by $\mathbf{y} = \varepsilon_0 \mathbf{x}$, meaning that all the scales within the medium are considered as small in this first step. This yields a cell problem that we can solve numerically (Part 3). The solution of such a cell problem is called the starting corrector $\chi_s(\mathbf{y})$. It can be seen as the static response of the medium to local unit strains expressed by the identity tensor \mathbf{I} .

(ii) From χ_s , two tensors \mathbf{G}_s and \mathbf{H}_s are built:

$$\mathbf{G}_s(\mathbf{y}) = \mathbf{I} + \frac{1}{2} [\nabla_y \chi_s(\mathbf{y}) + {}^t \nabla_y \chi_s(\mathbf{y})], \quad (26)$$

$$\mathbf{H}_s(\mathbf{y}) = \mathbf{C}(\mathbf{y}) : \mathbf{G}_s(\mathbf{y}). \quad (27)$$

\mathbf{G}_s can be seen as the unit strains plus the strains associated with the starting corrector. It is called *strain concentrator*. In the same way, \mathbf{H}_s is called *stress concentrator*.

(iii) Using $\mathcal{F}_3^{\varepsilon_0}$ (the extension of $\mathcal{F}^{\varepsilon_0}$ to 3-D, see Appendix C), the large-scale and small-scale variations are separated in \mathbf{G}_s and \mathbf{H}_s to form $\mathbf{G}^{\varepsilon_0}$ and $\mathbf{H}^{\varepsilon_0}$:

$$\mathbf{G}^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \mathbf{I} + [\mathbf{G}_s - \mathcal{F}_3^{\varepsilon_0} \{ \mathbf{G}_s \}](\mathbf{y}) : [\mathcal{F}_3^{\varepsilon_0} \{ \mathbf{G}_s \}(\mathbf{x})]^{-1}, \quad (28)$$

$$\mathbf{H}^{\varepsilon_0}(\mathbf{x}, \mathbf{y})$$

$$= [\mathcal{F}_3^{\varepsilon_0} \{ \mathbf{H}_s \}(\mathbf{x}) + [\mathbf{H}_s - \mathcal{F}_3^{\varepsilon_0} \{ \mathbf{H}_s \}(\mathbf{y})] : [\mathcal{F}_3^{\varepsilon_0} \{ \mathbf{G}_s \}(\mathbf{x})]^{-1}. \quad (29)$$

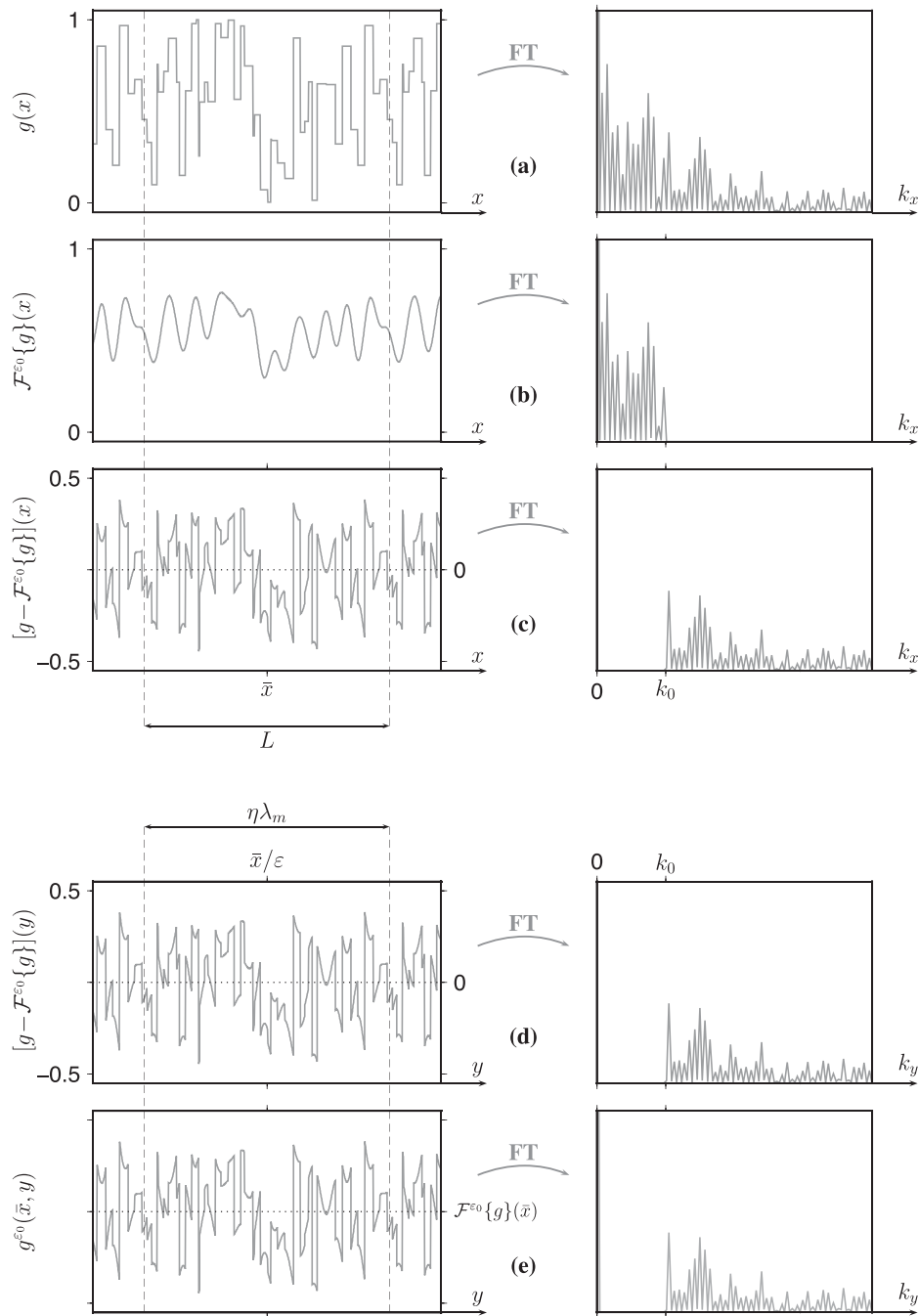


Figure 3. Separation of the scales within a 1-D medium of length L . g here represents either ρ or $\frac{1}{E}$. From any distribution of g extended to \mathbb{R} by periodicity (top row), we build $g^{e_0}(x, y)$ (bottom row) using a low-pass filter \mathcal{F}^{e_0} . In this figure, $g^{e_0}(x, y)$ is represented for a given $x = \bar{x}$. See the text for more details.

Through definitions (24) and (25), \mathbf{G}^{e_0} and \mathbf{H}^{e_0} are the strain and the stress concentrators associated with the first-order corrector χ^{e_0} . They are here constructed by separating the scales in \mathbf{G}_s and \mathbf{H}_s (eqs 28 and 29). From such a construction, one can demonstrate that \mathbf{H}^{e_0} and χ^{e_0} are indeed periodic in y and belong to Ψ_3 (Appendix D).

(iv) Using (25) in (22), we note that $\mathbf{C}^{e_0*} = \langle \mathbf{H}^{e_0} \rangle_3$. Introducing (29) into this latter equation, we end up with

$$\mathbf{C}^{e_0*}(\mathbf{x}) = \mathcal{F}_3^{e_0}\{\mathbf{H}_s\}(\mathbf{x}) : [\mathcal{F}_3^{e_0}\{\mathbf{G}_s\}(\mathbf{x})]^{-1}. \quad (30)$$

The zeroth-order effective elastic tensor given by the procedure therefore is obtained by filtering the stress and the strain associated with the starting corrector. This means that one just needs to go through steps (i), (ii) and (iv) to get this tensor in practice. Step (iii) would be necessary to obtain χ^{e_0} (through the construction of $\mathbf{C}^{e_0} = \mathbf{H}^{e_0} : \mathbf{G}^{e_0^{-1}}$ and the resolution of eq. 23 for instance) in order to access the first-order displacement. Because we focus on the zeroth-order solution in the present paper, we will not deal with this latter aspect. Furthermore, there is no demonstration for the

minor and major symmetries of $\mathbf{C}^{\varepsilon_0^*}$ yet. We do not address this point here, but we have checked the symmetries (up to a certain precision) in all the applications that we present in Part 4.

As regards the density, an equation similar to (B1) also emerges in 3-D, implying that ρ^{ε_0} must lie in Ψ_3 , so the scales can be separated as they were in 1-D:

$$\rho^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \mathcal{F}_3^{\varepsilon_0}\{\rho\}(\mathbf{x}) + [\rho - \mathcal{F}_3^{\varepsilon_0}\{\rho\}](\mathbf{y}). \quad (31)$$

Using (31) in (21), the effective density comes out:

$$\rho^{\varepsilon_0^*} = \mathcal{F}_3^{\varepsilon_0}\{\rho\}. \quad (32)$$

As in the 1-D case, $\rho^{\varepsilon_0^*}$ is obtained by simply filtering the initial density ρ .

3 IMPLEMENTATION OF THE 3-D NON-PERIODIC HOMOGENIZATION

While the non-periodic homogenization theory involves many specific concepts and quantities, its zeroth-order result can be stated quite shortly. From a given 3-D medium described by its density ρ and its elastic tensor \mathbf{C} , a long-wave equivalent density $\rho^{\varepsilon_0^*}$ can be computed by just filtering ρ (eq. 32), and a long-wave elastic tensor $\mathbf{C}^{\varepsilon_0^*}$ can be obtained by

(i) Calculate the starting corrector χ_s from the cell problem (23) using \mathbf{C} with periodic boundary conditions instead of $\mathbf{C}^{\varepsilon_0}$, that is,

$$\nabla_y \cdot \left\{ \mathbf{C} : \left[\mathbf{I} + \frac{1}{2} (\nabla_y \chi_s + {}^t \nabla_y \chi_s) \right] \right\} = \mathbf{0}. \quad (33)$$

(ii) Build \mathbf{G}_s and \mathbf{H}_s from χ_s using eqs (26) and (27), and then filter these two quantities to get $\mathbf{C}^{\varepsilon_0^*}$ (eq. 30).

Differential eq. (33) can be seen as a classic elastostatic problem with a specific load consisting in the divergence of the elastic tensor \mathbf{C} . Such a divergence yields a third-order tensor $\partial_i C_{ijkl}$. Thanks to the symmetry of \mathbf{C} ($C_{ijkl} = C_{jikl}$), this tensor reduces to six force vectors. To determine the six corresponding displacements, eq. (33) can be solved in two different ways: using a strong-form iterative scheme based on Fast Fourier Transform (Moulinec & Suquet 1998; Capdeville *et al.* 2015) or using a more classic weak-form finite-element approach (e.g. Hughes 2012). The latter is adapted to strongly discontinuous media. It is the method we implement here, relying on tetrahedral meshes. Because we want to investigate the behaviour of the solution with respect to the discretization, our code allows polynomial interpolations of degree 1, 2 or 3 (Worth *et al.* 2012) in addition with various quadrature rules (Felippa 2004). Moreover, both linear and quadratic tetrahedra are enabled, leading to either iso-, super- or subparametric elements. The obtained linear system involving six right-hand side members, we solve it using a direct solver. Among several codes, PARDISO (Schenk & Gärtner 2006) is chosen here. Finally, the low-pass filter $\mathcal{F}_3^{\varepsilon_0}$ is applied in the space domain to obtain $\rho^{\varepsilon_0^*}$ and $\mathbf{C}^{\varepsilon_0^*}$. To perform the 3-D convolution, we use the quadrature associated with the mesh employed in finite element analysis.

The periodicity imposed at the boundary $\partial\Omega$ of the elastic domain Ω involved in eq. (33) means that the medium is supposed to repeat itself periodically in the 3-D. When dealing with geological media, this condition is obviously not fulfilled. It is therefore replaced by either a homogeneous Neumann condition or a homogeneous Dirichlet condition. We choose the second option in our implementation so we impose $\chi_s = \mathbf{0}$ on $\partial\Omega$. The effect in the volume Ω of such an artificial condition decays exponentially (Dumontet 1990),

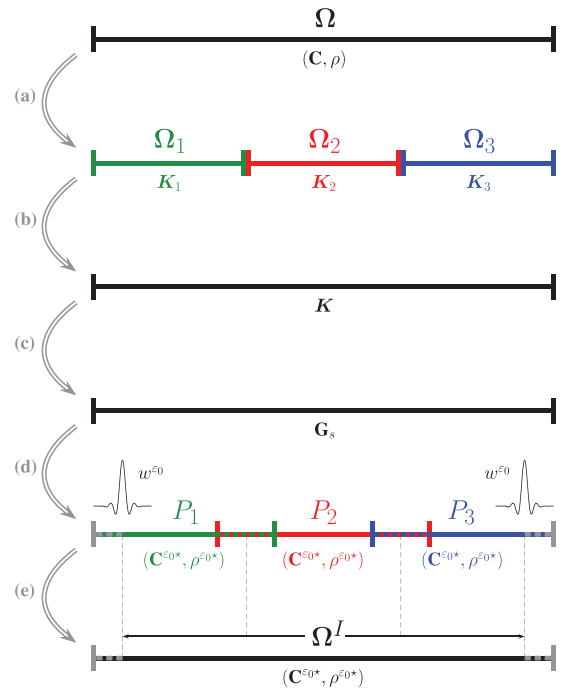


Figure 4. A classic distributed-memory workflow for solving the non-periodic homogenization problem. (a) The domain Ω is decomposed in n subdomains Ω_k . As an example here $n = 3$. In each subdomain, a local stiffness matrix \mathbf{K}_k is computed. (b) The local stiffness matrices are assembled to obtain the global stiffness matrix \mathbf{K} . (c) The linear system is inverted using a parallel solver. From the gradient of the solution, \mathbf{G}_s is formed. (d) Using the domain decomposition again, the effective properties $\mathbf{C}^{\varepsilon_0^*}$ and $\rho^{\varepsilon_0^*}$ are computed in each subdomain by filtering \mathbf{G}_s , $\mathbf{H}_s = \mathbf{C} : \mathbf{G}_s$ and ρ . In the outer domain (dotted grey), the filtering wavelet w^{ε_0} cannot be applied so $\mathbf{C}^{\varepsilon_0^*}$ and $\rho^{\varepsilon_0^*}$ cannot be computed. The same problem appears near the frontiers between the subdomains, so we enlarge these latter to build n overlapping parts P_k equipped with an outer part P_k^O and an inner part P_k^I in which the filter can be applied. (e) The result in the whole inner domain Ω^I is obtained by merging the results from the n inner parts.

so our numerical solution is meaningless in a thin layer from the border of the domain. Such meaningless values of χ_s do not matter, because the filtering process cannot be performed near $\partial\Omega$ anyway. Some elastic material to be convolved with the wavelet is actually missing there, so we are not able to compute the effective properties $\rho^{\varepsilon_0^*}$ and $\mathbf{C}^{\varepsilon_0^*}$ using (32) and (30). The layer Ω^O in which the filter cannot be applied is called the *outer domain*. Its thickness is equal to half of the wavelet support. Its complement $\Omega^I = \Omega - \Omega^O$ is called the *inner domain* (or the *interior of the domain*). The solutions proposed in the present paper make sense in Ω^I only. Further developments, such as those initiated by Capdeville & Marigo (2008, 2013), would be necessary to get meaningful effective properties in Ω^O .

When handling large models, the memory requirements for achieving the computation of the effective properties can be very large. This is mainly because solving large linear systems, even symmetric, is memory-demanding. This is also because high-order tensors are involved in the homogenization process. For these two reasons, a distributed-memory computation is necessary. A classic way of implementing it is presented in Fig. 4. It consists in

(a) Decomposing the domain Ω into n subdomains Ω_k such that $\Omega = \bigcup_{k=1}^n \Omega_k$. Each subdomain being handled by a processor, n local stiffness matrices \mathbf{K}_k are computed.

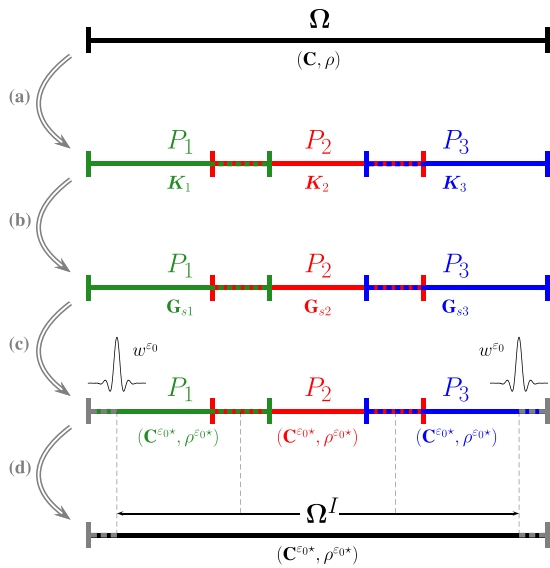


Figure 5. An alternative way of solving the non-periodic homogenization in parallel. (a) The domain Ω is decomposed in n parts P_k . As an example here $n = 3$. The parts overlap to equip them all with an outer part P_k^O and an inner part P_k^I . Moreover, a homogeneous Dirichlet condition is applied at the boundary of the parts to make each of them like a new domain in which a linear system with a stiffness matrix K_k can be defined. (b) Solving the linear system in each part leads to n tensors $\mathbf{G}_{s,k}$. (c) Filtering $\mathbf{G}_{s,k}$, $\mathbf{H}_s = \mathbf{C} : \mathbf{G}_{s,k}$ and ρ , the effective properties are obtained in each inner part P_k^I . (d) The results from the n inner parts are gathered to obtain the effective properties in the whole inner domain Ω^I .

(b) Assembling the local stiffness matrices to build the global stiffness matrix \mathbf{K} .

(c) Solving the linear system using a parallel solver to get the finite element solution χ_s of eq. (33). From χ_s , \mathbf{G}_s is formed following (26).

(d) Using the domain decomposition again to filter \mathbf{G}_s , $\mathbf{H}_s = \mathbf{C} : \mathbf{G}_s$ and ρ in each subdomain. From this filtering, the effective properties $\mathbf{C}^{\varepsilon_0^*}$ and $\rho^{\varepsilon_0^*}$ are derived. In the outer domain Ω^O (the dotted grey lines in Fig. 4), these properties cannot be computed because the filter cannot be applied. The same problem appears near the frontiers between the subdomains, so we have to either implement massive communications between the subdomains or enlarge each subdomain Ω_k by a buffer layer Ω_k^B which is as thick as half of the filtering wavelet support. This second option is sketched in Fig. 4. It yields n overlapping parts $P_k = \Omega_k \cup \Omega_k^B$. Along with possible pieces of Ω^O in Ω_k , the buffer layer Ω_k^B acts as an outer part $P_k^O = (\Omega^O \cap \Omega_k) \cup \Omega_k^B$. The effective properties can be calculated in the inner part $P_k^I = P_k - P_k^O$, so in the whole inner domain because $\bigcup_{k=1}^n P_k^I = \Omega^I$.

(e) Communicating between processors to gather the results from the n subdomains.

Such a workflow would certainly work, but it imposes the resolution of the whole linear system to fit the distributed-memory at once, meaning that a large stiffness matrix requires a large parallel computer. To overcome this limitation, we propose an alternative implementation which consists in working on the parts P_k all along the workflow (Fig. 5). Applying a homogeneous Dirichlet condition $\chi_s = \mathbf{0}$ at the boundary of the parts, each of them becomes like a new domain which is totally independent from the other parts. Thanks to the outer parts P_k^O , we get the effective properties in all the inner parts P_k^I , so in the whole inner domain $\Omega^I = \bigcup_{k=1}^n P_k^I$.

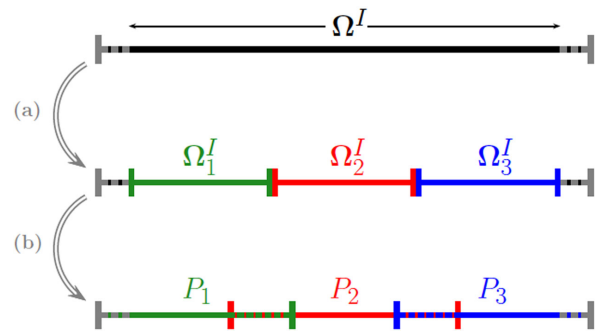


Figure 6. Strategy for building well-balanced parts. (a) We partition Ω^I instead of the whole domain Ω , then forming n subdomains Ω_k^I . As an example here $n = 3$. (b) Each subdomain is enlarged by a buffer layer Ω_k^B . At the border of the domain, the buffer is the outer domain Ω^O (dotted grey). We end up with n parts $P_k = \Omega_k^I \cup \Omega_k^B$ which all have the same size. Moreover, $P_k^I = \Omega_k^I$ and $P_k^O = \Omega_k^B$.

This alternative implementation is embarrassingly parallel because the parts are independent from each other all along the workflow, which means that they can be treated sequentially. This allows for the homogenization of large models on small computers, possibly laptops, provided the available RAM covers the memory requirement of every single part. Such a distributed-memory implementation is also proposed by Capdeville *et al.* (2015).

To speed up our code, multithreaded computations are performed whenever possible. Moreover, efficient algorithms based on k - d trees and stack data structures (e.g. Cormen *et al.* 2009) are used to search for elements or points across the finite element mesh. Last but not least, the partitioning of the mesh is performed on Ω^I instead of Ω to obtain well-balanced parts P_k (Fig. 6). The performance and capabilities of the code are illustrated through examples in the next part.

4 VALIDATION TESTS

We here handle three different models to test the accuracy of the homogenization method in 3-D and challenge our code. First, the case of fine layers is investigated. For such a medium, analytical expressions of the effective elastic parameters exist so we can just compare the result of the homogenization with the result of these expressions to validate our code. The second model we study is made of small cubes. Because no reference solution for the effective properties is available in this case, we base our validation on the comparison of waveforms computed in the initial medium on the one hand, and in the homogenized medium on the other hand. Finally, a 3-D geological model is handled to emphasize the efficiency of the homogenization in a realistic case.

4.1 Homogenization of a finely layered medium

The first model we consider for testing our 3-D homogenization code is a medium made of 60 isotropic layers randomly taken between 800 and 1200 m thick. Within each layer, the density is randomly chosen in the 2000–4000 kg m⁻³ range. Because the medium is isotropic, two parameters (e.g. the Lamé coefficients, or the S - and P -wave velocities) are sufficient to define the elastic tensor in each layer. We then randomly choose the S -wave velocity between 3000 and 5000 m s⁻¹, and the P -wave velocity between 5000 and 8000 m s⁻¹ (Fig. 7), with the constraint of having the Poisson's ratio in the 0.1–0.45 range to make it geologically realistic.

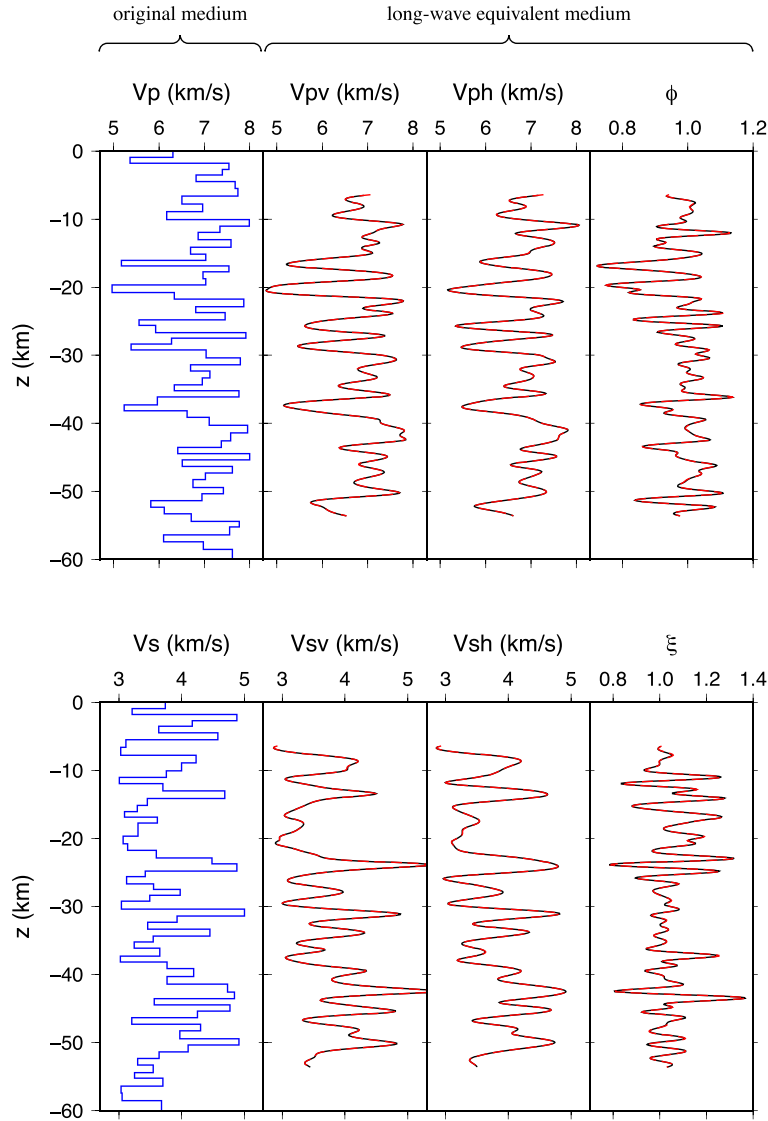


Figure 7. Comparison of effective properties computed with the 3-D non-periodic homogenization (dashed red) and the Backus theory (black) in the case of an original layered medium (blue). The two equivalent media both are anisotropic ($\phi = \frac{V_{PV}^2}{V_{PH}^2}$ and $\xi = \frac{V_{SH}^2}{V_{SV}^2}$ pointing out the rate of anisotropy) and agree very well with each other.

The homogenization of the model is performed using $\lambda_0 = 1600$ m. In this case, the thickness of the outer domain (i.e. half of the filtering wavelet support) is 6400 m (Appendix C), so we take the extent of the layers equal to 15 km in order to get an inner domain in which a solution can be computed. Following a rule of thumb for the spatial discretization (Appendix E), each layer is meshed by a single layer of tetrahedral elements equipped with interpolation functions of degree 3. As expected from Backus (1962), the resulting effective medium is transversely isotropic. In Fig. 7, we represent it in terms of vertically- and horizontally polarized wave velocity:

$$V_{PV} = \sqrt{\frac{C_{zzzz}^{\epsilon_0*}}{\rho^{\epsilon_0*}}},$$

$$V_{PH} = \sqrt{\frac{C_{xxxx}^{\epsilon_0*}}{\rho^{\epsilon_0*}}} = \sqrt{\frac{C_{yyyy}^{\epsilon_0*}}{\rho^{\epsilon_0*}}},$$

$$V_{SV} = \sqrt{\frac{C_{xzxz}^{\epsilon_0*}}{\rho^{\epsilon_0*}}} = \sqrt{\frac{C_{yzyz}^{\epsilon_0*}}{\rho^{\epsilon_0*}}},$$

$$V_{SH} = \sqrt{\frac{C_{xyxy}^{\epsilon_0*}}{\rho^{\epsilon_0*}}},$$

z being the vertical, layered direction, and (x, y) defining the horizontal plane. Moreover, we emphasize the amount of anisotropy by plotting $\phi = \frac{V_{PV}^2}{V_{PH}^2}$ and $\xi = \frac{V_{SH}^2}{V_{SV}^2}$.

The effective properties computed with the non-periodic homogenization in the layered case are expected to be the same as those proposed by Backus (1962). A comparison between the two solutions (Fig. 7) shows a very good match, meaning that our 3-D implementation of the non-periodic homogenization works properly in the present case. As mentioned in Part 3, no solution can be computed at the border of the domain because the convolution

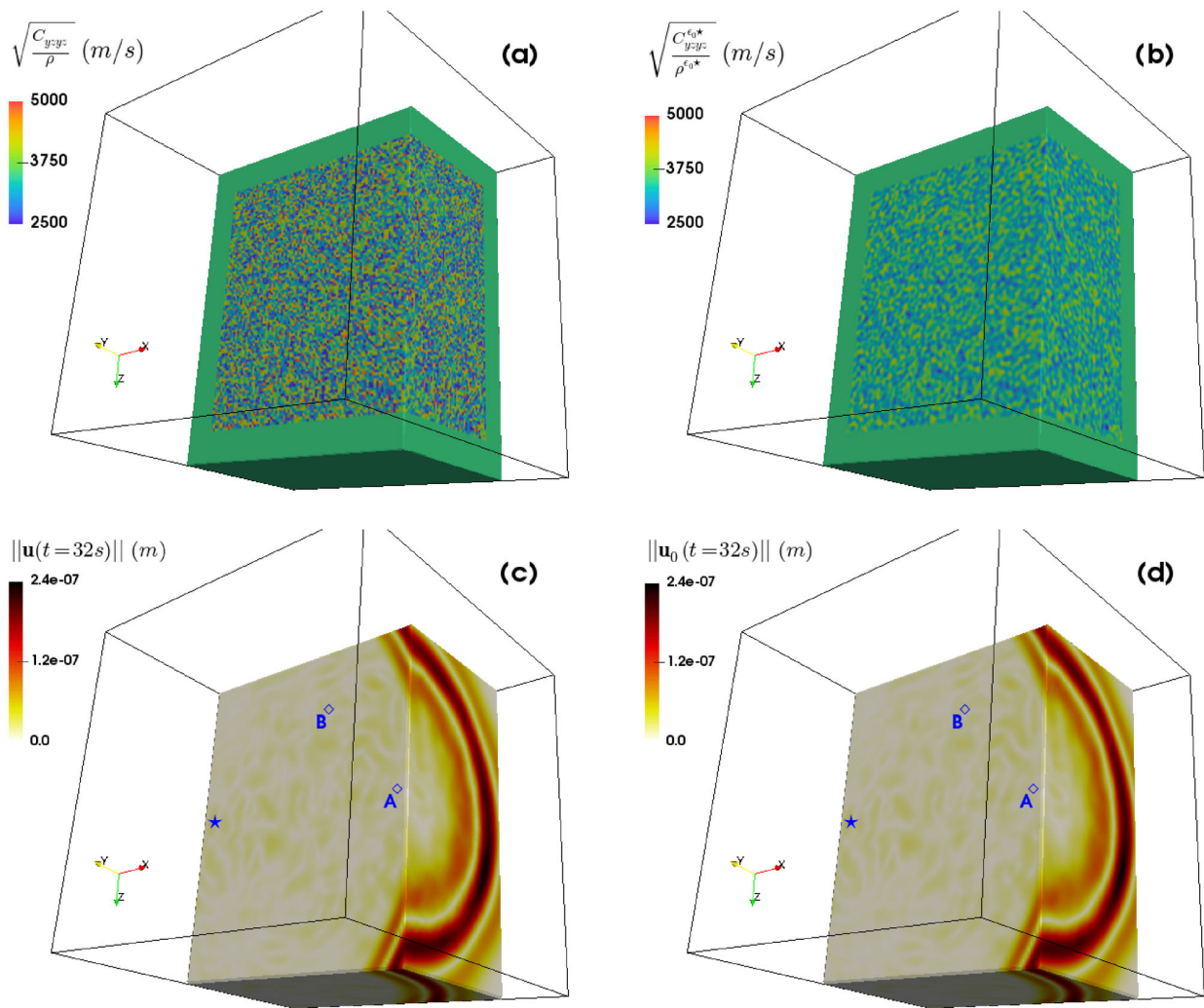


Figure 8. (a) Cut in the random cubes. The black lines emphasize the edges of the whole domain. (b) Cut in the homogenized medium. (c) Snapshot of the L^2 norm of the wavefield \mathbf{u} generated through the random cubes by a force along the z -axis at point x_S marked by the blue star. (d) Snapshot of the L^2 norm of the wavefield \mathbf{u}_0 generated through the homogenized medium by the exact same force.

involved in the filtering operation is not possible there. Nonetheless, extending the domain with a relevant material to make the convolution possible and to get satisfying effective properties at the border would be possible. Simply extending the boundary values of density and elastic coefficients would lead to the zeroth-order solution (Capdeville & Marigo 2007). To reach the first order, a continuous periodic extension (in which the boundary actually acts as a mirror) would have to be used (Leptev 2005; Capdeville & Marigo 2007). We do not implement such extensions in the present work.

4.2 Homogenization of random cubes

To push our validation further, we challenge our homogenization code to a highly heterogeneous medium made of small elastic cubes with random isotropic properties. Each cube is 1 km^3 large. 100 cubes are considered in each direction, which gives rise to a large cubic volume made of 1 000 000 random cubes. As shown in Fig. 8(a), this cubic volume is embedded in a 13 km thick homogeneous medium. Such a thickness corresponds to the support of the

filtering wavelet which will be used in the homogenization process. In each small cube, the properties are randomly picked between 2000 and 4000 kg m^{-3} for the density, 2500 and 5000 m s^{-1} for the S -wave velocity, and 4000 and 8000 m s^{-1} for the P -wave velocity. As in the previous example, the Poisson's ratio is constrained in the 0.1–0.45 range.

The homogenization of the random cubes is performed using $\lambda_0 = 1600 \text{ m}$ (i.e. $\lambda_m = 8000 \text{ m}$ and $\varepsilon_0 = 0.2$). The spatial discretization is ensured by dividing every single cube into six tetrahedra equipped with degree 3 interpolants. Using similar elements in the homogeneous region, we end up with 12 002 256 tetrahedra and 160 747 899 degrees of freedom (three components at 53 582 633 interpolation points). To achieve the computation, the domain is split into 100 overlapping subdomains. Calculating the effective properties in a single subdomain then requires about 116 GB and 4 hr on an Intel Xeon X5680 processor (6 cores, 3.33 GHz, 12 MB Cache, 6.4 GT s^{-1}). The obtained homogenized model is shown in Fig. 8(b).

To assess the quality of the effective medium calculated with the non-periodic homogenization technique, we perform a seismic

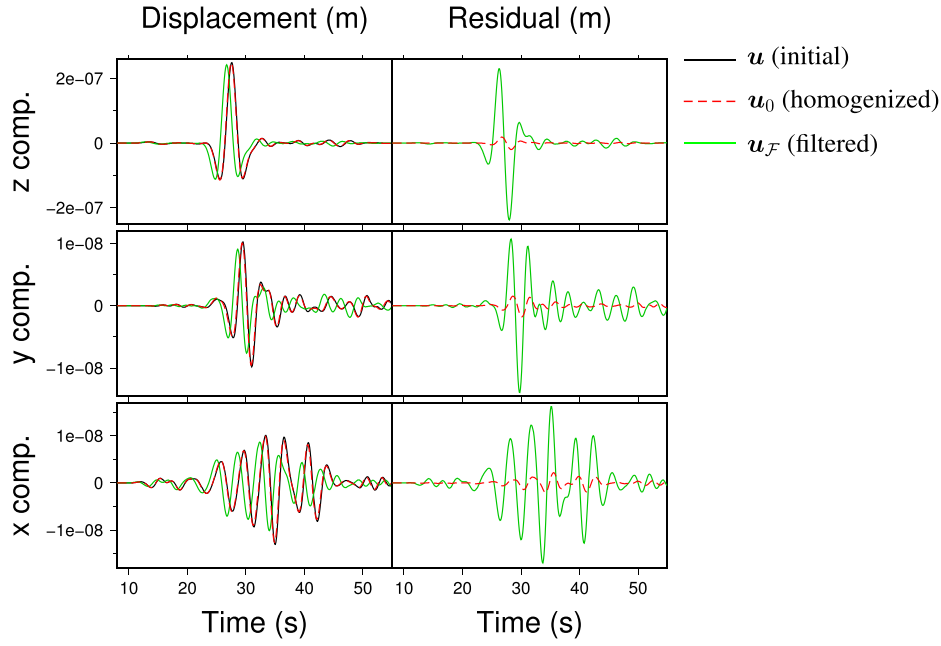


Figure 9. Comparison of the wavefields \mathbf{u} (black), \mathbf{u}_0 (dashed red) and \mathbf{u}_F (green) computed at receiver A (Figs 8c and d). The maximum and dominant frequencies of the Ricker function $R(t)$ emitted at the source are equal to 0.5 and 0.2 Hz, respectively.

wave simulation in it and we compare the obtained seismograms with traces computed in the initial model (i.e. the random cubes). In other words, we solve the initial problem

$$\begin{cases} \rho \ddot{\mathbf{u}} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{f} \\ \boldsymbol{\sigma} = \mathbf{C} : [\frac{1}{2}(\nabla \mathbf{u} + {}^t \nabla \mathbf{u})] \end{cases} \quad (34)$$

and the homogenized problem

$$\begin{cases} \rho^{\varepsilon_0*} \ddot{\mathbf{u}}_0 - \nabla \cdot \boldsymbol{\sigma}_0 = \mathbf{f} \\ \boldsymbol{\sigma}_0 = \mathbf{C}^{\varepsilon_0*} : [\frac{1}{2}(\nabla \mathbf{u}_0 + {}^t \nabla \mathbf{u}_0)] \end{cases} \quad (35)$$

to estimate the quality of ρ^{ε_0*} and $\mathbf{C}^{\varepsilon_0*}$ through the comparison of \mathbf{u}_0 to \mathbf{u} . In eq. (34) and (35), $\ddot{\mathbf{u}}$ represents the second time-derivative and \mathbf{f} is the external force. This latter is chosen as a simple Ricker function $R(t)$ along the z -axis at a given point \mathbf{x}_S in the homogeneous region: $\mathbf{f}(\mathbf{x}, t) = R(t) \delta(\mathbf{x} - \mathbf{x}_S) \mathbf{e}_z$. The dominant frequencies of $R(t)$ is chosen to be equal to 0.2 Hz. The two simulations are performed using a spectral element method with PML-type absorbing boundaries (Cupillard *et al.* 2012). Snapshots of the two obtained wavefields are shown in Figs 8(c) and (d). They look very similar, suggesting that our homogenized model is an accurate equivalent medium for the seismic wave propagation.

We carefully compare \mathbf{u}_0 to \mathbf{u} by looking at seismograms calculated at two particular receivers denoted by A and B in Figs 8(c) and (d). Receiver A is on a P -wave nodal plane, 92 km away from the source. The ballistic S -wave, which contains most of the seismic energy, appears on the z -component (Fig. 9). We see that this wave is very well-retrieved by the homogenized model, the difference between \mathbf{u}_0 and \mathbf{u} (i.e. the residual) reaching 8 per cent at most. Apart from the S -wave, a scattered wavefield is observed on the three components. Our homogenized model also reconstructs this wavefield very well. To emphasize the relevance of the homogenized solution, Fig. 9 also shows seismograms computed in a medium obtained by just filtering the initial density and elastic tensor. The displacement

and the stress propagating in this medium are denoted by \mathbf{u}_F and $\boldsymbol{\sigma}_F$, respectively. By definition, these two fields verify

$$\begin{cases} \mathcal{F}_3^{\varepsilon_0} \{\rho\} \ddot{\mathbf{u}}_F - \nabla \cdot \boldsymbol{\sigma}_F = \mathbf{f} \\ \boldsymbol{\sigma}_F = \mathcal{F}_3^{\varepsilon_0} \{\mathbf{C}\} : [\frac{1}{2}(\nabla \mathbf{u}_F + {}^t \nabla \mathbf{u}_F)] \end{cases} \quad (36)$$

The waveforms of \mathbf{u}_F at receiver A do not fit the wavefield \mathbf{u} in the initial model. Neither the phase nor the amplitude is properly reconstructed by the filtered medium, meaning that this latter does not hold the correct effective properties for the seismic wave propagation. Similar features are observed at receiver B (Fig. 10): \mathbf{u}_F is far from \mathbf{u} whereas the homogenized solution accurately recovers the whole seismograms, including a ballistic P -wave which appears at this receiver location.

The overall difference between \mathbf{u}_0 and \mathbf{u} can be evaluated quantitatively using the error

$$E_0 = \frac{1}{50} \sum_{r=1}^{50} \sqrt{\frac{\int_0^T (\mathbf{u}_0 - \mathbf{u})^2(\mathbf{x}_r) dt}{\int_0^T \mathbf{u}^2(\mathbf{x}_r) dt}}, \quad (37)$$

where \mathbf{x}_r is the randomly chosen location of receiver r . In eq. (37), T is the duration of the simulated propagation. It is equal to 55 s here. For $\varepsilon_0 = 0.2$ (i.e. $\lambda_0 = 1600$ m), we obtain $E_0 \simeq 0.006$. Computing the same kind of error for the naive solution \mathbf{u}_F , we get $E_F \simeq 0.055 \simeq 9E_0$. This result again emphasizes the much higher accuracy of the homogenized solution \mathbf{u}_0 . We also compute E_0 and E_F for $\varepsilon_0 = 0.4$ (i.e. $\lambda_0 = 3200$ m), $\varepsilon_0 = 0.8$ (i.e. $\lambda_0 = 6400$ m) and $\varepsilon_0 = 1.6$ (i.e. $\lambda_0 = 12800$ m). This allows us to study how fast \mathbf{u}_0 and \mathbf{u}_F converge to \mathbf{u} . From the homogenization theory, we expect $E_0 = O(\varepsilon_0)$. Fig. 11 shows that we actually get $E_0 \simeq O(\varepsilon_0^{3/2})$. This unexpectedly fast convergence of \mathbf{u}_0 toward \mathbf{u} could be due to the weak contribution of the higher-order terms in the present case study, as suggested by Capdeville *et al.* (2010b) in 2-D. Plotting also E_F in Fig. 11, it appears that the convergence of \mathbf{u}_F is way poorer.

The spectral element simulation of the wavefield \mathbf{u} in the initial model requires 126 hexahedra in each direction, that is, 2 000 376

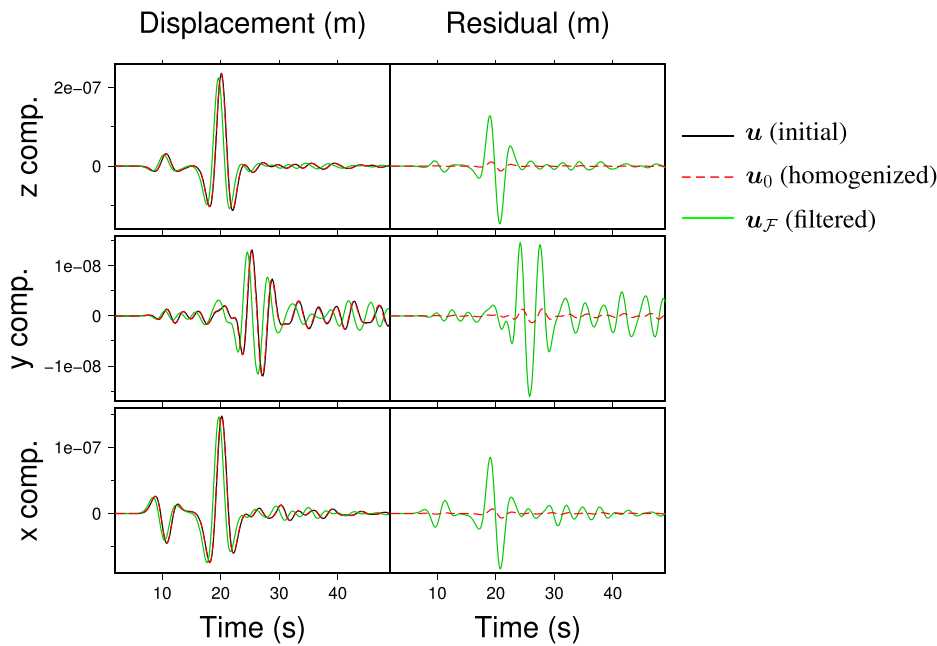


Figure 10. Comparison of the wavefields \mathbf{u} (black), \mathbf{u}_0 (dashed red) and $\mathbf{u}_{\mathcal{F}}$ (green) computed at receiver B (Figs 8c and d). Contrary to receiver A, B records a ballistic P -wave.

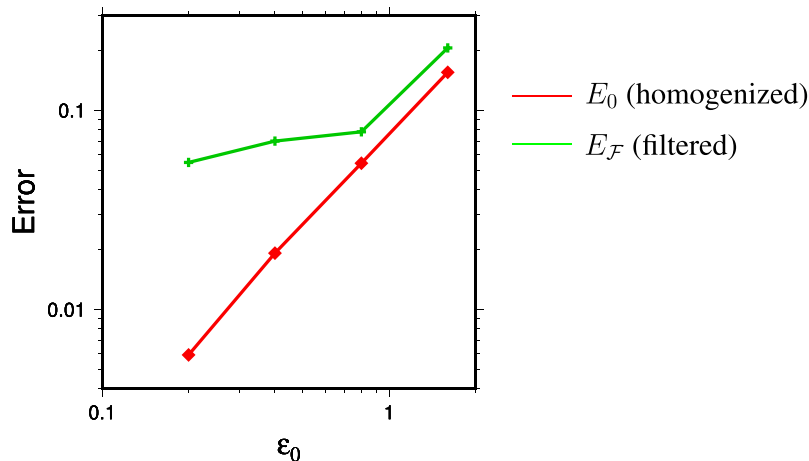


Figure 11. E_0 (red) and $E_{\mathcal{F}}$ (green) as a function of ε_0 . It is clear from this plot that \mathbf{u}_0 converges much faster than $\mathbf{u}_{\mathcal{F}}$ toward the target wavefield \mathbf{u} .

elements in total. This is because a spectral element mesh has to honour all the physical discontinuities of the model in study to make the computation accurate, so each $1 \text{ km} \times 1 \text{ km} \times 1 \text{ km}$ random cube has to be captured by a hexahedron. As a consequence, the obtained mesh highly oversamples the wavefield: because $\lambda_m = 8 \text{ km}$, elements as large as $8 \text{ km} \times 8 \text{ km} \times 8 \text{ km}$ with a polynomial order larger than 4 would be sufficient if the medium was smooth enough (i.e. if it only contained scales larger than λ_m). Because of the discontinuities, we here get a 512 times denser mesh, which yields a 4096 times higher numerical cost (a factor 512 in space times a factor 8 in time because of eq. 1). Computing a 55 s long simulation of \mathbf{u} then takes about 6 hr on ten Intel Xeon X5680 processors. In homogenized media, such as those computed here using $\varepsilon_0 = 0.2, 0.4, 0.8$ and 1.6 , coarser spectral element meshes can be used, then decreasing the computation cost. When using $\varepsilon_0 = 0.4$ for instance, we get an effective medium of the random cubes that only contains scales larger than $\lambda_0 = 3.2 \text{ km}$, so 3.2 km^3 large hexahedra then suit

an accurate spectral element simulation of \mathbf{u}_0 , which is 105 times less numerically demanding than the simulation of \mathbf{u} . Nevertheless, all the wave simulations presented in this paper are performed within the same fine hexahedral mesh to avoid possible numerical biases and thus make all our comparisons totally fair.

4.3 Homogenization of a realistic geological model

In this section, we use a subsurface model of the Ribaut area in France (Caumon *et al.* 2009) as an illustration of the 3-D non-periodic homogenization technique applied to a realistic geological medium (Fig. 12a). Simulating waves in the model as it is (i.e. composed of multiple faulted and folded horizons) is extremely challenging because of the fine grid required to accommodate thin layers and complex geometries formed by the discontinuities. Such a fine grid indeed makes the computation cost of any wave simulation enormous. Moreover, in the context of the spectral element method,

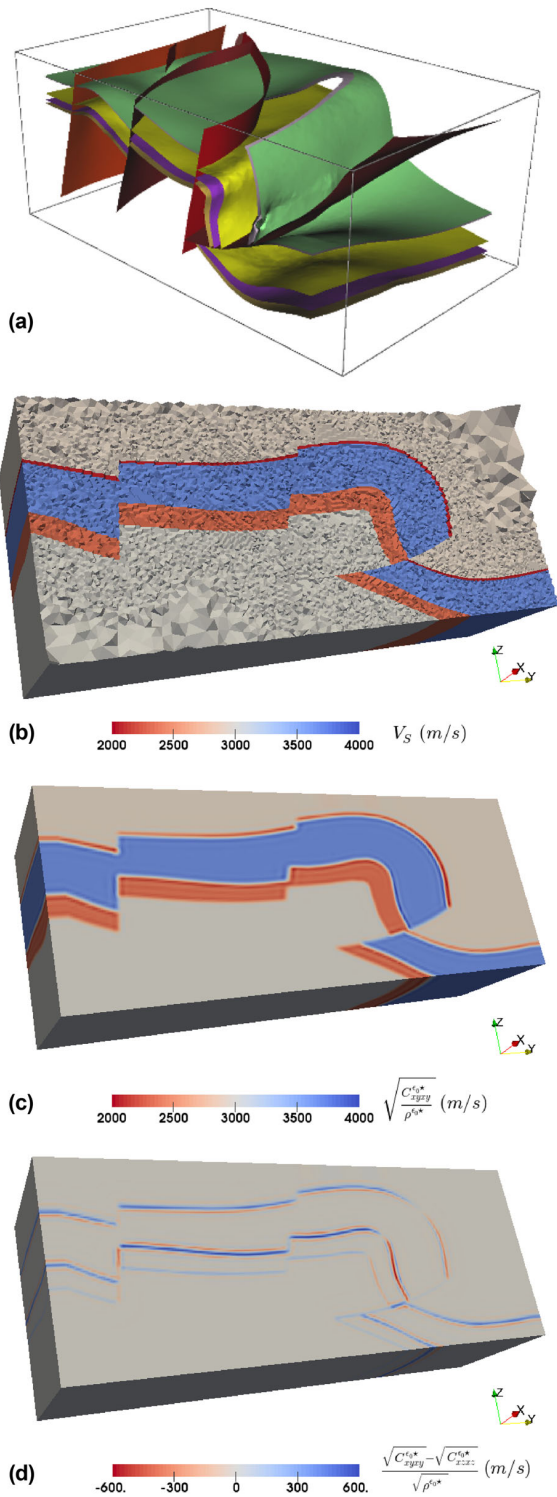


Figure 12. (a) Structural model of a highly faulted and folded region near Ribaute, Southern France. (b) Adaptive tetrahedral mesh of the model. This mesh is used to perform the homogenization. The background colour corresponds to the (randomly chosen) isotropic S -wave velocity within each layer of the model. (c) One of the S -wave velocities resulting from the homogenization. Another S -wave velocity is shown in Fig. 13. The difference between these two velocities is plotted in (d), which emphasizes the seismic anisotropy produced by the structure.

generating a hexahedral mesh which honours all the horizons and faults, if possible, demands tremendous efforts. For these reasons, we do not compute reference waveforms here; we just perform a spectral element simulation in the homogenized model to show how convenient working with effective properties is, and we compare the results with a simulation in a model obtained by brutally filtering the initial density and elastic tensor to exemplify that the choice of the upscaling technique matters.

The initial model is made of five homogeneous isotropic layers (Fig. 12b). To compute its effective properties using our finite-element code, it is meshed with VortexLib from RINGMesh (Pellerin *et al.* 2017). We choose $\lambda_0 = 30$ m and polynomial interpolants of degree 2, so the optimal volume for the tetrahedra is about 200 m^3 (Appendix E). While smaller elements are needed around the discontinuities for capturing their complex geometry, larger elements are allowed where the discontinuities have no influence, that is, at a distance larger than the size of the filtering wavelet support. We therefore take advantage of the refinement-derefinement technique proposed by Botella (2016) to generate an adaptive mesh which minimizes the finite-element computation cost. We end up with a 9 077 300 element mesh and 35 703 579 degrees of freedom. The homogenized model is obtained in 9 min on 120 Intel Xeon E5-2683 v4 (16 cores, 2.10 GHz, 40 MB Cache, 9.6 GT s^{-1}). The

result is shown in Fig. 12(c). We have plotted the $\sqrt{\frac{C_{xyxy}^{e0*}}{\rho^{e0*}}}$ component there to make the figure comparable to Fig. 12(b) where the isotropic S -wave velocity of the initial medium is represented. As expected, the homogeneous areas (i.e. the interior of the layers) are not changed by the homogenization process, whereas all the discontinuities (i.e. the faults and the interfaces between the layers) are smoothed. To illustrate the S -anisotropy produced by these discontinuities, $\frac{\sqrt{C_{xyxy}^{e0*}} - \sqrt{C_{xzxz}^{e0*}}}{\sqrt{\rho^{e0*}}}$ is plotted in Fig. 12(d). We observe that the anisotropy can reach 20 per cent where high velocity contrasts (i.e. at the bottom and the top of the fastest layer) occur.

Because the homogenized model is smooth, seismic waves within it can be simulated using a coarse mesh. As an example, a coarse hexahedral mesh supporting a spectral element simulation (Cupillard *et al.* 2012) is shown in Fig. 13. The elements there are 60 m^3 large and hold degree-8 polynomials to capture all the variations of the model. We see that the direct seismic wave front is highly deformed by the effective structure and that strong reflected and diffracted waves are generated, even though the medium contains no discontinuities. Fig. 13 also shows waveforms obtained at two stations. At station A, near the source, several S -waves reflected on discontinuities are observed. Zooming in two of them emphasizes important discrepancies between our homogenized solution and the waveform computed in a medium obtained by just filtering the density and the elastic tensor of the initial model. These discrepancies can be explained by the lack of anisotropy in this last medium. By construction, it is fully isotropic, so it cannot hold the important structure-induced anisotropy observed in our homogenized model (e.g. Fig. 12d). As a consequence, the waveforms computed in the two models at station A show significant differences in phase and amplitude. Similar observations are also made far from the source, at station B. In the light of the results presented in Section 4.2 and in previous 2-D studies (Capdeville *et al.* 2010b; Guillot *et al.* 2010), we can reasonably think that the waveforms computed in the homogenized model are the most accurate, but we cannot assess it because we do not provide any reference solution here. The main point of this section is to show an application of our homogenization

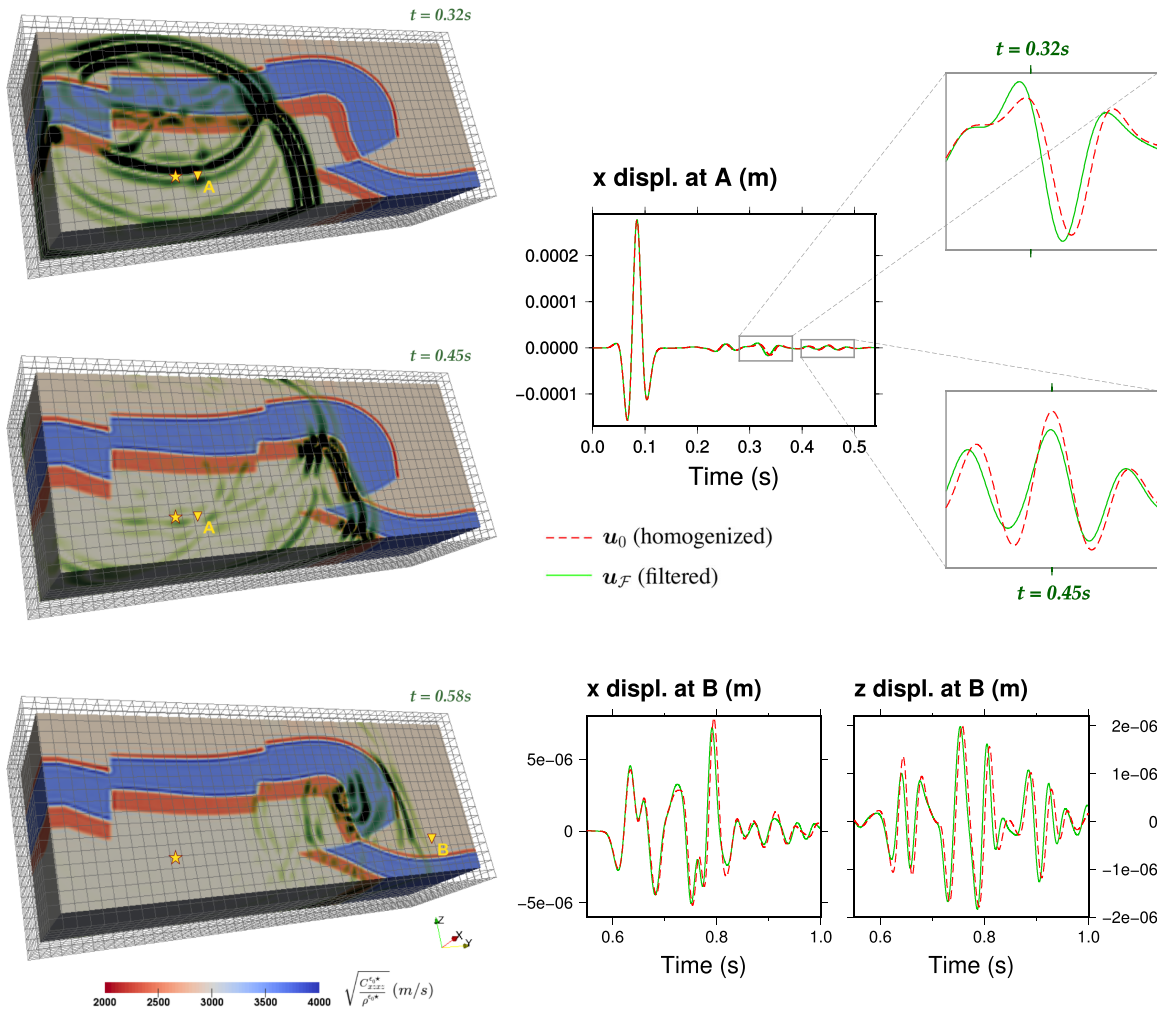


Figure 13. Left: three snapshots of the L^2 norm of the wavefield u_0 generated through the homogenized Ribaute model by a force along the x -axis acting at the point marked by the yellow star. The wave simulation is performed using a spectral element method on a regular hexahedral mesh. Top right: x -component of the displacement u_0 recorded near the source, at station A. The displacement $u_{\mathcal{F}}$ computed in a medium obtained by just filtering the density and the elastic tensor of the initial model is also plotted. As expected, the direct S -wave (arriving around $t = 0.05$ s) is identical in the two simulations. On the contrary, the reflected phases show significant differences. In particular, zooming in the waves reflected on the two major velocity contrasts (i.e. the bottom and the top of the fastest layer) shows that $u_{\mathcal{F}}$ arrives earlier than u_0 with a different amplitude. Bottom right: x - and z -components of both u_0 and $u_{\mathcal{F}}$ recorded far from the source, at station B. Because the direct S -wave has travelled through heterogeneities, it is now different in the two simulations. Obviously, discrepancies in the later, multiply reflected and scattered energy, are also observed.

code to a 3-D realistic geological model and to put the importance of the structure-induced anisotropy in evidence.

5 DISCUSSION AND CONCLUSIONS

Dealing with small-scale heterogeneities in seismic wave simulation is a difficult task because it usually involves enormous computation costs. To handle them, Graphics Processing Unit (GPU) and/or High Performance Computing (HPC) implementations of well-established numerical techniques such as the Spectral Element method (SEM), the Discontinuous Galerkin method (DGM) and the Finite Difference method (FDM), have been proposed (Komatitsch *et al.* 2010; Peter *et al.* 2011; Weiss & Shragge 2013; Gokhberg & Fichtner 2016; Remacle *et al.*

2016; Rietmann *et al.* 2017). Furthermore, the DGM allows local time-stepping and p -adaptivity (e.g. Dumbser *et al.* 2007; Etienne *et al.* 2010; Minisini *et al.* 2013; Diaz & Grote 2015) which mitigate stability constraint (1) and therefore reduce the overall computation cost. In the context of the SEM and the FDM, Pelties *et al.* (2011) proposed empirical laws on the mesh size to relax the need of honouring geological discontinuities. Complementary to all these numerical advances, the use of effective properties for the seismic wave propagation can drastically reduce the computation cost related to small-scale features while preserving a good accuracy. In the recent years, the non-periodic homogenization method emerged as a general technique to compute such effective properties. We here applied it in 3-D for the first time. In part 2, we recalled the theory of the homogenization, skipping some technical details to focus on the main ideas and concepts. Then we described an efficient,

embarrassingly parallel implementation of the method. In part 4, we challenged this implementation on various media, showing the high accuracy of the non-periodic homogenization and the ability of our code to handle large and complex 3-D models, with no restriction on the size and shape of the heterogeneities. The code is available at upon request.

Homogenizing complex geological media to ease the numerical simulation of full seismic wavefields has straight-forward applications in seismic risk assessment (e.g. Chaljub *et al.* 2010), survey design (e.g. Wei *et al.* 2012), structural model validation (e.g. Irakarama *et al.* 2017), seismic source characterization (e.g. Silwal & Tape 2016) and seismic tomography for keeping subwavelength geological details in the inversion (Fichtner *et al.* 2013b; Capdeville & Cance 2015). Because it tells what the waves ‘see’, the non-periodic homogenization also opens important perspectives in the interpretation of full waveform inversion results. In particular, the structure-induced (or, equivalently, ‘extrinsic’ or ‘apparent’ or ‘geometric’) anisotropy (Fichtner *et al.* 2013a; Wang *et al.* 2013) and the structure-induced attenuation (Ricard *et al.* 2014) could be downscaled, as recently initiated by Bodin *et al.* (2015), leading to probability distributions of small-scale features and to better estimations of the intrinsic anisotropy and attenuation at various scales, from the Earth’s core up to the subsurface. In addition, Capdeville *et al.* (2013) and Afanasiev *et al.* (2016) recently showed that the homogenization can help in regularizing full waveform inversion problems.

The non-periodic homogenization relies on a two-scale asymptotic expansion of the displacement and the stress involved in the elastic wave equation. In this paper, we have investigated the zeroth-order term and the associated upscaled properties. As shown by Capdeville *et al.* (2010a,b) and Guillot *et al.* (2010), adding the first-order term allows retrieving site effects (i.e. small-amplitude high-frequency non-propagating signals) at the receivers. These authors also show that a correction can be applied to the external force term to take into account the effects of the local small-scale structure at the source. A recent application of this correction can be found in Burgos *et al.* (2016).

We have claimed that our method does not require any constraint on the size and shape of the heterogeneities to be smoothed. In other words, the homogenization is able to upscale any media. The only limitation that has been noted so far occurs when trying to model subwavelength focusing in Helmholtz resonators. In this extreme case, the wavefield no longer holds a minimum wavelength, so the non-periodic homogenization fails (Zhao *et al.* 2016). Moreover, we have mentioned that our homogenization theory is not able to compute long-wave equivalent properties of heterogeneities laying near Neumann or Dirichlet surfaces yet. A proper solution only exists for smoothing a free-surface topography on top of a homogeneous material (Capdeville & Marigo 2013) or fine horizontal layers below a flat free-surface (Capdeville & Marigo 2008). Relevant extensions of the model beyond its boundaries can also lead to accurate effective properties (Leptev 2005; Capdeville & Marigo 2007). Pushing these approaches would surely help the non-periodic homogenization theory in handling heterogeneities near various boundary shapes and conditions.

Besides these physical and theoretical limitations, slight numerical weaknesses can be pointed out in our code. Even if it employs powerful algorithms and an efficient parallel scheme, it could be optimized using Basic Linear Algebra Subprograms and the Fortran column-major order more extensively, particularly when computing the stiffness matrix and when filtering the strain and the stress associated with the starting corrector. The performance of the code

could also benefit from the nested homogenization technique proposed by Capdeville *et al.* (2015) and from an iterative solver which would be more efficient than a direct solver when dealing with very large stiffness matrices. Most of all, developing an adaptive homogenization, taking into account the fact that the minimum wavelength usually varies within the medium, would be a major advance.

ACKNOWLEDGEMENTS

We deeply thank Q. Liu and C. Boehm for their sharp reviews and constructive comments that significantly improve the initial manuscript. We also thank A. Fichtner, L. Guillot, P. Cance, Y. Ricard, F. Chambat, J.-P. Montagner, J. Virieux, A. Kutsenko, Y. Masson, G. Caumon and the European COST action TIDES (ES1401) for emulative remarks and discussions. We are also grateful to A. Botella for advising some powerful search algorithms and for providing 3-D adaptive meshes which significantly reduced the computation cost of the homogenization. Many thanks to A. Borghi as well, who completely reset the Cauchy cluster at GeoRessources. Part of the calculations has also been led on the S-CAPAD cluster at IPGP, on the Centre de Calcul Intensif des Pays de la Loire (CCIPL) computers, and on the EXPLOR centre hosted by the University of Lorraine. Financial supports from the RING-Gocad consortium and the ANR projects ANR-10-BLAN-613 and ANR-16-CE31-0022-01 also enabled this work.

REFERENCES

- Abdulle, A. & Grote, M.J., 2011. Finite element heterogeneous multiscale method for the wave equation, *Multiscale Model. Simul.*, **9**(2), 766–792.
- Afanasiev, M., Boehm, C., May, D. & Fichtner, A., 2016. Using effective medium theory to better constrain full waveform inversion, *First Break*, **34**(7), 93–94.
- Allaire, G., 1992. Homogenization and two-scale convergence, *SIAM J. Math. Anal.*, **23**, 1482–1518.
- Allaire, G. & Capdeboscq, Y., 2000. Homogenization of a spectral problem in neutronic multigroup diffusion, *Comput. Methods Appl. Mech. Eng.*, **187**(1–2), 91–117.
- Allaire, G. & Habibi, Z., 2013. Homogenization of a conductive, convective, and radiative heat transfer problem in a heterogeneous domain, *SIAM J. Math. Anal.*, **45**(3), 1136–1178.
- Bacigalupo, A. & Gambarotta, L., 2014. Second-gradient homogenized model for wave propagation in heterogeneous periodic media, *Int. J. Solids Struct.*, **51**(5), 1052–1065.
- Backus, G., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.*, **67**(11), 4427–4440.
- Bensoussan, A., Lions, J.-L. & Papanicolaou, G., 1978. *Asymptotic Analysis for Periodic Structures*, North-Holland.
- Bodin, T., Capdeville, Y., Romanowicz, B. & Montagner, J.-P., 2015. Interpreting radial anisotropy in global and regional tomographic models, in *The Earth’s Heterogeneous Mantle*, pp. 105–144, eds Khan, A. & Deschamps, F., Springer.
- Botella, A., 2016. Génération de maillages non structurés volumiques de modèles géologiques pour la simulation de phénomènes physiques, *PhD thesis*, Université de Lorraine.
- Boutin, C. & Auriault, J.L., 1993. Rayleigh scattering in elastic composite materials, *Int. J. Eng. Sci.*, **31**(12), 1669–1689.
- Brossier, R., Operto, S. & Virieux, J., 2015. Velocity model building from seismic reflection data by full waveform inversion, *Geophys. Prospect.*, **63**(2), 354–367.
- Burgos, G., Capdeville, Y. & Guillot, L., 2016. Homogenized moment tensor and the effect of near-field heterogeneities on nonisotropic radiation in nuclear explosion, *J. geophys. Res.*, **121**(6), 4366–4389.

- Cance, P. & Capdeville, Y., 2015. Validity of the acoustic approximation for elastic waves in heterogeneous media, *Geophysics*, **80**(4), T161–T173.
- Capdeville, Y. & Cance, P., 2015. Residual homogenization for elastic wave propagation in complex media, *Geophys. J. Int.*, **200**, 984–997.
- Capdeville, Y. & Marigo, J., 2007. Second order homogenization of the elastic wave equation for non-periodic layered media, *Geophys. J. Int.*, **170**, 823–838.
- Capdeville, Y. & Marigo, J., 2008. Shallow layer correction for spectral element like methods, *Geophys. J. Int.*, **172**, 1135–1150.
- Capdeville, Y. & Marigo, J.-J., 2013. A non-periodic two scale asymptotic method to take account of rough topographies for 2-D elastic wave propagation, *Geophys. J. Int.*, **192**(1), 163–189.
- Capdeville, Y., Guillot, L. & Marigo, J., 2010a. 1-D non-periodic homogenization for the seismic wave equation, *Geophys. J. Int.*, **181**, 907–910.
- Capdeville, Y., Guillot, L. & Marigo, J., 2010b. 2-D non-periodic homogenization to upscale elastic media for P-SV waves, *Geophys. J. Int.*, **182**, 903–922.
- Capdeville, Y., Stutzmann, E., Wang, N. & Montagner, J.-P., 2013. Residual homogenization for seismic forward and inverse problems in layered media, *Geophys. J. Int.*, **194**(1), 470–487.
- Capdeville, Y., Zhao, M. & Cupillard, P., 2015. Fast Fourier homogenization for elastic wave propagation in complex media, *Wave Motion*, **54**, 170–186.
- Carcione, J., Picotti, S. & Santos, J., 2012. Numerical experiments of fracture-induced velocity and attenuation anisotropy, *Geophys. J. Int.*, **191**(3), 1179–1191.
- Casarotti, E., Stupazzini, M., Lee, S., Komatitsch, D., Piersanti, A. & Tromp, J., 2008. CUBIT and seismic wave propagation based upon the spectral-element method: an advanced unstructured mesher for complex 3D geological media, in *Proceedings of the 16th International Meshing Roundtable*, pp. 579–597, eds Brewer, M.L. & Marcum, D., Springer, Berlin.
- Caumon, G., Collon, P., Le Carlier de Veslud, C., Sausse, J. & Viseur, S., 2009. Surface-based 3D modeling of geological structures, *Math. Geosci.*, **41**(8), 927–945.
- Chaljub, E., Moczo, P., Tsuno, S., Bard, P.-Y., Kristek, J., Käser, M., Stupazzini, M. & Kristeková, M., 2010. Quantitative comparison of four numerical predictions of 3D ground motion in the Grenoble valley, France, *Bull. seism. Soc. Am.*, **100**(4), 1427–1455.
- Cormen, T.H., Leiserson, C.E., Rivest, R.L. & Stein, C., 2009. *Introduction to Algorithms*, 3rd edn, The MIT Press.
- Courant, R., Friedrichs, K. & Lewy, H., 1928. Über die partiellen Differenzgleichungen der mathematischen Physik, *Math. Ann.*, **100**, 32–74.
- Cupillard, P., Delavaud, E., Burgos, G., Festa, G., Vilotte, J.-P., Capdeville, Y. & Montagner, J.-P., 2012. RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale, *Geophys. J. Int.*, **188**, 1203–1220.
- Dal Maso, G., 1993. *An Introduction to Γ -convergence*, Birkhäuser.
- Diaz, J. & Grote, M.J., 2015. Multi-level explicit local time-stepping methods for second-order wave equations, *Comput. Methods Appl. Mech. Eng.*, **291**, 240–265.
- Dumbser, M., Käser, M. & Toro, E.F., 2007. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes—V. Local time stepping and p-adaptivity, *Geophys. J. Int.*, **171**(2), 695–717.
- Dumontet, H., 1990. Homogénéisation et effets de bords dans les matériaux composites, *PhD thesis*, Université Paris 6.
- Engquist, B., Holst, H. & Rumborg, O., 2009. *Multi-scale methods for wave propagation in heterogeneous media*, preprint [arXiv:0911.2638](https://arxiv.org/abs/0911.2638).
- Etienne, V., Chaljub, E., Virieux, J. & Glinsky, N., 2010. An hp-adaptive discontinuous Galerkin finite-element method for 3-D elastic wave modelling, *Geophys. J. Int.*, **183**, 941–962.
- Felippa, C., 2004. A compendium of FEM integration formulas for symbolic work, *Eng. Comput.*, **21**(8), 867–890.
- Fichtner, A. & Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models - a technique based on non-linear dispersion curve matching (DCM), *Geophys. J. Int.*, **173**(2), 519–533.
- Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental and global scale full-waveform inversion in the time-frequency domain, *Geophys. J. Int.*, **175**(8), 665–685.
- Fichtner, A., Kennett, B.L. & Trampert, J., 2013a. Separating intrinsic and apparent anisotropy, *Phys. Earth planet. Inter.*, **219**, 11–20.
- Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villasenor, A., 2013b. Multi-scale full waveform inversion, *Geophys. J. Int.*, **194**, 534–556.
- Fish, J. & Chen, W., 2001. Higher-order homogenization of initial/boundary-value problem, *J. Eng. Mech.*, **127**(12), 1223–1230.
- Fish, J. & Chen, W., 2004. Space-time multiscale model for wave propagation in heterogeneous media, *Comput. Methods Appl. Mech. Eng.*, **193**, 4837–4856.
- Gao, K., Chung, E., Gibson, R., Fu, S. & Efendiev, Y., 2015. A numerical homogenization method for heterogeneous, anisotropic elastic media based on multiscale theory, *Geophysics*, **80**(4), D385–D401.
- Gokhberg, A. & Fichtner, A., 2016. Full-waveform inversion on heterogeneous HPC systems, *Comput. Geosci.*, **89**, 260–268.
- Grechka, V., 2003. Effective media: A forward modeling view, *Geophysics*, **68**(6), 2055–2062.
- Grechka, V., 2007. Multiple cracks in vti rocks: Effective properties and fracture characterization, *Geophysics*, **72**(5), D81–D91.
- Grechka, V. & Kachanov, M., 2006. Effective elasticity of fractured rocks: A snapshot of the work in progress, *Geophysics*, **71**(6), W45–W58.
- Guillot, L., Capdeville, Y. & Marigo, J., 2010. 2-D non-periodic homogenization of the elastic wave equation: SH case, *Geophys. J. Int.*, **182**, 1438–1454.
- Hashin, Z. & Shtrikman, S., 1963. A variational approach to the elastic behavior of multiphase materials, *J. Mech. Phys. Solids*, **11**, 127–140.
- Hill, R., 1965. A self-consistent mechanics of composit materials, *J. Mech. Phys. Solids*, **13**, 213–222.
- Hornung, U., 1997. *Homogenization and Porous Media*, Vol. 6 of Interdisciplinary Applied Mathematics Series, Springer Verlag.
- Hughes, T.J., 2012. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Courier Corporation.
- Irakarama, M., Cupillard, P., Caumon, G. & Sava, P., 2017. Appraising structural interpretations using seismic data misfit functionals, in *79th EAGE Conference and Exhibition*, EAGE.
- Jordan, T.H., 2015. An effective medium theory for three-dimensional elastic heterogeneities, *Geophys. J. Int.*, **203**(2), 1343–1354.
- Komatitsch, D., Erlebacher, G., Göddeke, D. & Michéa, D., 2010. High-order finite-element seismic wave propagation modeling with MPI on a large GPU cluster, *J. Comput. Phys.*, **229**(20), 7692–7714.
- Kutsenko, A., Shuvalov, A. & Norris, A., 2013. On the quasistatic effective elastic moduli for elastic waves in three-dimensional phononic crystals, *J. Mech. Phys. Solids*, **61**(11), 2240–2259.
- Lekić, V., Panning, M. & Romanowicz, B., 2010. A simple method for improving crustal corrections in waveform tomography, *Geophys. J. Int.*, **182**, 265–278.
- Lepte, V., 2005. Two-scale extensions for non-periodic coefficients, [arXiv:math.AP/0512123](https://arxiv.org/abs/math/0512123).
- Lin, C., Saleh, R., Milkereit, B. & Liu, Q., 2017. Effective media for transversely isotropic models based on homogenization theory: With applications to borehole sonic logs, *Pure appl. Geophys.*, **174**(7), 2631–2647.
- Mauge, C. & Kachanov, M., 1994. Effective elastic properties of an anisotropic material with arbitrarily oriented interacting cracks, *J. Mech. Phys. Solids*, **42**, 561–584.
- Métivier, L., Bretaudeau, F., Brossier, R., Operto, S. & Virieux, J., 2013. Full waveform inversion and the truncated Newton method: quantitative imaging of complex subsurface structures, *Geophys. Prospect.*, **62**(6), 1353–1375.
- Minisini, S., Zhebel, E., Kononov, A. & Mulder, W.A., 2013. Local time stepping with the discontinuous Galerkin method for wave propagation in 3D heterogeneous media, *Geophysics*, **78**(3), T67–T77.
- Moulinec, H. & Suquet, P., 1998. A numerical method for computing the overall response of nonlinear composites with complex microstructure, *Comput. Methods Appl. Mech. Eng.*, **157**(1), 69–94.

- Nguetseng, G., 1989. A general convergence result for a functional related to the theory of homogenization, *SIAM J. Math. Anal.*, **20**(3), 608–623.
- Papanicolaou, G.C. & Varadhan, S. R.S., 1981. Boundary value problems with rapidly oscillating random coefficients, in *Random Fields, Vol. I, II (Esztergom, 1979)*, Vol. 27 of Colloq. Math. Soc. János Bolyai, pp. 835–873, North-Holland.
- Parnell, W. & Abrahams, I., 2008. Homogenization for wave propagation in periodic fibre-reinforced media with complex microstructure. I - Theory, *J. Mech. Phys. Solids*, **56**(7), 2521–2540.
- Pellerin, J., Botella, A., Bonneau, F., Mazuyer, A., Chauvin, B., Levy, B. & Caumon, G., 2017. RINGMesh: a programming library for developing mesh-based geomodeling applications, *Comput. Geosci.*, **104**, 93–100.
- Pelties, C., Käser, M., Hermann, V. & Castro, C.E., 2011. Regular versus irregular meshing for complicated models and their effect on synthetic seismograms, *Geophys. J. Int.*, **183**, 1031–1051.
- Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**, 721–739.
- Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, **167**(2), 495–503.
- Pratt, R., Shin, C. & Hicks, G., 1998. Gauss–Newton and full Newton methods in frequency domain seismic waveform inversion, *Geophys. J. Int.*, **133**, 341–362.
- Remacle, J.-F., Gandham, R. & Warburton, T., 2016. GPU accelerated spectral finite elements on all-hex meshes, *J. Comput. Phys.*, **324**, 246–257.
- Reuss, A., 1929. Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle, *J. Appl. Math. Mech.*, **9**(1), 49–58.
- Ricard, Y., Durand, S., Montagner, J.-P. & Chambat, F., 2014. Is there seismic attenuation in the mantle?, *Earth planet. Sci. Lett.*, **388**, 257–264.
- Rietmann, M., Grote, M., Peter, D. & Schenk, O., 2017. Newmark local time stepping on high-performance computing architectures, *J. Comput. Phys.*, **334**, 308–326.
- Sanchez-Palencia, E., 1980. *Non-homogenous Media and Vibration Theory*, Vol. 127 of Lecture notes in physics, Springer-Verlag.
- Santugini-Repique, K., 2007. Homogenization of the demagnetization field operator in periodically perforated domains, *J. Math. Anal. Appl.*, **334**(1), 502–516.
- Sayers, C., 1998. Long-wave seismic anisotropy of heterogeneous reservoirs, *Geophys. J. Int.*, **132**, 667–673.
- Sayers, C. & Kachanov, M., 1991. A simple technique for finding effective elastic constants of cracked solids for arbitrary crack orientation statistics, *Int. J. Solids Struct.*, **27**(6), 671–680.
- Sayers, C. & Kachanov, M., 1995. Microcrack-induced elastic wave anisotropy of brittle rocks, *J. geophys. Res.*, **100**(B3), 4149–4156.
- Schenk, O. & Gärtner, K., 2006. On fast factorization pivoting methods for symmetric indefinite systems, *Electron. Trans. Numer. Anal.*, **23**, 158–179.
- Schoenberg, M. & Helbig, K., 1997. Orthorhombic media: modeling elastic wave behavior in a vertically fractured earth, *Geophysics*, **62**(6), 1954–1974.
- Schoenberg, M. & Muir, F., 1989. A calculus for finely layered anisotropic media, *Geophysics*, **54**, 581–589.
- Schoenberg, M. & Sayers, C.M., 1995. Seismic anisotropy of fractured rock, *Geophysics*, **60**(1), 204–211.
- Silwal, V. & Tape, C., 2016. Seismic moment tensors and estimated uncertainties in southern Alaska, *J. geophys. Res.*, **121**(4), 2772–2797.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.
- Virieux, J., Calandra, H. & Plessix, R.-E., 2011. A review of the spectral, pseudo-spectral, finite-difference and finite-element modelling techniques for geophysical imaging, *Geophys. Prospect.*, **59**(5), 794–813.
- Voigt, W., 1889. Über die Beziehung zwischen den beiden Elastizitätsconstanten isotroper Körper, *Ann. Phys.*, **274**(12), 573–587.
- Wang, N., Montagner, J.-P., Fichtner, A. & Capdeville, Y., 2013. Intrinsic versus extrinsic seismic anisotropy: The radial anisotropy in reference earth models, *Geophys. Res. Lett.*, **40**(16), 4284–4288.
- Warner, M. & Guasch, L., 2016. Adaptive waveform inversion: Theory, *Geophysics*, **81**(6), R429–R445.
- Wei, W., Fu, L. & Blacquière, G., 2012. Fast multifrequency focal beam analysis for 3D seismic acquisition geometry, *Geophysics*, **77**(2), P11–P21.
- Weiss, R.M. & Shragge, J., 2013. Solving 3D anisotropic elastic wave equations on parallel GPU devices, *Geophysics*, **78**(2), F7–F15.
- Worth, D., Greenough, C. & Chin, S., 2012. Pragmatic software engineering for computational science, in *Handbook of Research on Computational Science and Engineering*, Vol. 1, pp. 119–149, eds Leng, J. & Sharrock, W., Information Science Reference.
- Zhao, M., Capdeville, Y. & Zhang, H., 2016. Direct numerical modeling of time-reversal acoustic subwavelength focusing, *Wave Motion*, **67**, 102–115.
- Zijl, W., Hendriks, M.A. & Hart, C.M.P., 2002. Numerical homogenization of the rigidity tensor in Hooke’s law using the node-based finite element method, *Math. Geol.*, **34**(3), 291–322.

APPENDIX A: THE LOW-PASS FILTER IN THE 1-D CASE

To separate the macroscopic and the microscopic scales, the non-periodic homogenization method requires the use of a low-pass filter $\mathcal{F}^{\varepsilon_0}$. Applied to any function $f: \mathbb{R} \rightarrow \mathbb{R}$, this filter removes all the scales smaller than λ_0 :

$$\mathcal{F}^{\varepsilon_0}\{f\}(x) = \int_{\mathbb{R}} f(x') w^{\varepsilon_0}(x - x') dx' \quad (\text{A1})$$

where w^{ε_0} is the wavelet whose spectrum is

$$\hat{w}^{\varepsilon_0}(k) = \begin{cases} 1 & \text{for } k \leq \frac{2\pi}{\lambda_0} \\ 0 & \text{for } k > \frac{2\pi}{\lambda_0}. \end{cases} \quad (\text{A2})$$

In this last expression, $\hat{\cdot}$ represents the Fourier transform. Note that the filter and the associated wavelet are indexed by ε_0 because λ_0 depends on the choice of ε_0 ($\lambda_0 = \varepsilon_0 \lambda_m$).

If defined by spectrum (A2), the wavelet w^{ε_0} has an infinite support, which is unmanageable in practice. For that reason, we introduce a cosine-taper to soften the cut-off:

$$\hat{w}^{\varepsilon_0}(k) = \begin{cases} 1 & \text{for } k \leq k_0 \\ \frac{1}{2} [1 + \cos(\pi \frac{k-k_0}{\Delta k})] & \text{for } k \in]k_0; k_0 + \Delta k] \\ 0 & \text{for } k > k_0 \end{cases} \quad (\text{A3})$$

where $k_0 = \frac{2\pi}{\lambda_0}$. With such a definition, w^{ε_0} can be truncated with a negligible error to get a compact support (Fig. A1).

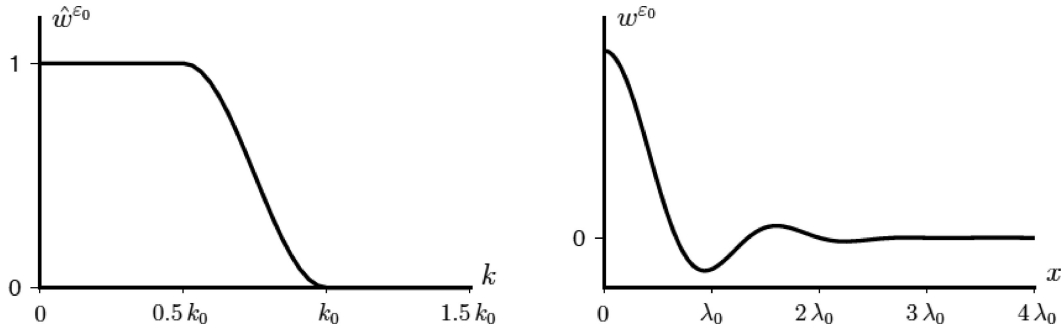
1000 *P. Cupillard and Y. Capdeville*

Figure A1. A 1-D spectrum \hat{w}^{ε_0} (left) and its corresponding wavelet w^{ε_0} (right). In this example, the length of the cosine-taper Δk is equal to $0.5 k_0$.

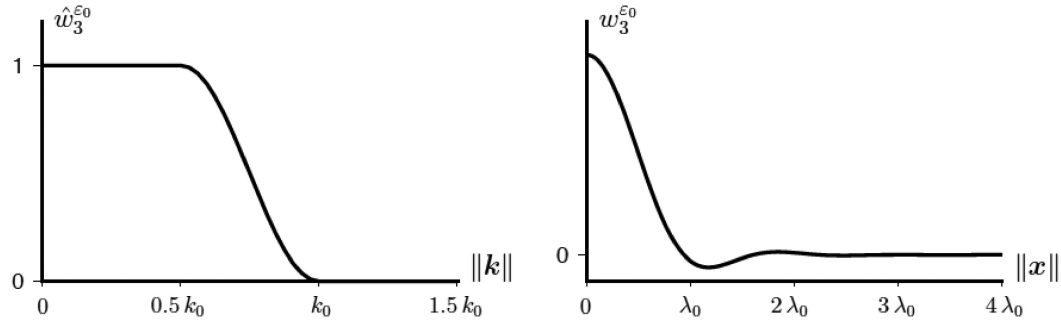


Figure A2. 3-D spectrum $\hat{w}_3^{\varepsilon_0}$ for $\Delta k = 0.5 k_0$ (left) and its corresponding 3-D wavelet $w_3^{\varepsilon_0}$ (right). Because both $\hat{w}_3^{\varepsilon_0}$ and $w_3^{\varepsilon_0}$ are spherical, we represent them as a function of $\|\mathbf{k}\|$ and $\|\mathbf{x}\|$, respectively. Comparing the obtained plots with Fig. A1, we note that the 3-D wavelet is different from just having the wavelet of the 1-D case in all the directions. This is simply because the (inverse) spherical Fourier transform is not the (inverse) 1-D Fourier transform.

APPENDIX B: ρ^{ε_0} MUST BELONG TO Ψ

Inserting eqs (2)–(4) and properties $E^{\varepsilon_0}(x, y)$ and $\rho^{\varepsilon_0}(x, y)$ instead of $E(x)$ and $\rho(x)$ into the initial elastodynamic problem yields a cascade of equations. For $i = 0$, it turns that

$$\rho^{\varepsilon_0}(x, y)\ddot{u}_0(x) - \nabla_x \sigma_0(x) - \nabla_y \sigma_1(x, y) = f(x), \quad (\text{B1})$$

where f is the external force and \ddot{u}_0 is the second time-derivative of the zeroth-order displacement. Because σ_1 has to lie in Ψ , ρ^{ε_0} must belong to Ψ as well.

APPENDIX C: THE LOW-PASS FILTER IN THE 3-D CASE

$\mathcal{F}_3^{\varepsilon_0}$ is the extension of the 1-D low-pass filter $\mathcal{F}^{\varepsilon_0}$ (Appendix A) to 3-D. Applied to any function $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, it removes all the scales smaller than λ_0 :

$$\mathcal{F}_3^{\varepsilon_0}\{f\}(\mathbf{x}) = \int_{\mathbb{R}^3} f(\mathbf{x}') w_3^{\varepsilon_0}(\mathbf{x} - \mathbf{x}') d\mathbf{x}', \quad (\text{C1})$$

where $w_3^{\varepsilon_0}$ is the wavelet whose 3-D spectrum is

$$\hat{w}_3^{\varepsilon_0}(\mathbf{k}) = \begin{cases} 1 & \text{for } \|\mathbf{k}\| \leq k_0 \\ \frac{1}{2} \left[1 + \cos\left(\pi \frac{\|\mathbf{k}\| - k_0}{\Delta k}\right) \right] & \text{for } \|\mathbf{k}\| \in]k_0; k_0 + \Delta k] \\ 0 & \text{for } \|\mathbf{k}\| > k_0. \end{cases} \quad (\text{C2})$$

k_0 still is the cut-off wave-number $\frac{2\pi}{\lambda_0}$. For all the applications within the present paper (Section 4), the length of the cosine-taper Δk has been chosen to be equal to $0.5 k_0$. When using this value, a truncation of the wavelet at $\|\mathbf{x}\| = 4 \lambda_0$ is satisfactory (Fig. A2).

APPENDIX D: $\mathbf{H}^{\varepsilon_0}$ AND χ^{ε_0} ARE PERIODIC IN y AND BELONG TO Ψ_3

As a solution of a periodic boundary value problem, χ_s is periodic in y , so $\nabla_y \chi_s$ is periodic in y . \mathbf{G}_s therefore is periodic in y , as well as \mathbf{H}_s , consequently. Finally, $\mathbf{G}^{\varepsilon_0}$ and $\mathbf{H}^{\varepsilon_0}$ are both periodic in y . By construction (eqs 28 and 29), they also belong to Ψ_3 .

Let us denote the cross product by \wedge . Because $\nabla_y \wedge \mathbf{G}_s = \mathbf{0}$ and $\nabla_y \wedge \mathcal{F}_3^{\varepsilon_0}\{f\} = \mathcal{F}_3^{\varepsilon_0}\{\nabla_y \wedge f\}$ for any tensorial function f , we have $\nabla_y \wedge \mathbf{G}^{\varepsilon_0} = \mathbf{0}$, so $\mathbf{G}^{\varepsilon_0}$ can be written as a constant plus a gradient of a displacement χ^{ε_0} . As a solution of eq. (23), $\nabla_y \chi^{\varepsilon_0}$ has no rotational component, so we end up with

$$\mathbf{G}^{\varepsilon_0} = \mathbf{I} + \frac{1}{2}(\nabla_y \chi^{\varepsilon_0} + {}^t \nabla_y \chi^{\varepsilon_0}). \quad (\text{D1})$$

By construction (eq. 28), $\langle \mathbf{G}^{\varepsilon_0} \rangle_3 = \mathbf{I}$ (that would not be the case if $\mathbf{G}^{\varepsilon_0}$ was built as E^{ε_0} and ρ^{ε_0} are—eqs 17 and 18—), so

$$\left\langle \frac{1}{2}(\nabla_y \chi^{\varepsilon_0} + {}^t \nabla_y \chi^{\varepsilon_0}) \right\rangle_3 = 0. \quad (\text{D2})$$

For any tensorial function $f \in \Psi_3$ periodic in y ,

$$\begin{aligned} \langle f \rangle_3 &= 0 \text{ and } \nabla_y \mathbf{g} \\ &= f \implies \mathbf{g} \text{ is periodic in } y \text{ and belongs to } \Psi_3, \end{aligned} \quad (\text{D3})$$

so χ^{ε_0} is periodic in y and belongs to Ψ_3 .

This demonstration can also be found in the seminal paper by Capdeville *et al.* (2010b), section 3.5.

APPENDIX E: HEURISTIC FOR SAMPLING A MODEL TO BE HOMOGENIZED

When honouring discontinuities, finite element meshes enable a correct account for the geometry of the medium in which the phenomenon to be simulated occurs. Nevertheless, the element size and the interpolation degree which guarantee the accurate capture of the phenomenon cannot be assessed rigorously in most of finite element applications. To choose these two parameters, rules of thumb are often used. In the case of the non-periodic homogenization, we know that our outputs $\rho^{\varepsilon_0^*}$ and $\mathbf{C}^{\varepsilon_0^*}$ only contain scales larger than λ_0 , so we sample them using a space-step $dx = \frac{\lambda_0}{4}$. This choice corresponds to twice the Nyquist rate. Because $\mathbf{C}^{\varepsilon_0^*}$ is obtained from the space-derivative of the solution χ_s of the finite element analysis, we

required an additional degree for this solution, so we want χ_s to be sampled by a space-step $\delta x = \frac{\lambda_0}{5}$. Such a sampling is ensured by imposing

$$\frac{a}{N} \leq \frac{\lambda_0}{5}, \quad (\text{E1})$$

where a is the length of the edges of the mesh and N is the interpolation degree. eq. (E1) is our heuristic for sampling a model to be homogenized by our finite element code. In 3-D, assuming the tetrahedral elements to be regular, we can rewrite this heuristic in terms of the volume of the elements v given a degree N :

$$v \leq \frac{\sqrt{2}}{12} \left(\frac{N\lambda_0}{5} \right)^3. \quad (\text{E2})$$

submitted to *Geophys. J. Int.*

Interpretation of time reversal focal spot based on point-source homogenization

Zoé Renat¹, Paul Cupillard¹, Yann Capdeville²

¹ *GeoRessources, UMR 7359, Université de Lorraine/CNRS, Nancy, France*

² *LPG, UMR 6112, Université de Nantes/CNRS, Nantes, France*

SUMMARY

The time reversal method is based on the back-propagation of seismic waveforms recorded at a set of receivers. When this set forms a closed surface and the elastic properties of the medium are correct, the seismic energy focuses at the source location, creating a focal spot. Such a spot is smooth in space, whereas the original wavefield usually shows a displacement discontinuity at the source. The goal of the present paper is to discuss the link between the focal spot and the original source using the concept of homogenized point-source recently proposed by Capdeville (2021). We show that the back-propagated wavefield is equivalent to the sum of two low-wavenumber fields resulting from the homogenization of the original point-source. In other words, the homogenized point-source is the equivalent force for producing the focal spot. In addition to the demonstration in the general 3D heterogeneous case, we present some numerical examples in 2D.

Key words: Source observation – Theoretical seismology – Computational seismology

Introduction

Seismic event localization and characterization can be performed using multiple methods and various data such as travel times or waveforms. Two major families of methods can be distinguished. The first one is based on travel time picking (e.g., ?). Because they rely on picking specific seismic phases, travel time-based methods for determining the location and the mechanism of seismic events (Garmany 1979; Hampson & Russell 1984; Kennett & Sambridge 1992; Sambridge & Mosegaard 2002) are highly sensitive to the quality of the data (i.e. signal to noise ratio). Moreover, they disregard significant pieces of information contained in the seismic recordings. To overcome these limitations, waveform-based methods have been developed (e.g., McMechan 1982; Schuster et al. 2004; Larmat et al. 2006; Grigoli et al. 2016; Shi et al. 2019; Willacy et al. 2019). They rely on complete or partial recorded signals. Their principle is to back-propagate data to either solve a full waveform inverse problem or focus waves at the source location (Li et al. 2020). Full waveform inversion for seismic source characterization (e.g., Willacy et al. 2019) consists in minimizing a misfit function between observed and synthetic data. To do so, a prior source model is necessary to produce a first synthetic dataset. Then the gradient of the misfit function is computed using an adjoint field calculation. The adjoint field is generated by the data residuals at all receivers acting as sources. Other waveform-based approaches include partial waveforms stacking (e.g., Kao & Shan 2004), match field processing (e.g., Cros et al. 2011), back-projection (e.g., Ishii et al. 2007) and time reversal (e.g., Yang & Zhu 2019). They are often called source imaging methods because they all somehow focus waves to create an "image" of the source. The main advantage of these methods is the absence of prior source model and the direct (i.e. non-iterative) calculation of the solution. However, they may require data processing to clean the input signals, with the notable exception of time reversal, which is poorly sensible to noise (e.g., Gajewski & Tessmer 2005; Cao et al. 2012).

The time reversal technique was introduced by Fink (1992) in ultrasound acoustics. It has been extended to elastic waves by Draeger & Fink (1997). It consists in reversing in time the recorded signals and back-propagating them from the receivers. Under some conditions, the waves follow the inverse path of the forward way and focus at the source position. Such a behavior can be explained by the invariability of the wavefield through a reverse time operator because of the second-order time derivative of the wave equation

$$\rho \ddot{u}_n - \partial_j (C_{njkl} \partial_k u_l) = f_n, \quad (1)$$

with ρ the density, \mathbf{u} the displacement, $\ddot{\mathbf{u}}$ the second-order time derivative of the displacement, \mathbf{C} the elastic tensor, and \mathbf{f} a force distribution acting in the medium. The result of the back-propagation is a spot at the source location, named the focal spot. The conditions to apply time reversal are the following:

- (i) The receivers form a closed surface, called time reversal mirror.
- (ii) The time reversal mirror does not perturb the wave propagation.

Interpretation of time reversal focal spot based on point-source homogenization 3

(iii) The medium is well known. Errors on the velocity model or interface positions would generate modifications in the wave path as compared to the path in the forward process.

(iv) The anelasticity is negligible. Anelastic attenuation would make a first-order time derivative appear in the wave equation so that the time reversibility would not be verified.

In practice, conditions (i), (ii) and (iv) are difficult to satisfy: there is no continuous surfaces of receivers in the underground, the velocity model presents uncertainties and the earth is anelastic. Nevertheless, some solutions exist to overcome these difficulties. Cassereau & Fink (1992) developed a mathematical model to perform time reversal with a plane mirror in acoustic media. Li & van der Baan (2016) applied it in the 2D elastic case using wells as time reversal mirror. More recently, Finger & Saenger (2021) used surface data as a plane mirror to locate microseisms in a geothermal reservoir. To quantify the influence of the velocity model uncertainties on the focalization. Gajewski & Tessmer (2005) carried out a sensitivity study. They showed that strong variations in the velocity model leads to a focal spot shifted in time and space, whereas a smooth model (compare to the true model) leads to only a shift in time. Furthermore, Bazargani & Snieder (2016) presented an optimization of time reversal which is able to overcome some uncertainties on the velocity model. As for the anelasticity, Zhu (2014); Bai et al. (2019) proposed a mathematical solution to compensate attenuation and validated it using numerical experiments of time reversal for source locations.

Time reversal is used in multiple domains such as medicine (Fink 2015), underwater communication (Kuperman et al. 1998) and geosciences (Li & van der Baan 2016). In the latter domain, applications mainly deal with earthquake location and characterization: Larmat et al. (2008) applied it on glacial earthquakes, Douma et al. (2015); Li & van der Baan (2016); Yang & Zhu (2019); Finger & Saenger (2021) on microseismicity, and Rietbrock & Scherbaum (1994) to characterize an aftershock series. Time reversal can also be used to follow fluid displacements in the ground using tremors (Steiner et al. 2008). Other studies investigate time reversal for detecting scatterers (Shustak & Landa 2017; Rabinovich et al. 2018; Givoli & Turkel 2012) or reconstructing the complete wavefield (van Manen et al. 2006; Masson et al. 2014; Masson & Romanowicz 2017). In addition to the location, the focal spot obtained with time reversal possibly contains pieces of information on the source mechanism and the medium heterogeneities at the source. However, the question of the link between this focal spot and the details of the source of the earthquake still remains. Some pieces of answers exist in the literature: Kremers et al. (2011) succeeded to extract the centroid time in addition to the spatial location in the case of a single point source from a time reversal simulation, but not in the case of multiple sources. Chambers et al. (2014) proposed a method to recover the space and time location as well as the moment tensor of a source. Nakahara & Haney (2015) showed a mathematical way to interpret the image of the source based on point spread functions.

In this article, we present a new way to interpret the time reversal focal spot in the light of the point-source homogenization. First, we present the theory of time reversal and a simple synthetic example of the problem we want to tackle. Second, we interpret the time reversal wavefield using the theory of point-source homogenization, which we illustrate with numerical 2D examples involving i)

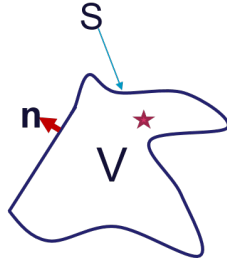


Figure 1. Set-up of the representation theorem. The star represents a point source. n is a unit-vector normal to the closed surface S . The theorem allows us calculating the displacement at any point in the volume V .

an explosion, ii) P- and S-waves and iii) an extended source. Finally, we discuss the implications and the limits of our results.

1 TIME REVERSAL

1.1 Theory

Relying on the four conditions presented in the introduction, we can go deeper into the physics of time reversal. Let us first invoke the representation theorem (Masson et al. 2014), which reads in the frequency domain as

$$\begin{aligned}
 u_i(\mathbf{x}) = & \int_V G_{in}(\mathbf{x}, \mathbf{x}') f_n(\mathbf{x}') dV \\
 & + \oint_S G_{in}(\mathbf{x}, \mathbf{x}') n_j C_{njkl} \partial'_k u_l(\mathbf{x}') dS' \\
 & - \oint_S u_n(\mathbf{x}') n_j C_{njkl} \partial'_k G_{il}(\mathbf{x}, \mathbf{x}') dS',
 \end{aligned} \tag{2}$$

with $\mathbf{G}(\mathbf{x}, \mathbf{x}')$ the Green tensor between the points \mathbf{x} and \mathbf{x}' , \mathbf{n} the unit normal vector to the closed surface S , and V the volume bounded by S (Fig. 1). This theorem establishes that the displacement at any position \mathbf{x} can be retrieved from three terms:

- a volume integral which involves the body force \mathbf{f} ;
- a surface integral which involves the normal traction $T_n = n_j C_{njkl} \partial'_k u_l$ at the surface S ;
- a surface integral which involves the tensor $M_{kl} = u_n n_j C_{njkl}$ at the surface S .

In the frequency domain, reversing the time means taking the complex conjugate, denoted by $*$. Applying this to Eq. (2) and benefiting from the Green tensor reciprocity, the time-reversed displacement is given by

$$\begin{aligned}
 u_i^*(\mathbf{x}) = & \int_V G_{in}(\mathbf{x}, \mathbf{x}') f_n^*(\mathbf{x}') dV \\
 & + \oint_S G_{in}(\mathbf{x}, \mathbf{x}') n_j C_{njkl} \partial'_k u_l^*(\mathbf{x}') dS' \\
 & - \oint_S u_n^*(\mathbf{x}') n_j C_{njkl} \partial'_k G_{il}(\mathbf{x}, \mathbf{x}') dS'.
 \end{aligned} \tag{3}$$

When implementing this numerically, the surface S is discretized and the integrals are approximated

by a quadrature:

$$\begin{aligned}
u_i^*(\mathbf{x}) &= \int_V G_{in}(\mathbf{x}, \mathbf{x}') f_n^*(\mathbf{x}') dV \\
&\quad + \sum_{r=1}^n \alpha_r G_{in}(\mathbf{x}, \mathbf{x}'_r) T_n^*(\mathbf{x}'_r) \\
&\quad - \sum_{r=1}^n \alpha_r M_{kl}^*(\mathbf{x}'_r) \partial'_k G_{il}(\mathbf{x}, \mathbf{x}'_r),
\end{aligned} \tag{4}$$

with \mathbf{x}'_r ($r=1, \dots, n$) the quadrature nodes with the associated weight α_r . Applying the Dirac property $f(\mathbf{x}) = \int_V f(\mathbf{x}') \delta(\mathbf{x}' - \mathbf{x}) d\mathbf{x}'$ to the last two terms and integrating by part the last term, we obtain

$$\begin{aligned}
u_i^*(\mathbf{x}) &= \int_V G_{in}(\mathbf{x}, \mathbf{x}') f_n^*(\mathbf{x}') dV \\
&\quad + \sum_{r=1}^n \int_V \alpha_r G_{in}(\mathbf{x}, \mathbf{x}') T_n^*(\mathbf{x}'_r) \delta(\mathbf{x}' - \mathbf{x}'_r) dV' \\
&\quad + \sum_{r=1}^n \int_V \alpha_r G_{il}(\mathbf{x}, \mathbf{x}') \partial'_k [M_{kl}^*(\mathbf{x}'_r) \delta(\mathbf{x}' - \mathbf{x}'_r)] dV'.
\end{aligned} \tag{5}$$

This last equation means that we can reconstruct the time-reversed displacement \mathbf{u}^* from

- the time-reversed body force \mathbf{f}^* , which actually acts as a sink;
- n monopoles derived from the time-reversed traction \mathbf{T}^* at each point \mathbf{x}_r on the surface S :

$$\mathbf{f}_r^m(\mathbf{x}) = \mathbf{T}^*(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_r) \alpha_r ; \tag{6}$$

- n dipoles derived from the time-reversed moment tensor \mathbf{M}^* at each point \mathbf{x}_r of the surface S :

$$\mathbf{f}_r^d(\mathbf{x}) = \nabla \cdot [\mathbf{M}^*(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_r)] \alpha_r . \tag{7}$$

In the time domain, Eq. (5) can be summarized as follows:

$$\mathbf{u}(\mathbf{x}, T - t) = \mathbf{u}^{tr}(\mathbf{x}, t) + \mathbf{u}^s(\mathbf{x}, t) \tag{8}$$

with $\mathbf{u}(\mathbf{x}, T - t)$ the forward displacement reversed in time (T being the propagation duration), $\mathbf{u}^{tr} = \mathbf{u}^m + \mathbf{u}^d$ the displacement generated from the closed surface (\mathbf{u}^m being the displacement created by the monopoles and \mathbf{u}^d the displacement created by the dipoles) and \mathbf{u}^s the displacement generated by the sink force. \mathbf{u}^{tr} is the field that can be computed in practice because the force \mathbf{f} is not available (this is our unknown). As a consequence, we do not reconstruct \mathbf{u} , but a smooth version of it, as we will see in the following.

1.2 A simple example in 2D

To shed a first light on the similarities and discrepancies between \mathbf{u} and \mathbf{u}^{tr} , we compute the two fields in a simple homogeneous 2D case with a perfect time reversal mirror. The model is a 10 km x 10 km square with absorbing boundaries (PML). The source is a point-source equipped with the identity moment tensor (i.e. an explosion) and a Ricker wavelet time function. The source and closed surface positions are summarized in Fig. 2. The simulations are performed using the spectral element

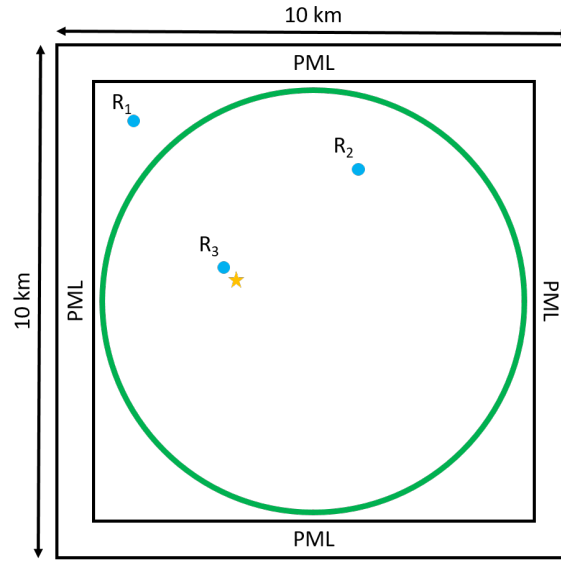


Figure 2. Set up of our first numerical experiment. The medium is homogeneous with $V_p = 3200 \text{ m/s}$, $V_s = 2000 \text{ m/s}$ and $\rho = 2200 \text{ kg/m}^3$. The source (orange star) is positioned in the area defined by the green circle which represents the closed surface on which the signal is recorded. The medium is surrounded by absorbing boundaries. The blue disks mark three positions at which the wavefield is output in each simulation (backward and forward) to do waveform comparisons.

software SPECFEM2D (Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999) following a three steps procedure:

- (i) creation of the closed surface of receivers and estimation of the integration weights α_r and the normal vectors \mathbf{n}_r ;
- (ii) forward simulation using SPECFEM2D and storage of both the displacement and the stress tensor at each receiver position to use them as input for the next step;
- (iii) backward simulation using SPECFEM2D, to obtain the focal spot.

Fig. 3 shows the forward simulation \mathbf{u} and the backward simulation \mathbf{u}^{tr} obtained with both monopole and dipole terms applied together. In the backward simulation, no waves are generated beyond the closed surface S (Fig. 3 at t_4). On the contrary, in the volume V embedded by the closed surface S , we observe a wavefield which converges to the source location and forms a focal spot. At t_4, t_3, t_2 , this wavefield perfectly corresponds to the forward displacement \mathbf{u} . After (in reverse time) t_2 , the backward wavefield diverges and no longer corresponds to the forward wavefield. If the sink term was implemented, the energy would be absorbed at the source location and the wavefield would match the forward simulation at all times (Masson et al. 2014). The absence of the sink actually creates the focal spot which is different from the forward near-field. To go further, we compare the displacement at three specific positions: beyond the closed surface, in the volume V far from the source position and in the volume V near the source position (Fig. 2).

When the receiver lays beyond the recording surface (Fig. 4, top row), the backward displacement

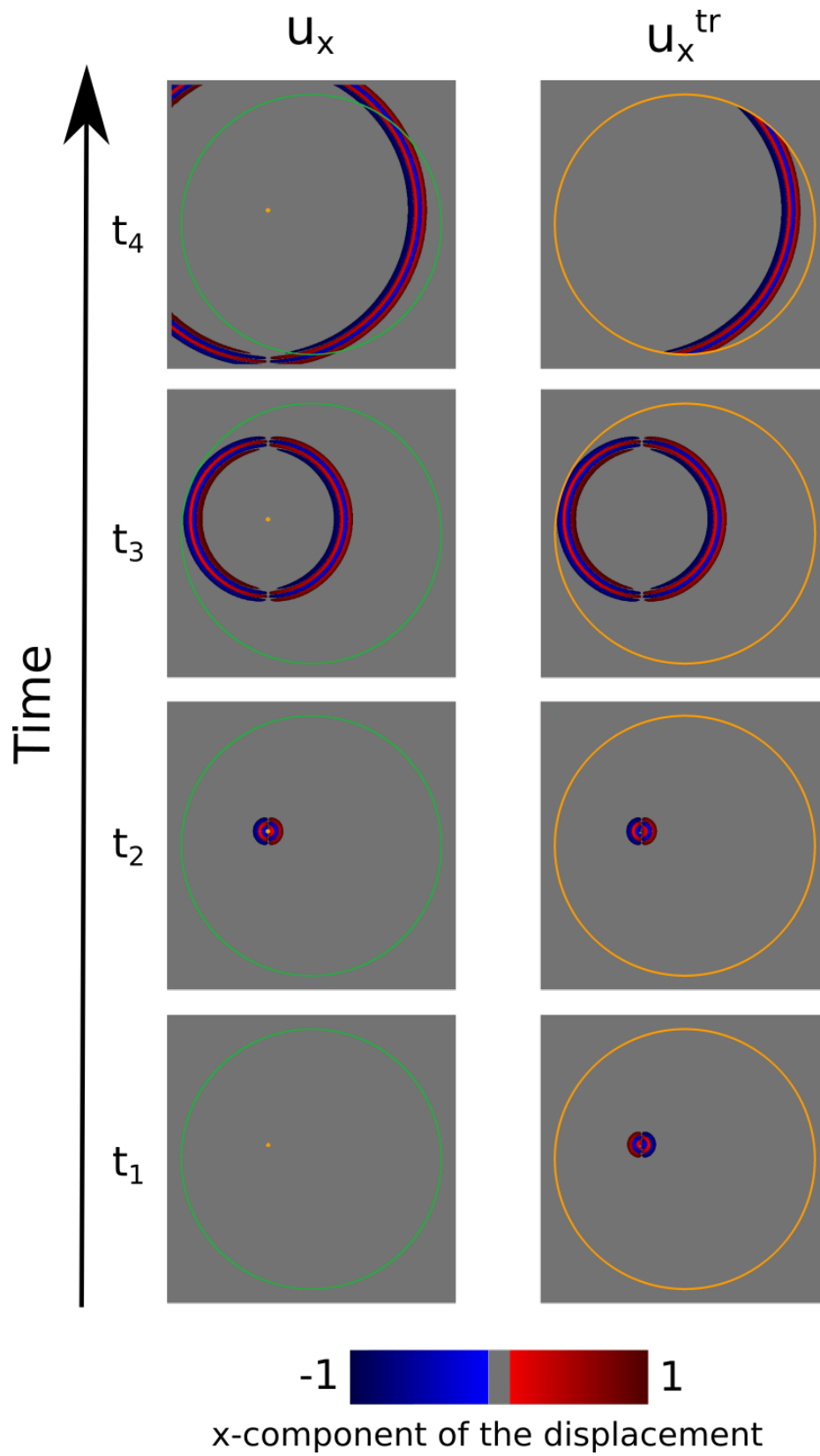


Figure 3. Snapshots of the x-component of the normalized displacement. The left hand-side column represents the forward displacement \mathbf{u} . The right hand-side column is the backward displacement \mathbf{u}^{tr} generated by monopoles (Eq. (6)) and dipoles (Eq. (7)). The receivers are represented in green, and the sources in orange.

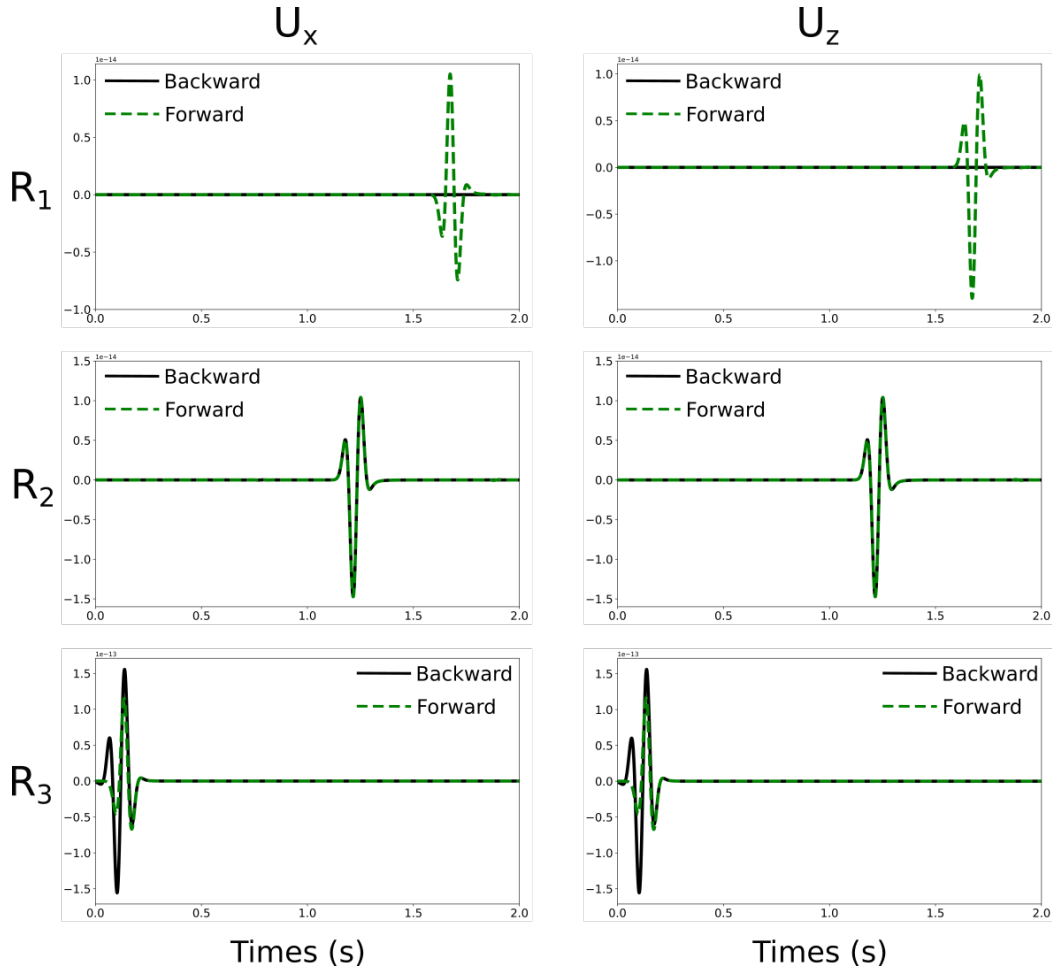


Figure 4. Comparison of forward \mathbf{u} and backward \mathbf{u}^{tr} waveforms at three different receiver positions (Fig. 2). The black line represents the backward displacement and the green dashed line the forward displacement. The first column shows the x-component and the second column shows the z-component.

is zero, meaning that $\mathbf{u}^m = -\mathbf{u}^d$. When the receiver is in the volume V and more than one wavelength away to the source location (Fig. 4, middle row), the backward displacement perfectly matches the forward displacement. Finally, when the receiver is less than one wavelength away to the source location (Fig. 4, bottom row), the backward displacement does not match the forward displacement: there is a smooth focal spot instead of a displacement discontinuity at the focalization time (i.e. the centroid time) in the backward simulation. To better see what happens near the source, we record the displacement along a receiver line positioned at the source depth and plot the displacement along this line for different time steps (Fig. 5). The displacement at t_4 and t_3 confirms the results presented in the previous figures: the backward displacement matches the forward displacement, even at the source location. However, at earlier times (t_1 and t_2), the two displacements do not match. The backward displacement is smoother than the forward displacement which presents a sharp discontinuity at the source position. This discontinuity corresponds to the seismic rupture process, which is mathematically

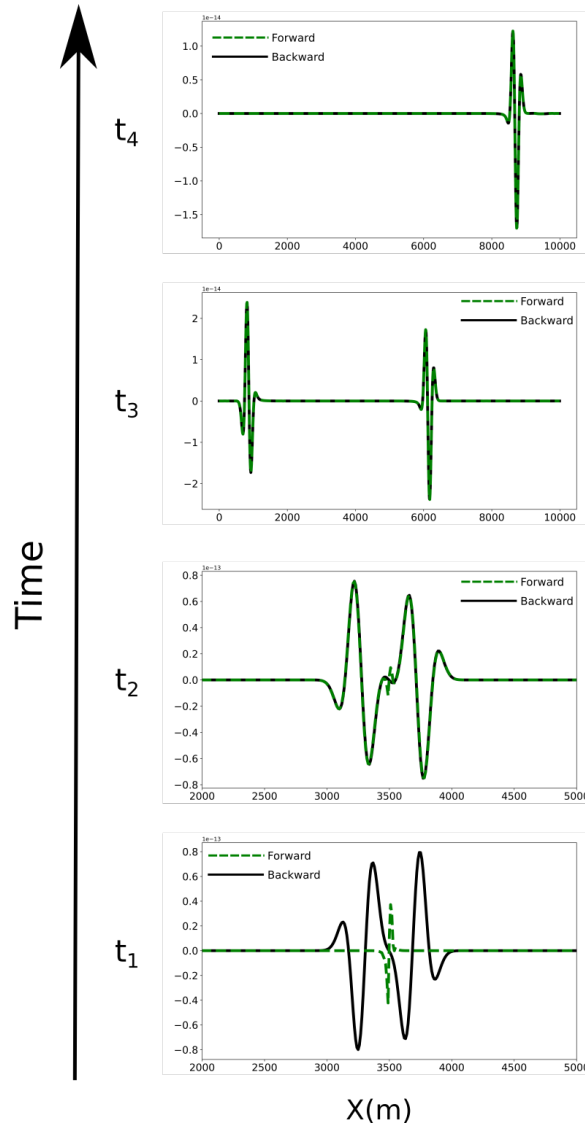


Figure 5. Plot of the x-component of the displacement along a line of receivers at the source depth for four different time steps. The steps are those chosen on Fig. 2. The black line represents the backward displacement \mathbf{u} and the dashed green line the forward displacement \mathbf{u}^{tr} .

modeled by a space-derivative of a Dirac function. In theory, this discontinuity is even sharper than what we observed in Fig. 5 because the simulation is limited by the mesh resolution (Capdeville 2021).

In smooth media such as the homogeneous one we are considering in this subsection, the smoothness of the focal spot is easily explained by the fact that \mathbf{u}^{tr} is generated by two smooth forces (Eq. (6) and Eq. (7)) so \mathbf{u}^m and \mathbf{u}^d are smooth. In more complex media which contains small-scale features and scatterers, the smoothness of the focal spot can be explained by the diffraction limit $\lambda_{min}/2$ (Cassereau & Fink 1992). Nevertheless, when specific scatterers (e.g., split ring resonators) lie near the source, this limit can be beaten and superresolution (i.e. subwavelength focusing) can be obtained (de Rosny & Fink 2002; Borcea et al. 2003; Fink 2006; Lerosey et al. 2007; Gélius & Asgedom 2011; Schuster et al.

10 *Zoé Renat, Paul Cupillard, Yann Capdeville*

2012; Zhao et al. 2016). In the present work, we disregard this case and put this interesting property of the time reversal on the side.

From Eq. (8), we can write the time reversal displacement as

$$\mathbf{u}^{tr}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, T - t) - \mathbf{u}^s(\mathbf{x}, t), \quad (9)$$

which means that the smooth time reversal wavefield can also be seen as a combination of two discontinuous wavefields: the forward and the sink. It is not trivial to prove that such a combination can generate a continuous wavefield. In the following part, we show that it can be demonstrated thanks to the theory of point-source homogenization developed by Capdeville (2021).

2 POINT-SOURCE HOMOGENIZATION

2.1 Principle and Mathematics

The non-periodic homogenization has proved to enable the computation of accurate effective properties for the elastic wave equation, which allows to perform wave simulations in complex geological models using coarse meshes and a tractable computational cost (e.g. Capdeville et al. 2010; Cupillard & Capdeville 2018; Capdeville et al. 2020). Homogenization consists in low-pass filtering relevant quantities derived from the model to be smoothed, namely the strain and the stress *concentrators*. To define the low-pass filter F^{ϵ_0} , a scaling parameter ϵ_0 is introduced:

$$\epsilon_0 = \frac{\lambda_0}{\lambda_{min}}, \quad (10)$$

where λ_0 is the wavelength under which the scales are considered as small (defined by the user) and λ_{min} the minimum wavelength to be propagated in the model.

Capdeville (2021) recently applied homogenization to a point-source. The idea is to filter the high-wavenumber content of a point-source in order to compute a smooth effective source. Point-source is a classic model for representing the mechanism of an earthquake. There are two major types of point-sources: the single force \mathbf{f} (Eq. (11)) and the couple force or stress source $\boldsymbol{\tau}$ (Eq. (12)) (?):

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{F}\delta(\mathbf{x} - \mathbf{x}_0)g(t), \quad (11)$$

$$\boldsymbol{\tau}(\mathbf{x}, t) = \mathbf{M}\delta(\mathbf{x} - \mathbf{x}_0)g(t), \quad (12)$$

with \mathbf{F} a force vector, \mathbf{M} a moment tensor, g a source time function, and δ the Dirac function. The homogenization aims at separating the small scales (hereafter denoted by the \mathbf{y} variable) and the large scales (hereafter denoted by the \mathbf{x} variable) within both the displacement and the stress fields. In other words, the homogenization looks for the solution of the elastodynamic problem in the following form: $\mathbf{u}(\mathbf{x}, \mathbf{y}, t)$ and $\boldsymbol{\sigma}(\mathbf{x}, \mathbf{y}, t)$. To do so, the method relies on asymptotic expansion of the two fields:

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{u}_0(\mathbf{x}, \mathbf{y}, t) + \epsilon_0\mathbf{u}_1(\mathbf{x}, \mathbf{y}, t) + \epsilon_0^2\mathbf{u}_2(\mathbf{x}, \mathbf{y}, t) + \dots, \quad (13)$$

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \boldsymbol{\sigma}_0(\mathbf{x}, \mathbf{y}, t) + \epsilon_0\boldsymbol{\sigma}_1(\mathbf{x}, \mathbf{y}, t) + \epsilon_0^2\boldsymbol{\sigma}_2(\mathbf{x}, \mathbf{y}, t) + \dots, \quad (14)$$

where the two space variables relate through the scaling parameter ϵ_0 : $\mathbf{y} = \frac{\mathbf{x}}{\epsilon_0}$. In what follows, we focus on the couple force, the derivation for the single force case being similar (Capdeville 2021).

First of all, it can be shown (Appendix A) that a Dirac function can be decomposed as

$$\delta(\mathbf{x}) = \frac{1}{\epsilon_0}(I - F)(\delta)(\mathbf{y}) + F^{\epsilon_0}(\delta)(\mathbf{x}), \quad (15)$$

where F^{ϵ_0} is the low pass-filter for the \mathbf{x} variable (muting the wavenumbers larger than $\frac{1}{\lambda_0}$) and F the low pass-filter for the \mathbf{y} variable (muting the wavenumbers larger than $\frac{1}{\lambda_{min}}$). As a consequence, we can assume that the couple force can be expanded as

$$\boldsymbol{\tau}(\mathbf{x}, t) = \frac{1}{\epsilon_0}\boldsymbol{\tau}_{-1}(\mathbf{x}, \mathbf{y})g(t) + \boldsymbol{\tau}_0(\mathbf{x})g(t). \quad (16)$$

To obtain the coefficients $\boldsymbol{\tau}_{-1}$ and $\boldsymbol{\tau}_0$, we introduce the expansions (13), (14) and (16) into the wave equation and Hooke's law, which leads to

$$\rho\ddot{\mathbf{u}}_i - \nabla_x \cdot \boldsymbol{\sigma}_i - \nabla_y \cdot \boldsymbol{\sigma}_{i+1} = \mathbf{0}, \quad (17)$$

$$\boldsymbol{\sigma}_i = \mathbf{c} : \left(\nabla_x(\mathbf{u}_i) + \nabla_y(\mathbf{u}_{i+1}) \right) - \boldsymbol{\tau}_i, \quad (18)$$

with \mathbf{c} the elastic tensor and i denoting each order in the expansions. For sake of simplicity, we here assume that \mathbf{c} is smooth, i.e. $\mathbf{c} = \mathbf{c}(\mathbf{x})$ (see Capdeville 2021, sections 2.4 and 3.4, for a discussion on how handling small scale mechanical properties in source homogenization). Solving the system (17) and (18) for each i , we end up with

$$\mathbf{u}(\mathbf{x}, \mathbf{y}, t) = \mathbf{u}_0(\mathbf{x}, t) + \boldsymbol{\theta}(\mathbf{y})g(t) + \mathcal{O}(\epsilon_0). \quad (19)$$

$\boldsymbol{\theta}(\mathbf{y})$ is called the corrector; it contains only small scales. Eq. (19) shows that the wavefield can be decomposed into a smooth wavefield \mathbf{u}_0 which propagates in the medium and a high-wavenumber part $\boldsymbol{\theta}$ which is not constrained by the dispersion relation. The corrector indeed satisfies the static partial differential equation

$$\nabla_y \cdot \mathbf{c} : \nabla_y(\boldsymbol{\theta}) = \nabla_y \cdot \boldsymbol{\tau}_{-1}. \quad (20)$$

with periodic boundary conditions. As for the zeroth-order displacement $\mathbf{u}^0(\mathbf{x}, t)$, it is the solution of the effective elastodynamic problem

$$\rho\ddot{\mathbf{u}}_0 - \nabla_x \cdot \boldsymbol{\sigma}_0 = \mathbf{0}, \quad (21)$$

$$\boldsymbol{\sigma}_0 = \mathbf{c} : \nabla_x(\mathbf{u}_0) - \boldsymbol{\tau}_0 g(t). \quad (22)$$

The procedure to determine the coefficients $\boldsymbol{\tau}_0$ and $\boldsymbol{\tau}_{-1}$ in practice is to solve Eq. (20) assuming a starting stress $\boldsymbol{\tau}_{-1,s} = \mathbf{M}\delta(\mathbf{y})$ and that all the variations of \mathbf{c} are small so that $\mathbf{c} = \mathbf{c}(\mathbf{y})$ (whatever the true distribution of scales in \mathbf{c} is). Doing so, we obtain the starting corrector $\boldsymbol{\theta}_s(\mathbf{y})$. By filtering it, the large and small scales can be separated within two quantities: the corrector $\boldsymbol{\theta}$ for the small scales and the source potential $\boldsymbol{\psi}$ for the large scales:

$$\boldsymbol{\theta}(\mathbf{y}) = (I - F)\boldsymbol{\theta}_s(\mathbf{y}), \quad (23)$$

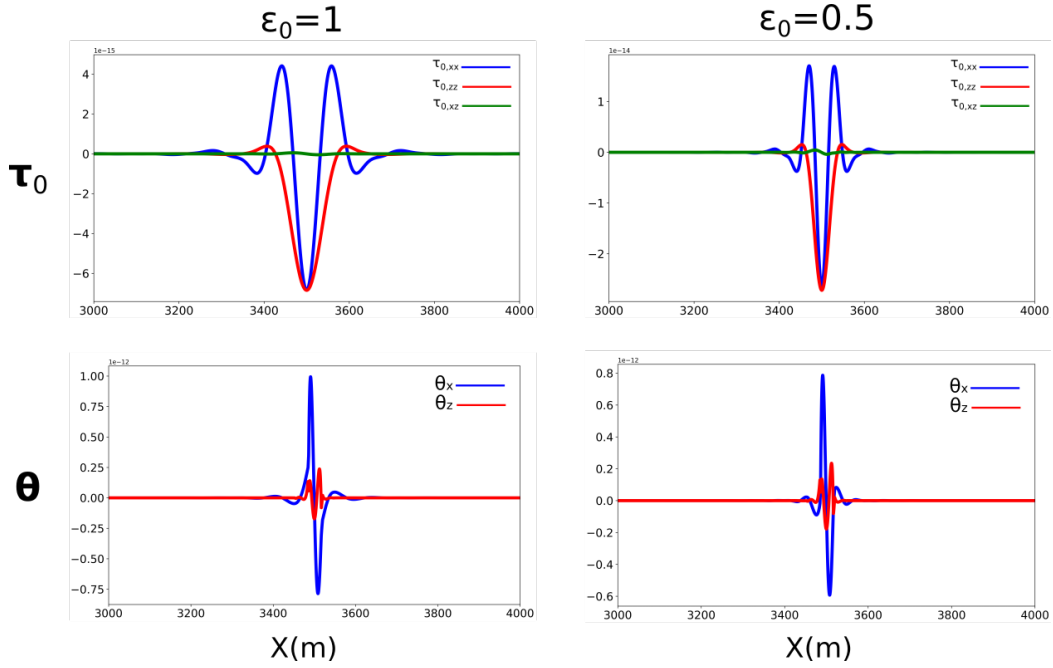


Figure 6. Results of point-source homogenization for $\mathbf{M} = \mathbf{I}$ in a 2D homogeneous medium using two values of ϵ_0 . The top row represents the effective couple force $\boldsymbol{\tau}_0$ and the bottom row shows the corrector $\boldsymbol{\theta}$ along the x-axis.

$$\boldsymbol{\psi}(\mathbf{x}) = F^{\epsilon_0} \left[\boldsymbol{\theta}_s \left(\frac{\mathbf{x}}{\epsilon_0} \right) \right]. \quad (24)$$

From the potential and the corrector, $\boldsymbol{\tau}_0$ and $\boldsymbol{\tau}_{-1}$ are built:

$$\boldsymbol{\tau}_{-1}(\mathbf{x}, \mathbf{y}) = \mathbf{c}(\mathbf{x}) : \nabla_{\mathbf{y}}(\boldsymbol{\theta}(\mathbf{y})), \quad (25)$$

$$\boldsymbol{\tau}_0(\mathbf{x}) = \mathbf{c}(\mathbf{x}) : \nabla_{\mathbf{x}}(\boldsymbol{\psi}(\mathbf{x})). \quad (26)$$

From $\boldsymbol{\tau}_0$, we easily derive the effective force \mathbf{f}_0 to apply in the effective wave propagation problem (21) and (22) to get \mathbf{u}_0 . Introducing (26) into (22) and then (22) into (21), we indeed have

$$\rho \ddot{\mathbf{u}}_0 - \nabla \cdot \mathbf{c} : \nabla(\mathbf{u}_0) = \mathbf{f}_0 \quad (27)$$

with $\mathbf{f}_0(\mathbf{x}, t) = -\nabla \cdot \boldsymbol{\tau}_0(\mathbf{x})g(t)$.

Fig. 6 shows $\boldsymbol{\tau}_0$ and $\boldsymbol{\theta}$ for $\mathbf{M} = \mathbf{I}$ in a 2D homogeneous case. We test two values for ϵ_0 : $\epsilon_0 = 1$ to see what happens when filtering at the minimum wavelength and $\epsilon_0 = 0.5$ because it corresponds to the diffraction limit. The top row shows that the spatial spread of the effective couple force decreases with ϵ_0 . Furthermore, we see that the potential $\boldsymbol{\psi}$ holds the large scales whereas the corrector $\boldsymbol{\theta}$ contains the small scales. Fig. 7 shows that the corrector contributes to the displacement only near the source. When this latter is emitting (Fig. 7, top row), the corrector has a major role because it contains the discontinuity at the source. When the source stops emitting (Fig. 7, bottom row), there is no difference between the reference and the zeroth-order displacement. In the next section, we link this theory to the focal spot of the time reversal.

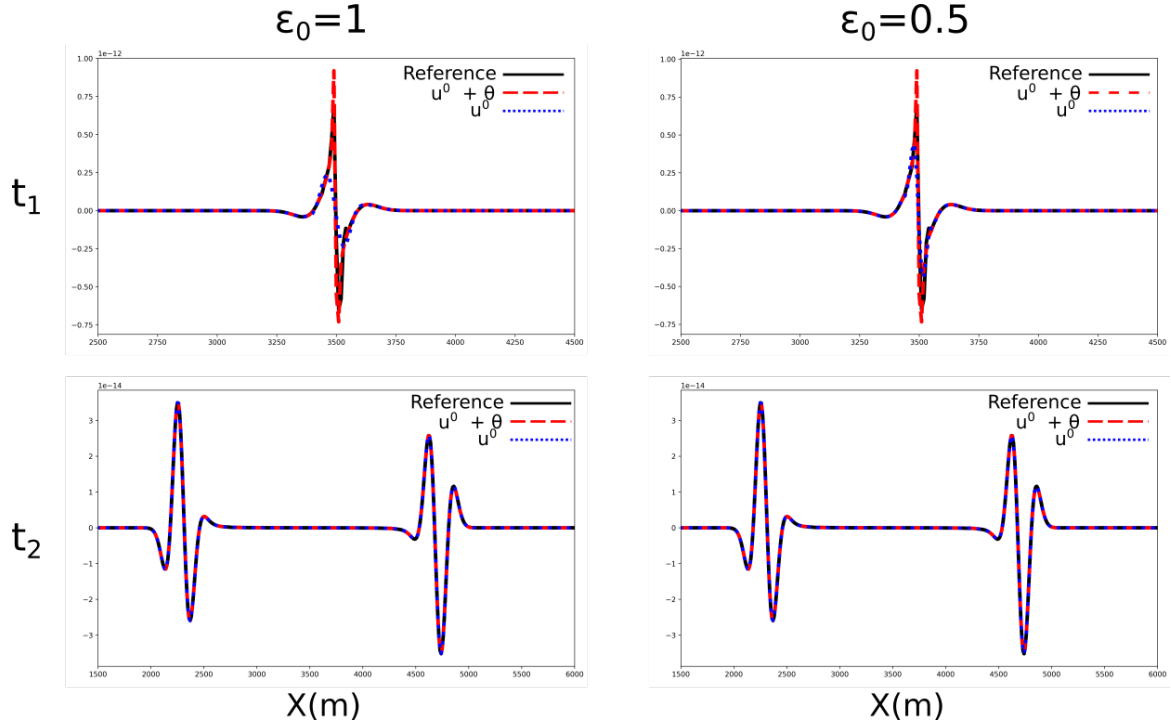


Figure 7. Comparison of the displacement generated by the homogenized point-source with a reference displacement computed in a fine mesh at the source. Two values of ϵ_0 are shown. The curves represent the x-component of the displacement at two different time steps: t_1 is a time near the source burst and t_2 is a time when the source has stopped emitting.

2.2 Application to time reversal

In the case of a stress source (12), the forward displacement \mathbf{u} is the response to the body force

$$\mathbf{f}(\mathbf{x}, t) = -\nabla \cdot \boldsymbol{\tau}(\mathbf{x}, t) = -\mathbf{M} \cdot \nabla \delta(\mathbf{x} - \mathbf{x}_0)g(t). \quad (28)$$

As we just saw, this displacement can be decomposed following Eq. (19). By reversing it in time, we can write the first term of the right-hand side of Eq. (9) as

$$\mathbf{u}(\mathbf{x}, \mathbf{y}, T - t) = \mathbf{u}_0(\mathbf{x}, T - t) + \boldsymbol{\theta}(\mathbf{y})g(T - t) + \mathcal{O}(\epsilon_0). \quad (29)$$

For the sink displacement $\mathbf{u}^s(\mathbf{x}, t)$, the body force is

$$\mathbf{f}^s(\mathbf{x}, t) = -\mathbf{M} \cdot \nabla \delta(\mathbf{x} - \mathbf{x}_0)g(T - t). \quad (30)$$

Using decomposition (19) again, we can then write the second term of the right-hand of Eq. (9) as

$$\mathbf{u}^s(\mathbf{x}, \mathbf{y}, t) = \mathbf{u}_0^s(\mathbf{x}, t) + \boldsymbol{\theta}^s(\mathbf{y})g(T - t) + \mathcal{O}^s(\epsilon_0). \quad (31)$$

Introducing (29) and (31) into Eq. (9), the time reversal displacement can be expressed as

$$\mathbf{u}^{tr}(\mathbf{x}, t) = \mathbf{u}_0(\mathbf{x}, T - t) + \boldsymbol{\theta}(\mathbf{y})g(T - t) + \mathcal{O}(\epsilon_0) - \mathbf{u}_0^s(\mathbf{x}, t) - \boldsymbol{\theta}^s(\mathbf{y})g(T - t) - \mathcal{O}^s(\epsilon_0). \quad (32)$$

The forces which generate the forward and the sink displacement (i.e. Eq. (28) and Eq. (30)) have the

exact same spatial part, so the two correctors $\boldsymbol{\theta}$ and $\boldsymbol{\theta}^s$ are equal. Therefore, Eq. (32) reduces to

$$\mathbf{u}^{tr}(\mathbf{x}, t) = \mathbf{u}_0(\mathbf{x}, T - t) - \mathbf{u}_0^s(\mathbf{x}, t) + \mathcal{O}'(\epsilon_0). \quad (33)$$

This last equation is the main result of this paper. It demonstrates the smoothness of the time reversal wavefield, including the focal spot, in a new way. Eq. (33) indeed shows that the back-propagated wavefield can be seen as the sum of two low-wavenumber fields resulting from the homogenization of the original point-source. In other words, the homogenized point-source is the equivalent force for producing the time reversal wavefield, including the focal spot.

To better understand how Eq. (33) works, we compute \mathbf{u}_0 and \mathbf{u}_0^s in the simple 2D case presented in Section 1.2. The point-source is homogenized using $\epsilon_0 = 0.5$ and all the wave simulations are performed in the same mesh. Fig. 8, left column, shows three snapshots of \mathbf{u}_0 , \mathbf{u}_0^s , $\mathbf{u}_0 - \mathbf{u}_0^s$ and \mathbf{u}^{tr} along a horizontal line. At $t < T - 2\tau_b$ (top row), the sink did not start acting yet, so $\mathbf{u}_0^s(\mathbf{x}, t) = \mathbf{0}$ and the time reversal displacement $\mathbf{u}^{tr}(\mathbf{x}, t)$ is equal to the zeroth-order forward displacement $\mathbf{u}_0(\mathbf{x}, T - t)$. At $t > T$ (bottom row), $\mathbf{u}_0(\mathbf{x}, T - t)$ vanishes and the time reversal displacement $\mathbf{u}^{tr}(\mathbf{x}, t)$ is the opposite of the zeroth-order sink displacement $-\mathbf{u}_0^s(\mathbf{x}, t)$. The middle row shows that in the time interval $T - 2\tau_b < t < T$, which corresponds to the focalization, the time reversal displacement is a combination of the two non-zero displacements \mathbf{u}_0 and \mathbf{u}_0^s . This numerical experiment illustrates Eq. (33). Moreover, it allows us defining the focal spot as an extended object both in space and time: the focal spot begins when the displacement is different from the forward displacement and ends when the displacement is the opposite of the sink displacement.

3 APPLICATIONS

To further illustrate the mathematical result of the previous part, we apply it to two more complex synthetic cases. In the first one, we add an S-wave by taking $\mathbf{M} \neq \mathbf{I}$; in the second one, we model an extended source using multiple point-sources. The elastic properties of the medium and the position of the time reversal mirror remain unchanged.

3.1 Model with both P- and S-waves

To add an S-wave in our numerical simulation, we input the following moment tensor:

$$\mathbf{M} = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}. \quad (34)$$

The source time function $g(t)$ is a Ricker wavelet with a dominant frequency of 10 Hz, leading to a minimum wavelength of 66 m.

Fig. 9 compares the backward simulation \mathbf{u}^{tr} to the forward simulation \mathbf{u} (see the two wavefields on the left-hand side and the curves in the central column). As expected, the introduction of an S-wave does not alter the conclusions we drew in Section 1.2. The time reversal still succeeds in reconstructing the far field (e.g., at t_3 and t_4). On the contrary, in the focalization time interval (e.g., at t_2), the

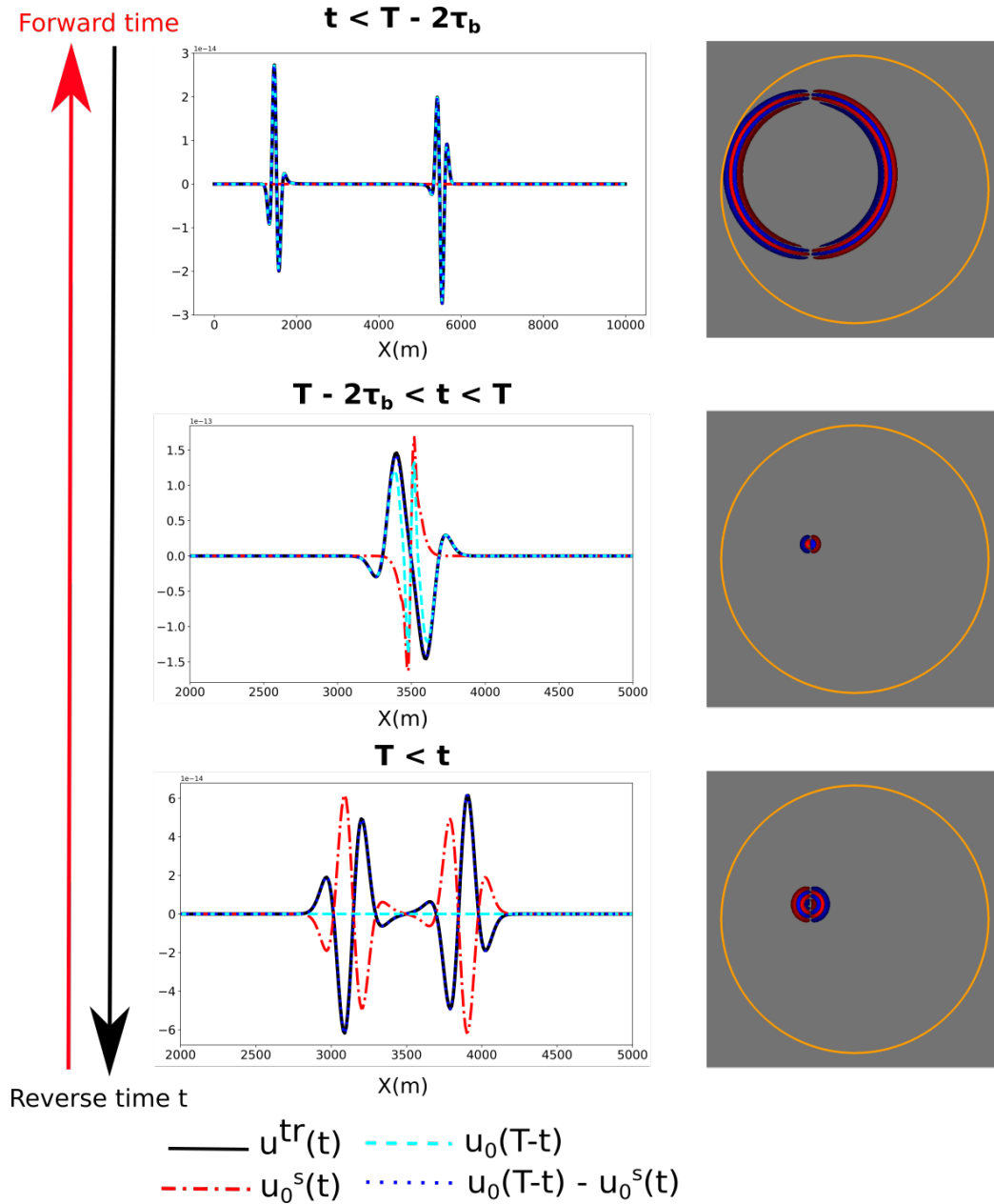


Figure 8. Comparison of the time reversal displacement u^{tr} to the difference between the two zeroth-order displacements u_0 and u_0^s . The curves (left) and the time reversal wavefield (right) represent the x-component at three different reverse time t . T is the forward simulation duration and τ_b is the centroid time of the source.

backward displacement does not match the forward displacement (central column) even if the two wavefields may look similar at a glance (left columns). The forward displacement indeed shows a space discontinuity whereas the backward displacement remains smooth. The bottom row of the figure (t_1) also shows this discrepancy, and it illustrates what happens when the sink term is not implemented: the backward wavefield diverges from the source position whereas the forward wavefield reduces to a point before vanishing.

The curves in the right-hand side column of Fig. 9 show a comparison between the backward

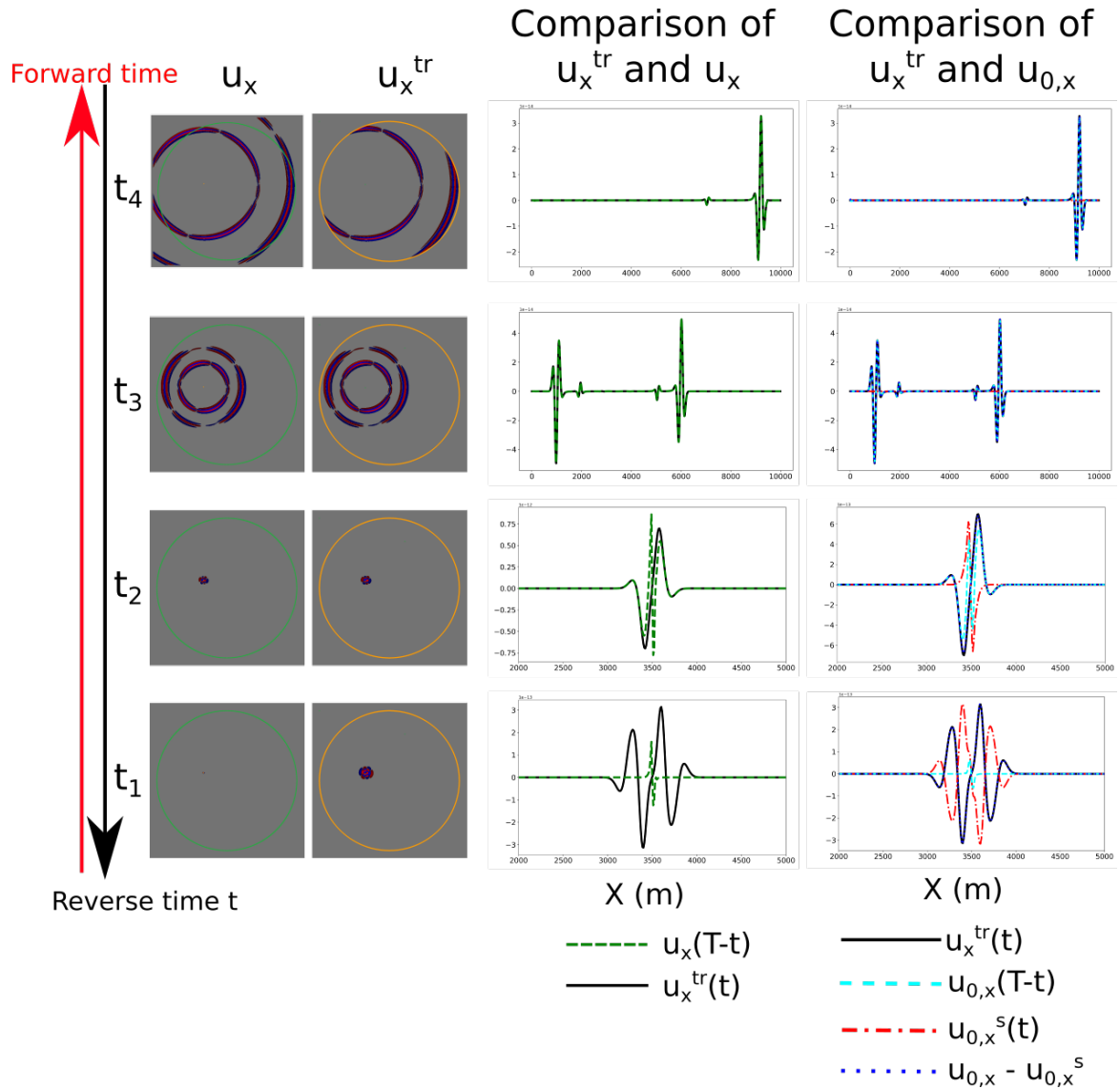


Figure 9. Comparison of the x-component of the backward displacement u^{tr} to the forward displacement u and to the point-source homogenization results u_0 and u_0^s . The two left-hand side columns show the forward and backward wavefields at four different time steps. The orange dots represent the sources and the green dots represent the receivers. The central column shows the forward and backward displacements along a line of receivers at forward source depth. The right-hand side column shows the two zeroth-order displacements u_0 and u_0^s , and compares the backward displacement to the difference $u_0(x, T-t) - u_0^s(x, t)$, which is the main result of the paper (Eq. (33)).

displacement u^{tr} and the zeroth-order displacement u_0 and u_0^s computed from the homogenized point-source. We observe that $u_0(x, T-t) - u_0^s(x, t)$ matches $u^{tr}(x, t)$ everywhere in space and time, which confirms Eq. (33). The same conclusion can be drawn using the z-component.

3.2 Involving an extended sources

Extended sources can represent finite fault ruptures. They can be discretized in space using elementary surfaces dS_i at position \mathbf{x}_i (e.g., Yue & Lay 2020). In this case, the driving force can be written as

$$\mathbf{f}(\mathbf{x}, t) = \sum_i \mathbf{f}_i(\mathbf{x}, t), \quad (35)$$

$$\mathbf{f}_i(\mathbf{x}, t) = \mathbf{M}_i \nabla \delta(\mathbf{x} - \mathbf{x}_i) g_i(t) dS_i, \quad (36)$$

where \mathbf{M}_i is the moment tensor of the i^{th} point-source positioned at \mathbf{x}_i (?).

To derive our main result (33) in the case of an extended source discretized as multiple point-sources, we rewrite Eq. (9):

$$\mathbf{u}^{tr}(\mathbf{x}, t) = \sum_i \mathbf{u}_i(\mathbf{x}, T - t) - \sum_i \mathbf{u}_i^s(\mathbf{x}, t), \quad (37)$$

where \mathbf{u}_i and \mathbf{u}_i^s are the response to \mathbf{f}_i and $\mathbf{f}_i^s = \mathbf{f}_i(\mathbf{x}, T - t)$, respectively. Using the decomposition (19) for each of these displacements, Eq. (37) becomes

$$\begin{aligned} \mathbf{u}^{tr}(\mathbf{x}, t) = & \sum_i \left(\mathbf{u}_{0,i}(\mathbf{x}, T - t) + \boldsymbol{\theta}_i(\mathbf{y}) g_i(T - t) \right) \\ & - \sum_i \left(\mathbf{u}_{0,i}^s(\mathbf{x}, t) + \boldsymbol{\theta}_i^s(\mathbf{y}) g_i(T - t) \right) + \mathcal{O}(\epsilon_0). \end{aligned} \quad (38)$$

As in the single point-source case, the correctors $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_i^s$ are equal for each point-source independently. That leads to

$$\mathbf{u}^{tr}(\mathbf{x}, t) = \sum_i \left(\mathbf{u}_{0,i}(\mathbf{x}, T - t) \right) - \sum_i \left(\mathbf{u}_{0,i}^s(\mathbf{x}, t) \right) + \mathcal{O}(\epsilon_0). \quad (39)$$

This last equation is equivalent to (33) in the case of multiple point-sources.

To illustrate Eq. (39), we perform a numerical experiment in the 10 km x 10 km homogeneous medium that we used in the previous sections. In this medium, we insert a horizontal finite fault discretized using four point-sources. We choose the four corresponding moment tensors to be equal to each other, i.e. $\mathbf{M}_i = \mathbf{M} \forall i$, where \mathbf{M} is chosen arbitrarily:

$$\mathbf{M} = \begin{bmatrix} 0.56 & 1 \\ 1 & 2.56 \end{bmatrix}. \quad (40)$$

We take the four source time functions g_i having the same Ricker shape, but we shift them according to the centroid time τ_i so that the rupture propagates toward the increasing point numbers: $\tau_1 < \tau_2 < \tau_3 < \tau_4$. Moreover, we want the fault segment displacements to overlap in time, so we choose $\Delta\tau = \tau_{i+1} - \tau_i = T_w/2$, with T_w the time support of the Ricker wavelet. Assuming a rupture speed equal to V_s , the length of the segments is calculated using $dS = V_s \Delta\tau$.

Fig. 10 shows that the time reversal focal spot obtained in the case of an extended source is more complex than the one obtained in the single point-source case. Despite this complexity, our theory holds: in the right-hand side column of the figure, Eq. (39) is confirmed at three different times in the focalization time interval. The only difference from the single point-source case is that the computations of $\mathbf{u}_0(\mathbf{x}, t)$ and $\mathbf{u}_0^s(\mathbf{x}, t)$ have to be carried out separately: \mathbf{u}_0 is computed based on the forward source

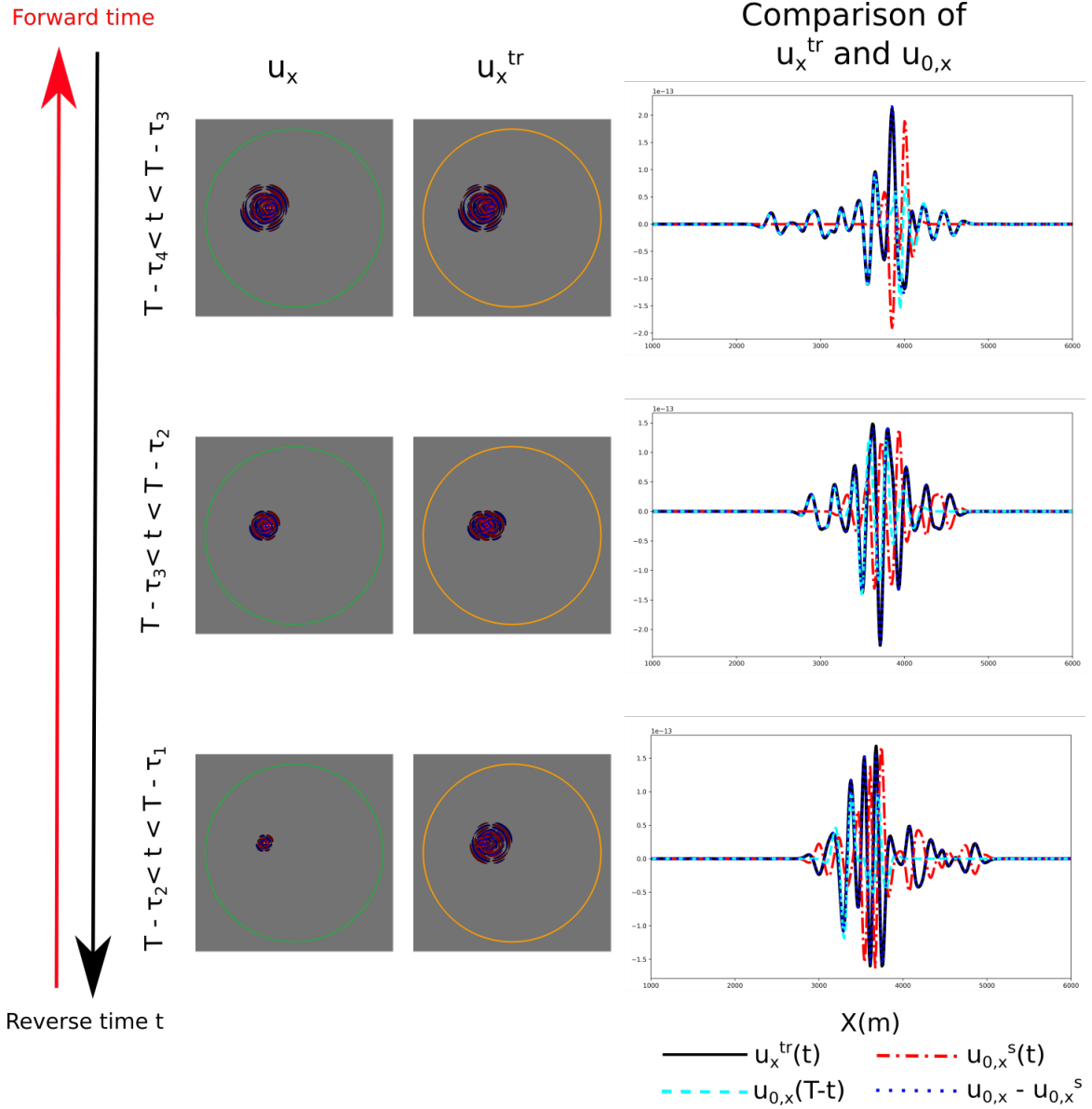


Figure 10. Left: The forward and backward wavefields in the case of an extended source made of four point-sources. Three times in the focalization time interval are presented. Right: Comparison between the backward displacement u^{tr} and the difference between the two zeroth-order displacements u_0 and u_0^s obtained from the homogenization the four point-sources. The black line is the backward u_x^{tr} , the read dashed one is the homogenized sink displacement $u_{0,x}^s$, the cyan dashed one is the homogenized forward displacement $u_{0,x}$ and the blue dashed one is the difference $u_{0,x} - u_{0,x}^s$.

order (i.e. (\mathbf{x}_1, τ_1) , (\mathbf{x}_2, τ_2) , ..., (\mathbf{x}_n, τ_n)) whereas u_0^s is computed in the reverse order of the source position (i.e. (\mathbf{x}_n, τ_1) , $(\mathbf{x}_{n-1}, \tau_2)$, ..., (\mathbf{x}_1, τ_n)).

4 DISCUSSION & CONCLUSION

As described by many works in the past, the time reversal displacement is shown to be smooth as compared to the displacement generated by an earthquake which presents a spatial discontinuity at the source location. Relying on the representation theorem and the point-source homogenization theory, we demonstrated a direct relation between the time reversal wavefield and the displacement generated by a homogenized point-source (Eq. (33)). By homogenized point-source, we mean the smooth effective source derived from the homogenization theory developed by Capdeville (2021). The relation is satisfied at all propagation times, including the focalization, which yields a new interpretation of the focal spot: this latter can be seen as the sum of two zeroth-order displacements generated by the smooth effective source.

Whereas considerable efforts have been made for spiking the focal spot (e.g., Kiser et al. 2011; Fukahata et al. 2014; Douma & Snieder 2014; Anderson et al. 2015; Bazargani & Snieder 2016; Nakata & Beroza 2016), our work leaves it intact and proposes to upscale the source model in order to derive smooth displacements that fit the focal spot. Such a process requires to solve the static equation (20), but then one no longer needs any small elements at the source location when simulating wave propagation numerically because the effective source is free of small scales. We did not illustrate this computational gain in the present paper; we rather used the same mesh for all the simulations to avoid possible numerical bias and make our comparisons as accurate as possible.

Once the upscaling process is known, downscaling the focal spot can be considered to estimate source model parameters. Adjoint tomography has been proved to successfully yield the moment tensor (Liu et al. 2004). As for the location, the inverse problem is highly non-linear (Kim et al. 2011), which makes adjoint tomography computationally expensive. Moreover, this technique is deterministic, requiring an initial model and leading to a single solution which can correspond to a local minimum of the misfit function. Our result opens the path to a Bayesian inversion that could overcome these limitations. In such an inversion, the focal spot would be the data and the point-source homogenization would be the first step of the forward modeling operation. The second step would be the computation of a few time-steps of the two zeroth-order wavefields, which consists in just one simulation in the case of a single point-source and in two simulations in the case of multiple point-sources. The main advantage of using a homogenized point-source as compared to a classic point-source is that the mesh does not have to be fine, so the simulation cost is low. The downscaling procedure we outline here is inspired by what Hedjazian et al. (2021) developed to estimate small scale elastic properties.

It is worth noting that adjoint tomography also turns receivers into sources to backpropagate signals (e.g., ?). Interestingly, Kim et al. (2011) use a time reversal simulation as an adjoint wavefield to show how this latter captures the characteristics of the source. From a more practical point of view, time reversal can be used to avoid the storage of the whole forward simulation. Recording the forward wavefield along a close surface indeed allows to reconstruct it during the adjoint simulation so that the two wavefields can be correlated on the fly to efficiently compute the gradient of the misfit function. Doing so, the inverse problem can be solved without storing a large amount of data at each iteration.

Our theoretical result (33) was illustrated using numerical experiments in a simple homogeneous medium, but it is also valid in heterogeneous media. The main limitation of our work actually is that it relies on a perfect mirror which does not exist in practical geoscience applications. This paper is purely theoretical, and further investigations are required to generalize our result to more realistic setting and to understand the focal spot when a partial mirror is involved (e.g., Cassereau & Fink 1993; Li & van der Baan 2016).

ACKNOWLEDGMENTS

This work was performed in the frame of the RING project (<http://ring.georessources.univ-lorraine.fr/>) at Université de Lorraine. We acknowledge the *Ministère de la Recherche et de l'Enseignement Supérieur*, the *Région Grand Est* and the *FEDER* for the PhD scholarship of Zoé Renat. The authors thank Guillaume Caumon for his help and advices during the project. Financial supports from the RING-Gocad consortium and the ANR projects ANR-10-BLAN-613 and ANR-16-CE31-0022- 01 also enabled this work.

REFERENCES

- Anderson, B. E., Douma, J., Ulrich, T., & Snieder, R., 2015. Improving spatio-temporal focusing and source reconstruction through deconvolution, *Wave Motion*, **52**, 151–159.
- Bai, T., Zhu, T., & Tsvankin, I., 2019. Attenuation compensation for time-reversal imaging in VTI media, *Geophysics*, **84**(4), C205–C216.
- Bazargani, F. & Snieder, R., 2016. Optimal source imaging in elastic media, *Geophys. J. Int.*, **204**(2), 1134–1147.
- Borcea, L., Papanicolaou, G., & Tsogka, C., 2003. Theory and applications of time reversal and interferometric imaging, *Inverse Probl.*, **19**(6), S139–S164.
- Cao, W., Hanafy, S. M., Schuster, G. T., Zhan, G., & Boonyasiriwat, C., 2012. High-resolution and super stacking of time-reversal mirrors in locating seismic sources, *Geophys. Prospect.*, **60**(1), 1–17.
- Capdeville, Y., 2021. Homogenization of seismic point and extended sources, *Geophys. J. Int.*, **226**(2), 1390–1416.
- Capdeville, Y., Guillot, L., & Marigo, J., 2010. 2-D non-periodic homogenization to upscale elastic media for P-SV waves, *Geophys. J. Int.*, **182**, 903–922.
- Capdeville, Y., Cupillard, P., & Singh, S., 2020. Chapter six - an introduction to the two-scale homogenization method for seismology, in *Machine Learning in Geosciences*, vol. 61 of **Advances in Geophysics**, pp. 217–306, eds Moseley, B. & Krischer, L., Elsevier.
- Cassereau, D. & Fink, M., 1992. Time-reversal of ultrasonic fields - Part iii: Theory of the closed time-reversal cavity, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **39**(5), 579–592.
- Cassereau, D. & Fink, M., 1993. Focusing with plane time-reversal mirrors: an efficient alternative to closed cavities, *J. Acoust. Soc. Am.*, **94**(4), 2373–2386.
- Chambers, K., Dando, B. D., Jones, G. A., Velasco, R., & Wilson, S. A., 2014. Moment tensor migration imaging, *Geophys. Prospect.*, **62**(4), 879–896.

- Cros, E., Roux, P., Vandemeulebrouck, J., & Kedar, S., 2011. Locating hydrothermal acoustic sources at Old Faithful Geyser using Matched Field Processing, *Geophys. J. Int.*, **187**(1), 385–393.
- Cupillard, P. & Capdeville, Y., 2018. Non-periodic homogenization of 3-D elastic media for the seismic wave equation, *Geophys. J. Int.*, **213**(2), 983–1001.
- de Rosny, J. & Fink, M., 2002. Overcoming the diffraction limit in wave physics using a time-reversal mirror and a novel acoustic sink, *Phys. Rev. Lett.*, **89**(12), 124301.
- Douma, J. & Snieder, R., 2014. Focusing of elastic waves for microseismic imaging, *Geophys. J. Int.*, **200**(1), 390–401.
- Douma, J., Niederleithinger, E., & Snieder, R., 2015. Locating events using time reversal and deconvolution: Experimental application and analysis, *J. Nondestruct. Eval.*, **34**(1).
- Draeger, C. & Fink, M., 1997. One-Channel Time Reversal of Elastic Waves in a Chaotic 2D-Silicon Cavity, *Phys. Rev. Lett.*, **79**, 407–410.
- Finger, C. & Saenger, E. H., 2021. Determination of the time-dependent moment tensor using time reverse imaging, *Geophysics*, **86**(2), KS63–KS77.
- Fink, M., 1992. Time reversal of ultrasonic fields - Part i: Basic principles, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **39**(5), 555–566.
- Fink, M., 2006. Time-reversal acoustics in complex environments, *Geophysics*, **71**(4), SI151–SI164.
- Fink, M., 2015. Acoustic imaging with time reversal methods: From medicine to ndt, in *AIP Conference Proceedings*, vol. 1650, pp. 13–23, American Institute of Physics.
- Fukahata, Y., Yagi, Y., & Rivera, L., 2014. Theoretical relationship between back-projection imaging and classical linear inverse solutions, *Geophys. J. Int.*, **196**(1), 552–559.
- Gajewski, D. & Tessmer, E., 2005. Reverse modelling for seismic event characterization, *Geophys. J. Int.*, **163**(1), 276–284.
- Garmany, J., 1979. On the inversion of travel times, *Geophys. Res. Lett.*, **6**(4), 277–279.
- Gelius, L.-J. & Asgedom, E., 2011. Diffraction-limited imaging and beyond - the concept of super resolution, *Geophys. Prospect.*, **59**(3), 400–421.
- Givoli, D. & Turkel, E., 2012. Time reversal with partial information for wave refocusing and scatterer identification, *Comput. Methods Appl. Mech. Eng.*, **213–216**, 223–242.
- Grigoli, F., Cesca, S., Krieger, L., Kriegerowski, M., Gammaldi, S., Horalek, J., Priolo, E., & Dahm, T., 2016. Automated microseismic event location using Master-Event Waveform Stacking, *Sci. Rep.*, **6**(1), 25744.
- Hampson, D. & Russell, B., 1984. First-break interpretation using generalized linear inversion, in *SEG Tech. Progr. Expand. Abstr. 1984*, pp. 532–534, Society of Exploration Geophysicists.
- Hedjazian, N., Capdeville, Y., & Bodin, T., 2021. Multiscale seismic imaging with inverse homogenization, *Geophys. J. Int.*, **226**(1), 676–691.
- Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E., 2007. Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array, *J. Geophys. Res.*, **112**(B11).
- Kao, H. & Shan, S.-J., 2004. The Source-Scanning Algorithm: mapping the distribution of seismic sources in time and space, *Geophys. J. Int.*, **157**(2), 589–594.
- Kennett, B. L. & Sambridge, M. S., 1992. Earthquake location - genetic algorithms for teleseisms, *Phys. Earth Planet. Inter.*, **75**, 103–110.
- Kim, Y., Liu, Q., & Tromp, J., 2011. Adjoint centroid-moment tensor inversions, *Geophys. J. Int.*, **186**(1),

264–278.

- Kiser, E., Ishii, M., Langmuir, C. H., Shearer, P. M., & Hirose, H., 2011. Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases, *J. Geophys. Res.*, **116**(B6), B06310.
- Komatitsch, D. & Tromp, J., 1999. Introduction to the Spectral Element Method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139**, 806–822.
- Komatitsch, D. & Vilotte, J. P., 1998. The Spectral Element Method: an efficient tool to simulate the seismic response of 2d and 3d geological structures, *Bull. Seism. Soc. Am.*, **88**, 368–392.
- Kremers, S., Fichtner, A., Brietzke, G. B., Igel, H., Larmat, C., Huang, L., & Käser, M., 2011. Exploring the potentials and limitations of the time-reversal imaging of finite seismic sources, *Solid Earth*, **2**(1), 95–105.
- Kuperman, W., Hodgkiss, W. S., Song, H. C., Akal, T., Ferla, C., & Jackson, D. R., 1998. Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror, *J. Acoust. Soc. Am.*, **103**(1), 25–40.
- Larmat, C., Montagner, J. P., Fink, M., Capdeville, Y., Tourin, A., & Clévéde, E., 2006. Time-reversal imaging of seismic sources and application to the great Sumatra earthquake, *Geophys. Res. Lett.*, **33**(19), 0–3.
- Larmat, C., Tromp, J., Liu, Q., & Montagner, J. P., 2008. Time reversal location of glacial earthquakes, *J. Geophys. Res.*, **113**(9), 1–9.
- Lorosey, G., de Rosny, J., Tourin, A., & Fink, M., 2007. Focusing beyond the diffraction limit with far-field time reversal, *Science*, **315**(5815), 1120–1122.
- Li, L., Tan, J., Schwarz, B., Staněk, F., Poiata, N., Shi, P., Diekmann, L., Eisner, L., & Gajewski, D., 2020. Recent advances and challenges of waveform-based seismic location methods at multiple scales, *Rev. Geophys.*, **58**(1), 1–47.
- Li, Z. & van der Baan, M., 2016. Microseismic event localization by acoustic time reversal extrapolation, *Geophysics*, **81**(3), KS123–KS134.
- Liu, Q., Polet, J., Komatitsch, D., & Tromp, J., 2004. Spectral-element moment tensor inversions for earthquakes in southern California, *Bull. Seism. Soc. Am.*, **94**(5), 1748–1761.
- Masson, Y. & Romanowicz, B., 2017. Fast computation of synthetic seismograms within a medium containing remote localized perturbations: a numerical solution to the scattering problem, *Geophys. J. Int.*, **208**(2), 674–692.
- Masson, Y., Cupillard, P., Capdeville, Y., & Romanowicz, B., 2014. On the numerical implementation of time-reversal mirrors for tomographic imaging, *Geophys. J. Int.*, **196**(3), 1580–1599.
- McMechan, G. A., 1982. Determination of source parameters by wavefield extrapolation, *Geophys. J. Int.*, **71**(3), 613–628.
- Nakahara, H. & Haney, M. M., 2015. Point spread functions for earthquake source imaging: An interpretation based on seismic interferometry, *Geophys. J. Int.*, **202**(1), 54–61.
- Nakata, N. & Beroza, G. C., 2016. Reverse time migration for microseismic sources using the geometric mean as an imaging condition, *Geophysics*, **81**(2), KS51–KS60.
- Rabinovich, D., Turkel, E., & Givoli, D., 2018. An augmented time reversal method for source and scatterer identification, *J. Comput. Phys.*, **375**, 99–119.
- Rietbrock, A. & Scherbaum, F., 1994. Acoustic imaging of earthquake sources from the Chalfant Valley, 1986, aftershock series, *Geophys. J. Int.*, **119**(1), 260–268.
- Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*,

40(3), 1–29.

Schuster, G. T., Yu, J., Sheng, J., & Rickett, J., 2004. Interferometric/daylight seismic imaging, *Geophys. J. Int.*, **157**(2), 838–852.

Schuster, G. T., Hanafy, S., & Huang, Y., 2012. Theory and feasibility tests for a seismic scanning tunnelling microscope, *Geophys. J. Int.*, **190**(3), 1593–1606.

Shi, P., Angus, D., Rost, S., Nowacki, A., & Yuan, S., 2019. Automated seismic waveform location using multichannel coherency migration (MCM)–I: theory, *Geophys. J. Int.*, **216**(3), 1842–1866.

Shustak, M. & Landa, E., 2017. Time reversal based detection of subsurface scatterers, in *SEG Tech. Progr. Expand. Abstr. 2017*, pp. 969–973, Society of Exploration Geophysicists.

Steiner, B., Saenger, E. H., & Schmalholz, S. M., 2008. Time reverse modeling of low-frequency microtremors: Application to hydrocarbon reservoir localization, *Geophys. Res. Lett.*, **35**(3), 1–7.

van Manen, D.-J., Curtis, A., & Robertsson, J. O., 2006. Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity, *Geophysics*, **71**(4), SI47–SI60.

Willacy, C., van Dedem, E., Minisini, S., Li, J., Blokland, J.-W., Das, I., & Droujinine, A., 2019. Full-waveform event location and moment tensor inversion for induced seismicity, *Geophysics*, **84**(2), KS39–KS57.

Yang, J. & Zhu, H., 2019. Locating and monitoring microseismicity, hydraulic fracture and earthquake rupture using elastic time-reversal imaging, *Geophys. J. Int.*, **216**(1), 726–744.

Yue, H. & Lay, T., 2020. Resolving complicated faulting process using multi-point-source representation: Iterative inversion algorithm improvement and application to recent complex earthquakes, *J. Geophys. Res.*, **125**(2), 1–24.

Zhao, M., Capdeville, Y., & Zhang, H., 2016. Direct numerical modeling of time-reversal acoustic subwavelength focusing, *Wave Motion*, **67**, 102–115.

Zhu, T., 2014. Time-reverse modelling of acoustic wave propagation in attenuating media, *Geophys. J. Int.*, **197**(1), 483–494.

APPENDIX A: TWO-SCALE DECOMPOSITION OF THE DIRAC FUNCTION

The goal of this appendix is to demonstrate Eq. (15). We first separate the small and large wavenumbers in a Dirac function using the low-pass filter F^{ϵ_0} :

$$\delta(\mathbf{x}) = F^{\epsilon_0}(\delta(\mathbf{x})) + (I - F^{\epsilon_0})(\delta(\mathbf{x})). \quad (\text{A.1})$$

Then we introduce F , the same filter than F^{ϵ_0} for the variable $\mathbf{y} = \frac{\mathbf{x}}{\epsilon_0}$. In other words, for any functions $h(\mathbf{x})$ and $\bar{h}(\mathbf{y}) = h(\epsilon_0\mathbf{y})$, we have

$$F^{\epsilon_0}(h(\mathbf{x})) = F(\bar{h}(\mathbf{x}/\epsilon_0)). \quad (\text{A.2})$$

Using this last equality in Eq. (A.1), it comes

$$\delta(\mathbf{x}) = F^{\epsilon_0}(\delta(\mathbf{x})) + (I - F)(\delta(\epsilon_0\mathbf{y})). \quad (\text{A.3})$$

Finally, we note the following property of the Dirac function:

$$\delta(\epsilon_0\mathbf{y}) = \frac{1}{\epsilon_0}\delta(\mathbf{y}). \quad (\text{A.4})$$

Introducing Eq. (A.4) in Eq. (A.3), we obtain Eq. (15).

Remerciements

Début 2008, lorsqu'il s'est agi de rédiger les remerciements de ma thèse, j'ai préféré prendre la route pour aller saluer quelques amis de province avant mon départ pour la Californie. Il n'y a donc, à ce jour, aucune trace écrite de la reconnaissance que je dois à ceux qui m'ont soutenu durant toutes ces années de recherche, de mon DEA en 2004 jusqu'à la rédaction du présent manuscrit. J'y remédie ici, *les larmes aux yeux, le cœur content*¹.

À Luis Rivera, dont la sympathie et la pédagogie exceptionnelle m'ont initié aux insoupçonnables subtilités de la sismologie quantitative. À son ami Jeroen Ritsema, pour m'avoir accueilli à l'IPGP et encadré avec chaleur et bienveillance.

À Jean-Pierre Vilotte et Gaetano Festa, pour la patience et la précision de leurs explications de la méthode des éléments spectraux.

À Geneviève Moguilny, sans doute l'une des personnes les plus remerciées de l'histoire de la recherche française, pour son aide en HPC, son manuel L^AT_EX et ses pots du vendredi. À Diego Mercerat, Éric Clévéde, Constanza Pardo, et tous les autres fidèles au rendez-vous.

À mes sœurs de thèse : Élise Delavaud, Alexandrine Gesret, Anne Bécel et Gaye Bayrakci.

À Jean-Paul Montagner, pour l'attention qu'il a toujours portée à ses étudiants. À Nikolai Shapiro, pour m'avoir soutenu au creux de la vague.

À Michel Campillo et Francisco Sánchez-Sesma, pour leurs rapports enthousiastes.

Au Real IPG (celui d'Alexandre Canitano, de Julien Roch, de Benoît Gotab, d'Eric Gayer, de Laurent Guillot), qui fut à l'Est parisien ce que le SC Bastia est à l'île de Beauté : une gloire locale faite de tacles piqués et de lourdes frappes non cadrées.

À Heiner Igel, pour l'émulation des *Initial Training Networks* SPICE et QUEST qu'il a portés et faits vivre. À tous les sismologues rencontrés dans ce cadre merveilleusement stimulant : Daniel Peter, Lapo Boschi, Ebru Bozdogan, Tarje Nissen-Meyer, Shравan Hanasoge, Amir Khan, Andreas Fichtner, Thomas Bodin, Pascal Audet... Tant de collaborations se sont construites lors des workshops annuels de ces deux ITN !

À Barbara Romanowicz, pour m'avoir accueilli en post-doc au sein du Berkeley Seismological Laboratory. Je ne cesse de mesurer l'énergie que Barbara déploie pour animer ce laboratoire et la richesse des discussions scientifiques que j'ai pu y avoir, avec Ved Lekic, Sanne Cottaar, Shan Dou, Scott French, Yder Masson, Huaiyu Yuan, Laurent Stehly, Aurélie Guilhem, Rob Porritt, Stéphanie Durand...

À Mathieu Hursin, mon premier pote outre-atlantique, mèche d'un feu d'artifice d'expatriés ayant joyeusement coloré mon Ouest américain, de LA à Yellowstone, de Vegas à Gualala, de Tahoe à Capitola : Gab, Raph, Raluca, Marine, Adrien, Clément, Thibault, Faiza, Pierre-Adrien, Eve, Cécile, François, Marion, Ivan, Matthias... Nous avons conquis ensemble bien plus qu'un continent !

À mes très chers amis géographes-marxistes-footeux : Erin, Alex, Stacey et Tony. Je regrette de ne plus vous voir...

À Daniela, Laurent, Joyce, Wanda et Kira, évidemment.

¹Trenet, C. (1947). Retour à Paris.

À Alexandre Fournier, pour m'avoir fait parvenir l'annonce du poste que j'occupe depuis près de dix ans.

À l'équipe RING et à tous ses satellites, pour m'avoir si chaleureusement accueilli et soutenu quand il le fallait. Pauline, Benoît, François, Julie, Christine, Denis, Julien, PH, Marianne, Marion, Yann, et tous les thésards passés et présents, tous les ingénieurs, toutes les secrétaires, vous constituez ou avez constitué un quotidien joyeux et bienveillant sans lequel je n'aurais su m'épanouir.

À Jean-Seb, pour son amitié, sa confiance, et son goût des choses bien faites qui alimente le mien propre. Aux amis musiciens de Nancy ou d'ailleurs avec lesquels j'ai partagé bien plus que des partitions et des concerts.

Au Wengé, dont la chaleur des cafés et des causeries a accompagné la rédaction de ce manuscrit. J'y suis assis à l'instant même.

Aux vieux copains du lycée. Aux presque-aussi-vieux copains de Strasbourg.

À Marie.

À mes grands-frères de science : Yann et Guillaume.

À ma belle-famille, qui a appris que *je dois bosser une heure ou deux* veut dire *j'en ai pour l'après-midi*.

À mes parents.

À Léonore.

À nos merveilleux enfants.

Merci !

Abstract

From high-frequency geophones used in near-surface exploration to broadband seismic stations deployed around the globe, hundreds of thousands of instruments record the seismic waves that propagate in our planet. The obtained data make it possible to study the mechanisms from which they originate (rupture of the earth's crust, magmatic upwelling, slow sliding of tectonic plates, glacier flow, ocean swell, etc) and to image the geological structures that the waves crossed in order to understand the dynamics of their formation, exploit their resources, monitor their evolution, etc. Whatever the scale and the object of study, simulating these data is necessary. Since the late 60s, numerical methods have been developed to model the complete waveform of seismic recordings. My past, present and future works are part of this perspective.

One of the most important issues for accurately modeling the propagation of seismic waves remains the correct account for realistic geological environments. These indeed contain multiple scales and often present a great geometrical complexity. In the last decade, the development of the non-periodic homogenization technique helped to deal with these aspects. Homogenization indeed makes it possible to calculate equivalent smooth media which can be easily integrated into numerical simulation methods. Most of all, it opens new paths towards the inversion of seismic data for the assessment of geological structures. These modeling and inversion aspects are here discussed and illustrated in different contexts (lithosphere imaging, subsurface exploration, fractured media, seismic risk assessment, source imaging) and for different types of data (earthquakes and noise correlations).

Résumé

Des géophones haute fréquence utilisés dans la prospection de la proche surface aux stations sismiques très large bande déployées tout autour du globe, des centaines de milliers d'instruments enregistrent les ondes sismiques qui se propagent dans notre planète. Les données obtenues permettent d'étudier les mécanismes qui en sont à l'origine (ruptures soudaines de la croûte terrestre, remontées magmatiques, glissements lents des plaques tectoniques, écoulements glaciaires, houle océanique, etc) et d'imager les structures géologiques traversées afin de comprendre la dynamique de leur formation, d'en exploiter les ressources, de surveiller leur évolution, etc. Quels que soient l'échelle et l'objet d'étude, simuler ces données est nécessaire. Depuis cinquante ans, des méthodes numériques se développent pour modéliser le plus précisément possible la forme d'onde complète des enregistrements sismiques. Mes travaux passés, présents et à venir s'inscrivent dans cette perspective.

L'un des verrous les plus importants pour modéliser de façon précise la propagation des ondes sismiques demeure la prise en compte correcte des environnements géologiques. Ceux-ci contiennent en effet de multiples échelles et présentent souvent une grande complexité géométrique. Au cours de la dernière décennie, le développement de l'homogénéisation non-périodique a contribué à gérer ces aspects. L'homogénéisation permet en effet de calculer des milieux lisses équivalents qui s'intègrent facilement aux méthodes de simulation numérique. Ce faisant, elle ouvre de nouvelles perspectives quant à l'inversion de données sismiques pour la reconnaissance des structures géologiques. Ces aspects de modélisation et d'inversion sont ici discutés et illustrés dans différents contextes (imagerie de la lithosphère, exploration de la subsurface, milieux fracturés, évaluation du risque sismique, imagerie de la source sismique) et pour différents types de données (de tremblements de terre et de corrélations de bruit).