Technical University of Denmark

DTU

Improved forward wave propagation and adjoint-based sensitivity kernel calculations using a numerically stable finite-element PML

Xie, Zhinan; Komatitsch, Dimitri; Martin, Roland; Matzen, René

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

Improved forward wave propagation and adjoint-based sensitivity kernel calculations using a numerically stable finite-element PML

Zhinan Xie,[1,2] Dimitri Komatitsch,[1] Roland Martin[3] and René Matzen[4]

[1]*LMA, CNRS UPR* 7051*, Université Aix-Marseille, Centrale Marseille, F*-13402 *Marseille Cedex* 20*, France. E-mail:* *xiezhinan1984@gmail.com*
[2]*Institute of Engineering Mechanics, China Earthquake Administration, Harbin* 150080*, China*
[3]*GET, CNRS UMR* 5563*, Université Toulouse III Paul Sabatier, Observatoire Midi-Pyrénées,* 14 *avenue Édouard Belin, F*-31400 *Toulouse, France*
[4]*Department of Mechanical Engineering, Solid Mechanics, Technical University of Denmark,* DK-2800 *Kgs. Lyngby, Denmark*

SUMMARY

In recent years, the application of time-domain adjoint methods to improve large, complex underground tomographic models at the regional scale has led to new challenges for the numerical simulation of forward or adjoint elastic wave propagation problems. An important challenge is to design an efficient infinite-domain truncation method suitable for accurately truncating an infinite domain governed by the second-order elastic wave equation written in displacement and computed based on a finite-element (FE) method. In this paper, we make several steps towards this goal. First, we make the 2-D convolution formulation of the complex-frequency-shifted unsplit-field perfectly matched layer (CFS-UPML) derived in previous work more flexible by providing a new treatment to analytically remove singular parameters in the formulation. We also extend this new formulation to 3-D. Furthermore, we derive the auxiliary differential equation (ADE) form of CFS-UPML, which allows for extension to higher order time schemes and is easier to implement. Secondly, we rigorously derive the CFS-UPML formulation for time-domain adjoint elastic wave problems, which to our knowledge has never been done before. Thirdly, in the case of classical low-order FE methods, we show numerically that we achieve long-time stability for both forward and adjoint problems both for the convolution and the ADE formulations. In the case of higher order Legendre spectral-element methods, we show that weak numerical instabilities can appear in both formulations, in particular if very small mesh elements are present inside the absorbing layer, but we explain how these instabilities can be delayed as much as needed by using a stretching factor to reach numerical stability in practice for applications. Fourthly, in the case of adjoint problems with perfectly matched absorbing layers we introduce a computationally efficient boundary storage strategy by saving information along the interface between the CFS-UPML and the main domain only, thus avoiding the need to solve a backward wave propagation problem inside the CFS-UPML, which is known to be highly ill-posed. Finally, by providing several examples we show numerically that our formulation is efficient at absorbing acoustic waves for normal to near-grazing incident body waves as well as surface waves.

Key words: Seismic tomography; Computational seismology; Wave propagation.

1 INTRODUCTION

Large-scale complex wave propagation problems in unbounded domains are frequently encountered in earthquake simulations, seismic tomography, geophysical exploration, ocean acoustics or non-destructive acoustic testing for instance. In the application of time-domain adjoint methods and imaging to improve our knowledge of large, complex underground tomographic models, it is important to model heterogeneous media accurately at the local and regional scale. In order to limit the computational cost of such iterative imaging problems or even very high frequency forward problems for a given model solved for instance based on a time-domain finite-element (FE) or finite-difference (FD) method, one aims at reducing the problem size by introducing efficient artificial absorbing boundaries to truncate the semi-infinite or infinite medium. In the case of convex artificial boundaries, one acceptable assumption that simplifies the setup of the truncated problem is that all waves leaving the truncated domain are purely outgoing and should thus never reenter it (Liao & Wong 1984). Along such artificial boundaries, infinite-domain truncating methods should thus be introduced to reduce or even eliminate the spurious

boundary reflections, which is important in particular if thin mesh slices are used and/or if receivers are located at large offset (Martin & Komatitsch 2009).

Simple summation of different runs performed with Dirichlet (rigid) and then Neumann (free) totally reflecting boundary conditions, leading to cancellation of reflections with opposite sign was introduced by Smith (1974) but is not satisfactory because, while it is exact for reflections that occur in the corner of the model under study, it does not work for waves reflected off opposite model edges nor for multiple (high-order) corner reflections. In addition, such an approach is expensive because it requires summing eight runs in 3-D and four runs in 2-D; it has therefore been abandoned nowadays. Four main types of methods have been introduced since then (Givoli 2004), namely, boundary integral methods (BIMs), infinite element methods (IEMs), absorbing layer methods (ALMs) and absorbing boundary condition (ABC) methods, which can be further divided into global ABCs and local ABCs. BIM, IEM, together with global ABCs are only of limited application in the case of complex models due to the fact that: (i) they are valid only for homogeneous unbounded domains; (ii) they are valid only for a particular shape of the artificial boundary and (iii) their computational cost is often high. More details can be found for instance in Ting & Miksis (1986), Chen *et al.* (2004), Teng (2003) and Grote & Kirsch (2007) regarding BIM, Astley & Hamilton (2006) regarding IEM and Givoli & Keller (1990), Grote & Keller (1995), Keller & Grote (2000) and Givoli & Patlashenko (2004) regarding global ABCs. Thus, only local ABCs and ALMs are currently widely used.

The main idea behind the design of a local ABC is to use the outgoing nature of waves that need to be absorbed, which implies that the values on artificial boundaries can be extrapolated based upon past time values. Thus, in its continuous form before discretization by a numerical scheme, all spatial derivatives along the normal direction of an artificial boundary are one-sided. Since the end of the 1970s, different families of local ABCs have been developed, such as viscous boundary conditions (Lysmer & Kuhlemeyer 1969), viscous-spring boundary conditions (Deeks & Randolph 1994; Liu & Li 2005), paraxial conditions (Clayton & Engquist 1977; Stacey 1988), asymptotic operators (Bayliss & Turkel 1980), multitransmitting boundary conditions (Liao *et al.* 1984) and Higdon boundary conditions (Higdon 1986, 1990). The first two are only of first order, while the other four families provide a hierarchy of local ABCs with increasing accuracy order. A long-standing problem in the application of a local ABC is that it is hard to construct stable numerical schemes for high-order local ABC due to the presence of high-degree one-sided normal derivatives. Furthermore, due to the difficulty in establishing the well-posedness of local ABC formulations, instabilities encountered in their implementation are not well understood nor solved. There is thus ongoing research work on that topic, and significant progress has been made in recent years. First, by reformulating paraxial conditions, Higdon boundary conditions or asymptotic operators (Guddati & Tassoulas 2000; Givoli & Neta 2003; Hagstrom & Warburton 2004; Hagstrom *et al.* 2008; Hagstrom & Warburton 2009), the so-called practical high-order local ABCs have been proposed. In these conditions, high-degree normal derivatives are eliminated based on the use of auxiliary variables, defined only on the artificial boundary and assumed to satisfy the interior wave equation. Secondly, by adding a Lysmer–Kuhlemeyer operator inside the Higdon boundary conditions, a long-time stable high-order ABC has been derived and implemented for a 2-D elastic guided-wave problem by Baffet *et al.* (2012). However, for a layered infinite domain or even a homogeneous infinite domain governed by the second-order wave equation written in displacement, long-time stable high-order ABC formulations are still not available (Rabinovich *et al.* 2011, 2013). Other stable implementations of high-order ABCs have been obtained from space–time localization of global ABCs (Alpert *et al.* 2002; Du & Zhao 2010; Grote & Sim 2011), however these ABCs inherit the limitations of global ABCs. This is the reason why simple first-order ABCs are still widely used nowadays, but one needs to impose them sufficiently far from the region of interest in order to get relatively accurate results because their efficiency at absorbing waves is often relatively poor. Unfortunately, this is computationally inefficient.

In parallel, ALM is undergoing rapid development. The main idea of an ALM is to replace the infinite domain outside the artificial boundary with an absorbing layer of finite width. Thus, ideally on the artificial boundary waves can enter from the main domain into the absorbing layer without generating any reflection, and then damp out exponentially. ALM can be divided into two main groups: (i) sponge layers, which include the tapering absorbing layer (Cerjan *et al.* 1985) and physical absorbing layers (Israeli & Orszag 1981; Kosloff & Kosloff 1986; Sochacki *et al.* 1987; Sarma *et al.* 1998; Semblat *et al.* 2011); (ii) Perfectly matched layers (PML), the main advantage of PML over sponge layers being that in the continuous and infinite case waves enter into the absorbing layer without generating any reflection irrespective of their frequency and incidence angle (hence the name 'perfectly matched'). For a PML of finite size, damped-out waves will reflect off the outer edge of the PML, on which a Dirichlet rigid condition, that is, a condition with total reflection is implemented, and will be damped again on their way back to the entrance of the PML; thus theoretically these waves are not zero, but their amplitude is damped by an exponential factor that depends twice the thickness of the PML (since the waves travel through it twice) and thus in practice they are extremely small. Sponge layers are not perfectly matched, which makes their efficiency small in practice because significant spurious reflections occur. The use of smooth attenuation profiles can partially alleviate this difficulty, but doing so requires thick regions and consequently additional storage and computation. PML was originally developed by Bérenger (1994) in electromagnetism. Then, the interpretation and rederivation of PML in terms of complex-coordinate-stretching or complex-field-stretching methods (Chew & Weedon 1994; Teixeira & Chew 1998) led to its rapid development and application in other fields such as acoustics, elastodynamics, poroelastodynamics and hydrodynamics. Here, we will focus on the application of PML in the case of the linear elastic wave equation. As historically PML was first derived for Maxwell's equations written as a first-order system, by analogy a large amount of research work on PML has been developed for the first-order velocity–stress form of the elastic wave equation. Thus, its implementation is often based on velocity–stress FD methods or mixed velocity–stress FE methods (Chew & Liu 1996; Bécache *et al.* 2001; Collino & Tsogka 2001; Festa & Nielsen 2003; Marcinkovich & Olsen 2003; Cohen & Fauqueux 2005; Ma & Liu 2006; Drossaert & Giannopoulos 2007a; Komatitsch & Martin 2007). In this paper, we will work with the second-order

displacement-based form of the elastic wave equation because, among other advantages (Kreiss *et al.* 2002) it is a natural and very efficient framework for the spectral-element method (SEM) for both forward and adjoint simulations (Komatitsch & Tromp 1999; Tromp *et al.* 2008; Peter *et al.* 2011), and in addition less computational work is involved. To our knowledge, the first work on this system for PML is Komatitsch & Tromp (2003), in which based on the complex-coordinate-stretching approach the authors derived a classical split-field PML (SPML) formulation by splitting the displacement into four components. Here and below we refer to a 'classical PML' as a PML derived by using the classical coordinate stretching function first introduced by Chew & Weedon (1994). However, it turns out that the formulation of Komatitsch & Tromp (2003) has long-time instability problems, which were not known at the time. Later, by interpreting the Newmark time scheme as a time-staggered velocity–stress algorithm, Festa & Vilotte (2005) successfully implemented a velocity–stress complex-frequency-shifted SPML (CFS-SPML) formulation based on a SEM, while still using a classical SEM based on the second-order wave equation written in displacement inside the main domain. However, it is not clear how to extend their work to other, more accurate, time schemes, and Festa *et al.* (2005) mention instabilities for that formulation. Similarly, here and below we refer to a 'CFS-PML' as a PML derived by using the CFS coordinate stretching function first introduced by Kuzuoglu & Mittra (1996) and extended by Roden & Gedney (2000), that is, a PML that has a more sophisticated coordinate transform that can act as a filter to improve the behaviour at grazing incidence after discretization by a numerical scheme. The first classical unsplit-field PML (UPML), that is, a scheme that does not require to split the field into several (artificial) components, was directly derived based on the second-order wave equation written in displacement by Basu & Chopra (2004). It can be directly implemented in classical displacement-based FE methods (Bao *et al.* 1998). That formulation was extended and implemented in Li & Matar (2010) and Kucukcoban & Kallivokas (2011, 2013). However, long-time instability problems are reported in their implementations.

Instability problems are not specific to PML derived based on the second-order wave equation written in displacement; similar problems exist in the velocity–stress PML. Long-time instability is also observed in the case of an isotropic medium, for instance Festa & Nielsen (2003) and Marcinkovich & Olsen (2003) report long-time instability in the classical velocity–stress SPML implemented based on a staggered FD scheme. The mathematical analysis of Joly (2012) and Kreiss & Duru (2013) in the case of an isotropic medium has shown that the long-time instability appears in the classical PML due to the use of the classical coordinate stretching functions, which makes the resulting PML formulation weakly well-posed, meaning that a growth of total energy, that is, an instability, exists but remains bounded by a constant number that is independent of the spatial discretization step and time step for a fixed simulation time. This implies that inside the PML solutions can potentially grow with time. These instabilities can then spread into the main domain. In addition, and independently, all the classical PML models such as SPML, UPML, CFS unsplit-field PML (CFS-UPML) and CFS-SPML have also been shown to be ill-posed and unstable for some types of (very) anisotropic media (Bécache *et al.* 2003; Komatitsch & Martin 2007; Martin *et al.* 2008; Meza-Fajardo & Papageorgiou 2008; Kreiss & Duru 2013). To try to overcome this difficulty, a condition called Multi-axial PML (M-PML) was proposed in Meza-Fajardo & Papageorgiou (2008) but it was soon proven that it is not perfectly matched and thus is not a PML (Dmitriev & Lisitsa 2011), it is only an (improved) sponge.

However, in practice, a long-time numerically stable implementation of the velocity–stress CFS-UPML based on a second-order accurate velocity–stress FD technique has been described in Komatitsch & Martin (2007), and in the 2-D case a long-time numerically stable implementation of CFS-UPML based on the second-order wave equation written in displacement and implemented based on a classical low-order FE method has been described in Matzen (2011). By 'numerically stable' here and in the whole paper, we mean that all the numerical examples given seem very stable, but no mathematical stability proof is available (at least so far). However, careful design and selection of the complex-shifted coordinate stretching function and its parameters are needed to avoid the potentially singular parameters that exist in that formulation. Another issue regarding work found in the literature is that the CFS-UPML displacement-based formulation for FE techniques has never been developed in 3-D, and also not developed for adjoint wave propagation problems (not even in 2-D).

Due to the difficulties encountered to design a numerically stable CFS-UPML displacement-based formulation, in particular one suitable for adjoint problems and for FE implementation, in practice in many cases time-domain adjoint-based tomography in regional or local models is still performed using simple and far less accurate viscous boundary condition or tapered ALMs (see, e.g. Fichtner *et al.* 2009b; Douma *et al.* 2010; Tape *et al.* 2010; Zhu *et al.* 2012; Zhu & Tromp 2013; Luo *et al.* 2014). This is not satisfactory because it makes difficult to use an automatic waveform misfit picking algorithm to pick the different seismic or acoustic phases of the recorded waveforms, since they are contaminated by the spurious waves reflected off the artificial boundaries of the model. Thus, one can then only use data recorded by the receivers in a short-time range, before the spurious waves reach them, or use receivers located far from the artificial boundaries. This means that for inverse problems with such poorly accurate absorbing boundaries the computational cost increases very significantly because one needs to use a substantially larger domain to try to diminish the effect of spurious reflections. Keeping in mind that hundreds of forward and adjoint simulations are needed in order to solve an inverse problem iteratively in practice, this quickly becomes very problematic.

Thus, in this paper, in order to meet the challenges encountered when applying adjoint methods in the time domain to imaging problems in a region with absorbing layers, we first improve and extend the 2-D convolution formulation of CFS-UPML introduced by Matzen (2011), and then derive the 3-D case. We introduce both a convolution version and an auxiliary differential equation (ADE) version (Gedney & Zhao 2010; Martin *et al.* 2010) of this improved formulation. Both formulations can be implemented in displacement-based numerical methods, but the ADE form allows for extension to higher order time schemes and is also easier to implement. Secondly, we rigorously derive the CFS-UPML formulation for adjoint wave propagation problems. Thirdly, we define a computationally efficient boundary storage strategy by saving all the information needed along the interface between the CFS-UPML and the main domain only. This enables us to perform on-the-fly calculations of sensitivity kernels for imaging problems. Fourthly, we implement the proposed CFS-UPML using both a classical

low-order mass-lumped FE method and a high-order Legendre SEM. In the case of the convolution formulation, we use a Newmark scheme combined with a second-order recursive convolution scheme for time integration, and in the case of the ADE formulation we resort to a two-level low-dispersion and low-dissipation Runge–Kutta (LDDRK) scheme.

2 FORWARD WAVE PROPAGATION PROBLEM

In this section, we will make the 2-D convolution formulation of the CFS-UPML derived in previous work more flexible by providing a new treatment to analytically remove singular parameters in the formulation, which will then allow for a far more flexible choice regarding the complex-shifted coordinate stretching function. We will then also extend this new formulation to 3-D. Furthermore, we will derive the corresponding ADE formulation of CFS-UPML, which will enable us to use the same time and space numerical scheme in the CFS-UPML absorbing layer as in the main domain. We will show that both formulations can be implemented in second-order displacement-based formulations, while work found in the literature often focuses on the first-order velocity–stress formulation only.

2.1 Classical wave equation

In a heterogeneous solid region of the Earth, the second-order elastic wave equation expressed in terms of the displacement vector can be written in strong form as:

$$\rho \ddot{\mathbf{u}} = \nabla \cdot \boldsymbol{\sigma} + \mathbf{f}, \tag{1a}$$

$$\boldsymbol{\sigma} = \mathbf{c} : \nabla \mathbf{u}, \tag{1b}$$

where $\mathbf{u}$ is the displacement vector, $\boldsymbol{\sigma}$ is the symmetric second-order stress tensor, $\mathbf{c}$ is the fourth-order elastic constitutive tensor, ρ is mass density and $\mathbf{f}$ is an external force representing the seismic source. '∇' is the gradient operator, the '$\cdot$' symbol represents the dot product, the scalar product of the gradient operator with a tensor field represents its divergence, the '$:$' symbol represents a double tensorial contraction operation and a dot over a symbol represents its derivative with respect to time. The material properties of the solid, $\mathbf{c}$ and ρ, can be spatially heterogeneous and are assumed to be known time-independent quantities that define the geological medium under study. $\mathbf{c}$ is symmetric, with minor and major symmetries, and is positive definite. In the isotropic case, $\mathbf{c}$ can be written in terms of the Kronecker delta δ_{ij} as:

$$c_{ijkl} = \left(\kappa - \frac{2}{3}\mu\right)\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) = \lambda\,\delta_{ij}\delta_{kl} + \mu\,(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}); \qquad i, j, k, l = 1, 2, 3, \tag{2}$$

where c_{ijkl} are the components of $\mathbf{c}$, κ is the bulk modulus, λ is Lamé's first parameter and μ is Lamé's second parameter, also called the shear modulus. Throughout the paper, the Einstein summation convention over repeated indices is implicitly assumed except when stated otherwise.

2.2 CFS-UPML

The derivation of PML can be formalized using the concept of complex coordinate stretching: PML and the related artificially damped wave equations can be interpreted as an analytically continuous extension in the complex space of the classical elastic wave equation defined in real space (Chew & Weedon 1994). In this section, we improve and extend the 2-D convolution CFS-UPML formulation derived in Matzen (2011) by analytically removing singularities that can potentially arise in some of the parameters that appear in the time-domain form of CFS-UPML, which then facilitates the flexible choice of CFS coordinate stretching functions. Here, we will only show the main steps of the derivation in 3-D, details being given in Appendix A, and for completeness we give the 2-D version in Appendix B. In the frequency domain, as classically done, let us first define a complex coordinate system by stretching the real coordinate based on a general complex-coordinate-stretching function, then map the Fourier-transformed eq. (1) into that new complex coordinate system to transform the original elastic wave equation into an artificially damped wave equation. In order to avoid having to implement the PML in complex coordinates, we then rewrite the resulting wave equation in real coordinates based on the chain rule. By using a complex-coordinate-stretching function of the CFS type (Kuzuoglu & Mittra 1996; Roden & Gedney 2000; Drossaert & Giannopoulos 2007b; Komatitsch & Martin 2007; Martin *et al.* 2008, 2010), we get the frequency-domain CFS-UPML. An inverse Fourier transform of these reformulated equations then gives a new time-domain CFS-UPML that can be directly implemented in displacement-based numerical methods. Below, we introduce both the convolution and the ADE formulation of that new time-domain CFS-UPML.

2.2.1 Frequency-domain equation

In Cartesian coordinate form, the Fourier-transformed eq. (1) can be written as

$$-\rho\omega^2\hat{\mathbf{u}} = \nabla \cdot \hat{\boldsymbol{\sigma}}, \tag{3a}$$

$$\hat{\boldsymbol{\sigma}} = \mathbf{c} : \nabla\hat{\mathbf{u}}, \tag{3b}$$
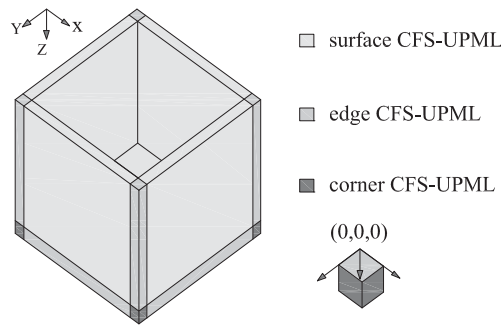
Figure 1. Semi-infinite domain truncated by a CFS-UPML absorbing layer. The top surface is a free (and thus totally reflecting) surface. We denote the different CFS-UPML surfaces as CFS-PML(s_x, 1, 1), CFS-PML(1, s_y, 1), CFS-PML(1, 1, s_z); the edge CFS-UPML as CFS-PML(s_x, s_y, 1), CFS-PML(s_x, 1, s_z), CFS-PML(1, s_y, s_z) and the corner CFS-UPML as CFS-PML(s_x, s_y, s_z).

where '^' over a symbol represents its Fourier transform, $\mathbf{V}_i = \partial_{x_i}$, $i = 1, 2, 3$ and $\partial_{x_1} = \partial/\partial_x$, $\partial_{x_2} = \partial/\partial_y$, $\partial_{x_3} = \partial/\partial_z$. In the CFS-UPMLs as well as in the whole domain, the initial displacement, velocity and source term are assumed to be zero, that is, the medium is initially at rest. Without losing generality, we show the derivation of the corner CFS-UPML in the first octant, that is, in the region where $x_i \geq 0$ (Fig. 1). We first introduce the new complex coordinate as

$$\tilde{x}_i(x_i) = \int_0^{x_i} s_i(x_i') \, dx_i', \tag{4}$$

where the s_i are non-zero complex-coordinate-stretching functions. After directly mapping (3) into the newly introduced complex coordinates, we rewrite the resulting equations in real coordinates based on the chain rule $\partial_{\tilde{x}_i} = 1/s_i \, \partial_{x_i}$ and get:

$$-\rho\omega^2\hat{\mathbf{u}} = \tilde{\mathbf{V}} \cdot \tilde{\boldsymbol{\sigma}}, \tag{5a}$$

$$\tilde{\boldsymbol{\sigma}} = \mathbf{c} : \tilde{\mathbf{V}}\hat{\mathbf{u}}, \tag{5b}$$

where $\tilde{\mathbf{V}}_i = (1/s_i) \, \partial_{x_i}$ (no summation). In order for the layer to be a PML, the above equation should have the following properties: (i) along the interface between the PML and the main domain, the solutions of (3) and (5) should be equal, that is, the reflection coefficient on the interface should be zero; (ii) wave solutions of (5) should be exponentially damped inside the PML; (iii) no growing solution should be supported by the time-domain counterpart of (5), that is, total energy should decrease in the absorbing layer and no numerical instability should appear (i.e. no growing energy should be observed). As shown by Teixeira & Chew (1997), the first property is naturally satisfied by any general choice of s_i, since (5) has exactly the same form as (3). Thus, waves coming from the main domain enter the CFS-UPML without generating any spurious reflection. The two other properties must be enforced by a proper choice of the s_i. Theoretically, many different choices could be made. However, in practice complicated s_i will make (5) difficult or even impossible to be Fourier-transformed back into a time domain in a closed (analytical) form, and thus only two main types of coordinate-stretching functions are widely used:

classical (original) type: $s_i(x_i) = 1 + \dfrac{d_i(x_i)}{\mathbf{i}\,\omega}$, $\qquad$ (6a)

CFS type: $s_i(x_i) = \kappa_i(x_i) + \dfrac{d_i(x_i)}{\alpha_i(x_i) + \mathbf{i}\,\omega}$, $\qquad$ (6b)

where **i** is the complex number and the d_i are damping factors along the normal direction x_i (Fig. 2) that cause the amplitude of a propagating wavefield with any incidence angle to decay exponentially along that normal direction inside the CFS-UPML. The α_i are frequency-shifting factors that make the damping rates of waves inside the CFS-UPML depend on frequency, thus providing a Butterworth-type filter, and the κ_i are scaling factors that can be used to improve the attenuation of evanescent waves (Zhang & Shen 2010). It has been found (Festa & Vilotte 2005) that a strictly positive α_i can improve the damping rate of near-grazing incident propagating waves, while decreasing the damping rate of low-frequency propagating waves. The scaling factors κ_i tend to bend grazing incidence propagating waves into waves that travel closer to
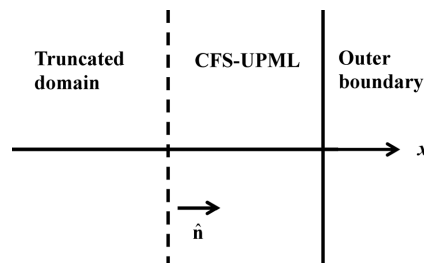

Figure 2. Definition of the local normal **n̂** to the interface between the main domain and the CFS-UPML.

normal incidence, and thus it can also improve the absorption of waves at grazing incidence. However, increasing the value of κ_i will slightly reduce the damping rate of near-normal incidence waves (Zhang & Shen 2010).

The CFS function was introduced in Kuzuoglu & Mittra (1996) with the main idea of restoring the causality of PML, since the analysis performed by these authors pointed out that the original PML model derived using classical damping functions is not causal and can thus become unstable in time-domain simulations. Although their proof has later been shown to be wrong (Teixeira & Chew 1999), later research has shown that PML models derived based upon classical coordinate-stretching functions are only weakly well-posed (Joly 2012) and that growing (i.e. unstable) solutions can potentially exist and be triggered for instance by roundoff numerical errors. However, a long-time numerically stable implementation of the velocity–stress CFS-UPML based on a second-order accurate 3-D velocity–stress FD technique has been achieved in Komatitsch & Martin (2007), and in the 2-D case a long-time numerically stable implementation of CFS-UPML based on the second-order wave equation written in displacement and implemented based on a classical low-order FE method has been developed in Matzen (2011). In this paper, we therefore resort to a CFS function; more specifically, we define the profile of α_i as

$$\alpha_i(x_i) = \alpha_i^{\max} \left[1 - (x_i/L)^p \right], \tag{7}$$

where L is the thickness of the PML, the $\alpha_i^{\max}$ are positive real constants and p is a positive real number. Such a damping profile makes the CFS-UPML efficient at absorbing near-grazing incident waves and also efficient at absorbing low-frequency waves (Gedney 1998). Fixed (i.e. Dirichlet) boundary conditions are imposed on the outer edges of the PML.

2.2.2 Time-domain equation

In order to obtain a formulation in the time domain based on the displacement vector only we need to reformulate (5) by multiplying by $s_1 s_2 s_3$ on both sides of the equation of motion (Jiao *et al.* 2003; Basu & Chopra 2004; Matzen 2011):

$$- \rho \omega^2 s_1 s_2 s_3 \hat{\mathbf{u}} = s_1 s_2 s_3 \tilde{\nabla} \cdot \tilde{\boldsymbol{\sigma}}, \tag{8a}$$

$$\tilde{\boldsymbol{\sigma}} = \mathbf{c} : \tilde{\nabla} \hat{\mathbf{u}}. \tag{8b}$$

The inverse Fourier transform of (8) leads to the time-domain convolution formulation of CFS-UPML:

$$\rho \, L(t) * \mathbf{u} = \nabla \cdot \bar{\boldsymbol{\sigma}}, \tag{9a}$$

$$\bar{\sigma}_{ij} = \bar{c}_{ijkl} * \partial_k u_l, \tag{9b}$$

where

$$L(t) = F^{-1} \left[-\omega^2 s_1 s_2 s_3 \right], \quad \bar{c}_{ijkl} = c_{ijkl} F^{-1} \left[\frac{s_1 s_2 s_3}{s_i s_k} \right]. \tag{10}$$

Based on the original expression of the CFS coordinate stretching function (6b), the full derivation of $L(t)$ and $\bar{c}_{ijkl}$ is difficult and can fail when repeated poles are present in the terms $-\omega^2 s_1 s_2 s_3$ or $s_1 s_2 s_3 / s_i s_k$. In such a case, with that original formulation it is not possible to cleanly remove all the singular parameters from $L(t)$ and $\bar{c}_{ijkl}$. To be able to do that, it is convenient to rewrite the stretching functions in terms of $\beta_i = \alpha_i + d_i / \kappa_i$, such that:

$$s_i(x_i) = \kappa_i(x_i) \frac{\beta_i(x_i) + \mathbf{i}\,\omega}{\alpha_i(x_i) + \mathbf{i}\,\omega} \tag{11}$$

for $i = 1, 2, 3$. With the aid of formal calculation software such as Mathematica or Maple, the exact expression of (9) can then be derived. For the $L(t)$ term, for instance, in the simplest case $\alpha_1 \neq \alpha_2 \neq \alpha_3$, that is, when poles of $-\omega^2 s_1 s_2 s_3$ are all unique we have:

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_1 t} H(t) + \bar{a}_4 e^{-\alpha_2 t} H(t) + \bar{a}_5 e^{-\alpha_3 t} H(t), \tag{12a}$$

$$\bar{a}_0 = \kappa_1 \kappa_2 \kappa_3, \quad \bar{a}_1 = a_0 \left(\Gamma_{\alpha_1}^{\beta_1} + \Gamma_{\alpha_2}^{\beta_2} + \Gamma_{\alpha_3}^{\beta_3} \right), \quad \bar{a}_2 = a_0 \left(\Gamma_{\alpha_1}^{\beta_1} \Gamma_{\alpha_1, \alpha_2}^{\beta_2} + \Gamma_{\alpha_2}^{\beta_2} \Gamma_{\alpha_2, \alpha_3}^{\beta_3} + \Gamma_{\alpha_3}^{\beta_3} \Gamma_{\alpha_1, \alpha_3}^{\beta_1} \right), \tag{12b}$$

$$\bar{a}_3 = a_{123}, \quad \bar{a}_4 = a_{213}, \quad \bar{a}_5 = a_{312}, \quad a_{ijk} = \kappa_i \kappa_j \kappa_k \alpha_i^2 \Gamma_{\alpha_i}^{\beta_i} \Gamma_{\alpha_i}^{\beta_j} \Gamma_{\alpha_i}^{\beta_k} / \left(\Gamma_{\alpha_i}^{\alpha_j} \Gamma_{\alpha_i}^{\alpha_k} \right), \tag{12c}$$

where $\Gamma_b^a = a - b$, $\Gamma_{b,c}^a = a - b - c$. Furthermore, in that case all convolution terms such as $[e^{-\alpha_i t} H(t)] * u_i$ that arise when convolving $L(t)$ with the different components of the displacement vector can be computed efficiently, without having to store the whole past time steps of the simulation on the computer, based on an incremental recursive convolution technique (Luebbers & Hunsberger 1992) owing to the exponential algebraic property of the convolution kernel function $e^{-\alpha_i t} H(t)$:

$$e^{-\alpha_i (t \pm t_0)} H(t \pm t_0) = e^{-\alpha_i t} e^{\pm \alpha_i t_0}, \qquad t \pm t_0 \geq 0, \tag{13}$$

where t_0 is a real constant.

However, there are also more complicated cases in which some poles appear more than once. For the $-\omega^2 s_1 s_2 s_3$ term, for instance, some poles appear twice in the following cases: (1) $\alpha_1 = \alpha_2$; (2) $\alpha_1 = \alpha_3$; (3) $\alpha_2 = \alpha_3$ and poles can even appear three times when (4) $\alpha_1 = \alpha_2 = \alpha_3$.

The repeated poles then give rise to convolution kernel functions of the form $e^{-\alpha_i t} t H(t)$ or $e^{-\alpha_i t} t^2 H(t)$. For example, when $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_0$ we have

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_0 t} H(t) + \bar{a}_4 e^{-\alpha_0 t} t H(t) + \bar{a}_5 e^{-\alpha_0 t} t^2 H(t), \tag{14a}$$

$$\bar{a}_0 = \kappa_1 \kappa_2 \kappa_3, \quad \bar{a}_1 = \bar{a}_0 \left(\Gamma_{\alpha_0}^{\beta_1} + \Gamma_{\alpha_0}^{\beta_2} + \Gamma_{\alpha_0}^{\beta_3} \right), \quad \bar{a}_2 = \bar{a}_0 \left(\Gamma_{\alpha_0}^{\beta_1} \Gamma_{\alpha_0,\alpha_0}^{\beta_2} + \Gamma_{\alpha_0}^{\beta_2} \Gamma_{\alpha_0,\alpha_0}^{\beta_3} \right), \tag{14b}$$

$$\bar{a}_5 = \frac{1}{2} \kappa_1 \kappa_2 \kappa_3 \alpha_0^2 \Gamma_{\alpha_0}^{\beta_1} \Gamma_{\alpha_0}^{\beta_2} \Gamma_{\alpha_0}^{\beta_3}, \quad \bar{a}_4 = -2 \frac{\partial a_5}{\partial \alpha_0}, \quad \bar{a}_3 = -\frac{1}{2} \frac{\partial a_4}{\partial \alpha_0}. \tag{14c}$$

Thus, if we now convolve $L(t)$ with components of the displacement vector, convolution terms of the following form will arise:

$$\left[e^{-\alpha_0 t} t H(t) \right] * u_i, \quad \left[e^{-\alpha_0 t} t^2 H(t) \right] * u_i. \tag{15}$$

Substituting the above convolution kernel function into (13), we can verify that it does not satisfy the exponential algebraic property, thus (15) cannot be computed directly based on a recursive convolution technique. However, this difficulty can be circumvented by using the linear property of the convolution:

$$\left[e^{-\alpha_0 t} t H(t) \right] * u_i = t \left[e^{-\alpha_0 t} H(t) \right] * u_i + \left[e^{-\alpha_0 t} H(t) \right] * [t u_i] \tag{16}$$

and

$$\left[e^{-\alpha_0 t} t^2 H(t) \right] * u_i = t^2 \left[e^{-\alpha_0 t} H(t) \right] * u_i - 2t \left[e^{-\alpha_0 t} H(t) \right] * [t u_i] + \left[e^{-\alpha_0 t} H(t) \right] * [t^2 u_i]. \tag{17}$$

Based on (16) and (17), convolving (14) with the displacement field then results in

$$L(t) * u_i(t) = a_0 \ddot{u}_i + a_1 \dot{u}_i + a_2 u_i + a_3 \left[e^{-\alpha_0 t} H(t) \right] * u_i + a_4 \left[e^{-\alpha_0 t} H(t) \right] * [t u_i] + a_5 \left[e^{-\alpha_0 t} H(t) \right] * [t^2 u_i], \tag{18a}$$

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3 + t \bar{a}_4 + t^2 \bar{a}_5, \quad a_4 = -\bar{a}_4 - 2t \bar{a}_5, \quad a_5 = \bar{a}_5, \tag{18b}$$

which then makes all convolution terms suitable for efficient computation based on a recursive convolution technique. For completeness, in Appendix A we give the full expression of that CFS-UPML convolution formulation.

Based on this convolution formulation, we can easily obtain the ADE formulation by taking the derivative of the equation that governs the convolution terms. We will show in Section 4 that such a formulation can naturally be extended to higher order time schemes, which is not the case of the convolution formulation. As an example, taking the convolution terms in (18) and denoting $[e^{-\alpha_0 t} H(t)] * u_i$, $[e^{-\alpha_0 t} H(t)] * [t u_i]$ and $[e^{-\alpha_0 t} H(t)] * [t^2 u_i]$ as R_1, R_2, R_3, we have

$$\dot{R}_1 = -\alpha_0 R_1 + u_i, \tag{19a}$$

$$\dot{R}_2 = -\alpha_0 R_2 + t u_i, \tag{19b}$$

$$\dot{R}_3 = -\alpha_0 R_3 + t^2 u_i. \tag{19c}$$

In a similar way, we can derive the ADE for the convolution terms that arise in $L(t) * \mathbf{u}$ and $\bar{c}_{ijkl} * \partial_k u_l$, and after doing so all these ADE share the same form:

$$\dot{R}_m = -b R_m + g(t), \tag{20}$$

where m is an integer. More details about the ADE formulation can be found in Appendix B for the 2-D case and in Appendix A for the 3-D case.

3 ADJOINT WAVE PROPAGATION PROBLEM

Adjoint-based tomography is an efficient way of using differences between observed and simulated data to iteratively improve initial models, based on sensitivity kernels expressed in terms of Fréchet derivatives for a chosen objective misfit function (Chavent *et al.* 1975; Tarantola & Valette 1982; Lailly 1983; Tarantola 1984, 1987, 1988; Akçelik *et al.* 2003; Tromp *et al.* 2005, 2008; Fichtner *et al.* 2009a; Fichtner 2010; Peter *et al.* 2011). It is also part of more complex full waveform imaging (Bamberger *et al.* 1982; Tarantola & Valette 1982; Lailly 1983; Tarantola 1984, 1987, 1988; Virieux & Operto 2009; Fichtner 2010; Monteiller *et al.* 2013). In what follows, by 'sensitivity kernels' we mean their classical definition as the volume density of the Fréchet derivative of the misfit functions with respect to model parameter defined for the whole model [(see, e.g. Greiner & Reinhardt 1996, example 3 of their 'Section 2.3: Functional derivatives') and Fichtner (2010, p. 112)]. From a computational point of view, one of the main advantages of adjoint-based tomography is that only two numerical simulations are needed for the computation of sensitivity kernels because of the introduction of the adjoint problem, for which the simulation is independent of the number of receivers and of seismic observables. As a consequence of formulating the adjoint problem as shifted and reversed in time, these sensitivity kernels can be computed by convolution of the simulated forward and adjoint wavefields (Tarantola 1984, 1987, 1988; Tromp

et al. 2005, 2008; Fichtner 2010; Peter *et al.* 2011). Alternatively, one could ignore the reverse and shifted time operation to obtain a different adjoint field that, combined with the forward field by cross-correlation, will yield the sensitivity kernels (Plessix 2006). However, as we will recall below, in order to obtain a self-adjoint problem we have to use the first approach.

Once computed, the sensitivity kernels can then be used either in a linear and local gradient-based optimization technique or in a global and non-linear (stochastic) technique to minimize differences between data and synthetics by improving the initial model iteratively. Popular examples of such local gradient-based optimization techniques are conjugate gradients, Least-Squares problems by means of bidiagonalization and QR decomposition (LSQR) and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

For a given misfit function, adjoint wave problems can be set up based upon the Lagrange multiplier method introduced in Arora & Haug (1979) and used for the time-domain wave equation in Le Dimet & Talagrand (1986) and Liu & Tromp (2006, 2008). The full description of adjoint wave problems should contain: (1) adjoint sources; (2) adjoint wave equations and (3) infinite or semi-infinite domain truncation methods in the case of a model with artificial absorbing boundaries. Derivations of infinite or semi-infinite domain truncation methods for the adjoint wave equation with ABCs (including PML) are uncommon in the literature. To our knowledge, the first article about this topic is Chung *et al.* (2000), then Rickard *et al.* (2003) and Liu & Tromp (2006). The former two are for Maxwell's equations written in the strong form, and use the split field approach to formulate the problem; the last one only considers viscous boundary conditions, which are not as efficient as PML. Thus, after defining the misfit function, in the following section to our knowledge we will derive the CFS-UPML adjoint wave equation written at the second order in displacement for the first time. We will show that the problem remains self adjoint in the presence of CFS-UPML.

3.1 Misfit function

In practice, the available amount of real data is limited and may not be sufficient to fully constrain the tomography problem. In order to extract as much information as possible from real data, we should define the misfit function carefully. A lot of research has thus been done on that aspect of adjoint methods, see, for instance Tromp *et al.* (2008); Fichtner (2010); Bozdağ *et al.* (2011) and references therein. Since strong-motion seismometers, developed to record large-amplitude vibrations, can record velocity and acceleration in addition to displacement, we can define a relatively general misfit functional G as a function of residual signal displacement (state variables) $\mathbf{u}_r$, velocity $\dot{\mathbf{u}}_r$ and acceleration $\ddot{\mathbf{u}}_r$ and a field filtered based on the convolution operator $\hat{\mathbf{u}}_r$ between the real and the synthetic data. We thus have:

$$\Phi = \int_0^T \int_{\Omega_t} G(\mathbf{u}_r, \dot{\mathbf{u}}_r, \ddot{\mathbf{u}}_r, \hat{\mathbf{u}}_r) \mathrm{d}^3\mathbf{x}\mathrm{d}t, \tag{21}$$

where Φ is the misfit function, Ω_t denotes the truncated domain, T is the total duration (i.e. the final time) of the time-domain simulation and the convolution filter operator is expressed as $\hat{\mathbf{u}}_r = \int_{-\infty}^{\infty} \mathbf{u}_r(t) F(t - t') \mathrm{d}t' = F * \mathbf{u}_r(t)$, where $F(t)$ denotes the filter operation. Commonly used filters to handle time-domain signals comprise Fourier, Hilbert and Laplace transforms, among others. In practice, both the data and the synthetics will be windowed and weighted over the time interval. In what follows, we assume that such operations have been performed.

3.2 Adjoint wave equation, adjoint sources and adjoint CFS-UPML formulation

Following the Lagrange multiplier method (Liu & Tromp 2006, 2008), we consider a heterogeneous solid region model of the Earth in which a PML is introduced in order to absorb outgoing waves. Let us denote by Ω_u the volume of the main domain together with the CFS-UPML. Inside that domain the propagation of the synthetic wavefield $\mathbf{u}^s$ is governed by

$$\rho L(t) * \mathbf{u}^s = \nabla \cdot \bar{\boldsymbol{\sigma}}^s + \mathbf{f},$$

$$\bar{\sigma}_{ij}^s = \bar{c}_{ijkl} * \partial_k u_l^s, \tag{22}$$

where $\bar{c}_{ijkl}$ denotes the fourth-order elastic tensor stretched based on PML, whose detailed expression is given by eqs (A20a), (A21a), (A22a), (A23a) and (A24a) in Appendix A. In the non-singular case, the operator $L(t)$ is expressed by

$$L(t) = a_0 \ddot{\delta}(t) + a_1 \dot{\delta}(t) + a_2 \delta(t) + L^c(t) \tag{23}$$

with

$$L^c(t) = \left[a_3 \mathrm{e}^{-\alpha_{x_1} t} + a_4 \mathrm{e}^{-\alpha_{x_2} t} + a_5 \mathrm{e}^{-\alpha_{x_3} t} \right] H(t). \tag{24}$$

Its expression in the singular case follows from eqs (A7a), (A8a), (A9a) and (A10a). On the free outer boundary of Ω_u, which includes the free boundary inside the truncated domain and its extension inside the CFS-UPML, the traction vector must vanish, which leads to the boundary condition

$$\hat{\mathbf{n}} \cdot \bar{\boldsymbol{\sigma}}^s = \mathbf{0} \qquad \text{on } \partial\Omega_u = \Gamma_u^F, \tag{25}$$

where $\hat{\mathbf{n}}$ denotes the unit outward normal to the surface. On the remaining outer boundaries of Ω_u, we implement a Dirichlet (fixed) boundary condition

$$\mathbf{u}^s = \mathbf{0} \qquad \text{on } \partial\Omega_u = \Gamma_u^D. \tag{26}$$

In addition, inside the CFS-UPML, one has initial conditions for (22) expressing the fact that the medium is at rest at the initial time

$$\mathbf{u}^s(\mathbf{x}, 0) = \mathbf{0}, \quad \dot{\mathbf{u}}^s(\mathbf{x}, 0) = \mathbf{0}. \tag{27}$$

Since our goal is to minimize the misfit function (21) given the constraint that the displacement be governed by the elastic wave equation (22), we formulate this optimization problem in terms of the augmented misfit response:

$$\Phi_A = \int_0^T \int_{\Omega_t} G(\mathbf{u}_r, \dot{\mathbf{u}}_r, \ddot{\mathbf{u}}_r, \hat{\mathbf{u}}_r) d^3\mathbf{x} dt - \int_0^T \int_{\Omega_u} \boldsymbol{\lambda} \cdot [\rho L(t) * \mathbf{u}^s - \nabla \cdot \bar{\boldsymbol{\sigma}}^s - \mathbf{f}] d^3\mathbf{x} dt, \tag{28}$$

where $\boldsymbol{\lambda} \equiv \boldsymbol{\lambda}(\mathbf{x}, t)$ denotes the adjoint (Lagrange multiplier) vector. The augmented misfit response is always equal to the original misfit response in eq. (21) because $\rho L(t) * \mathbf{u}^s - \nabla \cdot \hat{\boldsymbol{\sigma}}^s - \mathbf{f} = \mathbf{0}$ is ensured from solving the wave equation (22). This means that their variation must be identical, that is, $\delta\Phi = \delta\Phi_A$. A clever choice of $\boldsymbol{\lambda}$, which can be chosen freely since we solve for $\rho L(t) * \mathbf{u}^s - \nabla \cdot \hat{\boldsymbol{\sigma}}^s - \mathbf{f} = \mathbf{0}$, can simplify the sensitivity analysis considerably as demonstrated in the following.

Using the chain rule, the variations of eq. (28) are given by

$$\delta\Phi_A = \int_0^T \int_{\Omega_u} \left[\frac{\partial G}{\partial u_i^s} \delta u_i^s + \frac{\partial G}{\partial \dot{u}_i^s} \delta \dot{u}_i^s + \frac{\partial G}{\partial \ddot{u}_i^s} \delta \ddot{u}_i^s + \frac{\partial G}{\partial \hat{u}_i^s} \delta \hat{u}_i^s \right] d^3\mathbf{x} dt - \int_0^T \int_{\Omega_u} \lambda_i \left[\delta\rho L(t) * u_i^s - \partial_j \left(\delta\bar{c}_{ijkl} * \partial_k u_l^s \right) \right] d^3\mathbf{x} dt$$

$$- \int_0^T \int_{\Omega_u} \lambda_i \left[\rho L(t) * \delta u_i^s - \partial_j \left(\bar{c}_{ijkl} * \partial_k \delta u_l^s \right) - \delta f_i \right] d^3\mathbf{x} dt, \tag{29}$$

where the convolution operator is defined by

$$\bar{c}_{ijkl}(t) * \partial_k u_l^s(t) = \int_0^t \bar{c}_{ijkl}(t - t') \partial_k u_l^s(t') dt'. \tag{30}$$

Due to the fact that the observed data $\mathbf{u}^o$, $\dot{\mathbf{u}}^o$, $\ddot{\mathbf{u}}^o$ and $\hat{\mathbf{u}}^o$ do not vary with the choice of the misfit function, we have

$$\frac{\partial G}{\partial u_i^o} = \frac{\partial G}{\partial \dot{u}_i^o} = \frac{\partial G}{\partial \ddot{u}_i^o} = \frac{\partial G}{\partial \hat{u}_i^o} = 0. \tag{31}$$

Using integration by parts for both spatial and temporal variations, invoking Gauss' theorem and swapping the order of integration over t and t' as shown in Appendix C, after reordering the terms we get

$$\delta\Phi_A = \int_0^T \int_{\Gamma_u^F} \left[\hat{n}_j \left(\delta\bar{c}_{ijkl} * \partial_k u_l^s \right) + \hat{n}_j \left(\bar{c}_{ijkl} * \partial_k \delta u_l^s \right) \right] \lambda_i d^2\mathbf{x} dt - \int_0^T \int_{\Gamma_u^F} \hat{n}_j (\bar{c}_{ijkl} * \partial_k \lambda_l) \delta u_i^s d^2\mathbf{x} dt$$

$$+ \int_{\Omega_u} \left[\left(\frac{\partial G}{\partial \ddot{u}_i^s} - \rho a_0 \lambda_i \right) \delta \dot{u}_i^s \right]_0^T d^3\mathbf{x} + \int_{\Omega_u} \left[\left(\frac{\partial G}{\partial \dot{u}_i^s} - \frac{d}{dt}\frac{\partial G}{\partial \ddot{u}_i^s} + \rho a_0 \dot{\lambda}_i - \rho a_1 \lambda_i \right) \delta u_i^s \right]_0^T d^3\mathbf{x}$$

$$+ \int_0^T \int_{\Omega_u} \left[\partial_j \left(\bar{c}_{ijkl} * \partial_k \lambda_l \right) - \rho \bar{L}(t) * \lambda_i + \frac{\partial G}{\partial u_i^s} - \frac{d}{dt}\frac{\partial G}{\partial \dot{u}_i^s} + \frac{d^2}{dt^2}\frac{\partial G}{\partial \ddot{u}_i^s} + F * \frac{\partial G}{\partial \hat{u}_i^s}(-t) \right] \delta u_i^s d^3\mathbf{x} dt$$

$$+ \int_0^T \int_{\Omega_t} \left[\lambda_i \delta f_i - \delta\rho \lambda_i \ddot{u}_i^s - \partial_i \lambda_j \left(\delta\bar{c}_{ijkl} * \partial_k u_l^s \right) \right] d^3\mathbf{x} dt, \tag{32}$$

where

$$\bar{L}(t) = a_0 \ddot{\delta}(t) - a_1 \dot{\delta}(t) + a_2 \delta(t) + L^c(t),$$

in which the sign of the first-order time derivative term has changed compared to eq. (23) for $L(t)$. The implicit system derivatives $\delta\mathbf{u}^s$ and $\delta\dot{\mathbf{u}}^s$ are cancelled out from the sensitivity expression in eq. (32) by solving the adjoint response in conjunction with the initial conditions $\dot{u}_i^s(\mathbf{x}, 0) = u_i^s(\mathbf{x}, 0) = 0$. Consequently, the adjoint response is obtained by solving the following wave equation:

$$\rho \bar{L}(t) * \boldsymbol{\lambda} = \nabla \cdot \bar{\boldsymbol{\sigma}}^\lambda + \mathbf{f}_A(t)$$

$$\bar{\sigma}_{ij}^\lambda = \bar{c}_{ijkl} * \partial_k \lambda_l$$

$$\mathbf{f}_A(t) = \frac{\partial G}{\partial \mathbf{u}_r}(t) - \frac{d}{dt}\frac{\partial G}{\partial \dot{\mathbf{u}}_r}(t) + \frac{d^2}{d^2 t}\frac{\partial G}{\partial \ddot{\mathbf{u}}_r}(t) + F * \frac{\partial G}{\partial \hat{u}_i^s}(-t) \tag{33}$$

with boundary conditions

$$\hat{\mathbf{n}} \cdot \bar{\boldsymbol{\sigma}}^{\lambda} = 0 \quad \text{on} \quad \Gamma_{\mathrm{u}}^{\mathrm{F}}; \qquad \boldsymbol{\lambda} = 0 \quad \text{on} \quad \Gamma_{\mathrm{u}}^{\mathrm{D}} \tag{34}$$

and terminal values

$$\boldsymbol{\lambda}(\mathbf{x}, T) = \mathbf{0}, \qquad \dot{\boldsymbol{\lambda}}(\mathbf{x}, T) = \mathbf{0}. \tag{35}$$

Introducing $\mathbf{u}^{\dagger}(\mathbf{x}, t) = \boldsymbol{\lambda}(\mathbf{x}, T - t)$, we can deduce that the adjoint stress tensor is given by

$$\bar{\sigma}_{ij}^{\dagger}(t) = \bar{c}_{ijkl} * \partial_k u_l^{\dagger}(t) \tag{36}$$

and the terminal value problem in eq. (33) can be transformed, by replacing $\bar{L}(t)$ with $L(t)$ from eq. (23), into the following initial value problem:

$$\rho\, L(t) * \mathbf{u}^{\dagger} = \nabla \cdot \bar{\boldsymbol{\sigma}}^{\dagger} + \mathbf{f}_{\mathrm{A}}(T - t),$$

$$\bar{\sigma}_{ij}^{\dagger} = \bar{c}_{ijkl} * \partial_k u_l^{\dagger},$$

$$\mathbf{f}_{\mathrm{A}}(t) = \frac{\partial G}{\partial \mathbf{u}_{\mathrm{r}}}(t) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial \dot{\mathbf{u}}_{\mathrm{r}}}(t) + \frac{\mathrm{d}^2}{\mathrm{d}^2 t}\frac{\partial G}{\partial \ddot{\mathbf{u}}_{\mathrm{r}}}(t) + F * \frac{\partial G}{\partial \hat{u}_i^{\mathrm{s}}}(-t) \tag{37}$$

with boundary conditions

$$\hat{\mathbf{n}} \cdot \bar{\boldsymbol{\sigma}}^{\dagger} = 0 \quad \text{on} \quad \Gamma_{\mathrm{u}}^{\mathrm{F}}; \qquad \mathbf{u}^{\dagger} = 0 \quad \text{on} \quad \Gamma_{\mathrm{u}}^{\mathrm{D}} \tag{38}$$

and terminal values

$$\mathbf{u}^{\dagger}(\mathbf{x}, 0) = \mathbf{0}, \qquad \dot{\mathbf{u}}^{\dagger}(\mathbf{x}, 0) = \mathbf{0}. \tag{39}$$

Hence, the equation governing absorption of the outgoing adjoint wavefield by the PML is exactly the same equation as that for the forward wavefield, that is, the problem remains self-adjoint even in the presence of PML, albeit with the adjoint force $\mathbf{f}_{\mathrm{A}}$ determined by the chosen misfit response. The natural choice for finding the adjoint field is thus to prefer (37) over (33) because using the latter would break the useful self-adjoint character of the problem. Upon solving the initial value problem, (37) reduces the variation $\delta\Phi = \delta\Phi_{\mathrm{A}}$ to

$$\delta\Phi = -\int_0^T \int_{\Omega_{\mathrm{t}}} \nabla\boldsymbol{\lambda}{:}\mathbf{c}{:}\nabla\mathbf{u} + \delta\rho\boldsymbol{\lambda} \cdot \ddot{\mathbf{u}}\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t = \int_{\Omega_{\mathrm{t}}} \left(\delta\log\rho\, K_{\rho}^{\mathrm{s}} + \delta\log\mathbf{c}{::}\mathbf{K}_c^{\mathrm{s}}\right) \mathrm{d}^3\mathbf{x} \tag{40}$$

with

$$\mathbf{K}_c^{\mathrm{s}} = -\int_0^T \nabla\mathbf{u}^{\dagger}(\mathbf{x}, t)\nabla\mathbf{u}(\mathbf{x}, T - t)\,\mathrm{d}t. \tag{41}$$

Expressing the elasticity tensor for an isotropic earth model in terms of bulk and shear moduli, that is, $c_{ijkl} = (\kappa - 2/3\mu)\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$, leads to

$$\delta\log\mathbf{c}{::}\mathbf{K}_c^{\mathrm{s}} = \delta\log\kappa^{\mathrm{s}}\, K_{\kappa}^{\mathrm{s}} + \delta\log\mu^{\mathrm{s}}\, K_{\mu}^{\mathrm{s}}. \tag{42}$$

The variation of the misfit function can then be written as

$$\delta\Phi = \int_{\Omega_{\mathrm{t}}} \left(\delta\log\rho\, K_{\rho}^{\mathrm{s}} + \delta\log\kappa^{\mathrm{s}}\, K_{\kappa}^{\mathrm{s}} + \delta\log\mu^{\mathrm{s}}\, K_{\mu}^{\mathrm{s}}\right) \mathrm{d}^3\mathbf{x}, \tag{43}$$

where the sensitivity kernels K_{ρ}^{s}, K_{κ}^{s} and K_{μ}^{s} are given by (see, e.g. Tromp *et al.* 2005, 2008; Liu & Tromp 2008)

$$K_{\rho}^{\mathrm{s}} = -\rho(\mathbf{x})\int_0^T \mathbf{u}^{\dagger}(\mathbf{x}, t) \cdot \ddot{\mathbf{u}}(\mathbf{x}, T - t)\,\mathrm{d}t, \tag{44}$$

$$K_{\kappa}^{\mathrm{s}} = -\kappa(\mathbf{x})\int_0^T \nabla \cdot \mathbf{u}^{\dagger}(\mathbf{x}, t)\nabla \cdot \mathbf{u}(\mathbf{x}, T - t)\,\mathrm{d}t, \tag{45}$$

$$K_{\mu}^{\mathrm{s}} = -2\mu(\mathbf{x})\int_0^T \mathbf{D}^{\dagger}(\mathbf{x}, t){:}\mathbf{D}(\mathbf{x}, T - t)\,\mathrm{d}t, \tag{46}$$

where $\mathbf{D} = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T) - \frac{1}{3}\nabla \cdot \mathbf{u}\mathbf{I}$ is the traceless deviatoric strain and $\mathbf{D}^{\dagger}$ its adjoint.

The definition of the sensitivity kernels can vary depending on the choice of parameters made to describe the model. For example, Liu & Tromp (2008) show in the context of seismic tomography, when the misfit function is designed to quantify the traveltime difference and phase delay between real and synthetic data, that the natural choice of model parameters is density ρ, shear wave speed V_s and compressional wave speed V_p, since traveltime differences and phase delays are more sensitive to these parameters. Based on the relationship:

$$\delta\log\rho K_{\rho}^{\mathrm{s}'} + \delta\log V_p K_{V_p}^{\mathrm{s}} + \delta\log V_s K_{V_s}^{\mathrm{s}} = \delta\log\rho\, K_{\rho}^{\mathrm{s}} + \delta\log\kappa^{\mathrm{s}}\, K_{\kappa}^{\mathrm{s}} + \delta\log\mu^{\mathrm{s}}\, K_{\mu}^{\mathrm{s}}, \tag{47}$$

one obtains the corresponding kernels as

$$K_\rho^{s'} = K_\rho + K_\mu + K_\kappa, \tag{48a}$$

$$K_{V_s}^s = 2\left(K_\mu - \frac{4}{3}\frac{\mu}{\kappa}K_\kappa\right), \tag{48b}$$

$$K_{V_p}^s = 2\left(1 + \frac{4}{3}\frac{\mu}{\kappa}\right)K_\kappa, \tag{48c}$$

with

$$\delta\Phi = \int_{\Omega_s} \left(\delta\log\rho K_\rho^{s'} + \delta\log V_p K_{V_p}^s + \delta\log V_s K_{V_s}^s\right)\mathrm{d}^3\mathbf{x}, \tag{49}$$

where $K_\rho^{s'}$ is called an impedance kernel, which is shown to be powerful in the context of reverse-time migration and can be used to locate and define interfaces with strong impedance contrasts (Zhu *et al.* 2009; Douma *et al.* 2010).

3.3 Boundary storage strategy for on-the-fly calculation of the sensitivity kernel

The numerical computation of the sensitivity kernels of eq. (44) requires that the forward wavefield in reverse time (also sometimes called the backpropagated wavefield) and the adjoint wavefield be simultaneously available. Several solutions have been proposed in the literature to implement that: (i) the most straightforward approach one can think of is to first calculate and save to disk the forward wavefield as a function of space and time, and read it back when needed in the adjoint simulation; unfortunately this approach is currently unrealistic in the 3-D case because the total amount of data to store as well as the resulting input/output time are extremely big; it is worth mentioning however that this approach is already feasible in 2-D, and should be feasible in 3-D in a decade or so; (ii) reconstruct the forward wavefield in reverse time from the saved last snapshot of the forward simulation simultaneously during the computation of the adjoint wavefield. This approach does not require any significant storage to disk, however it requires twice the amount of computer memory as the first approach (because the forward run must be redone backwards simultaneously with the adjoint run) and it also increases the computation time by a factor of 3/2 because three runs need to be performed instead of two (since the forward run needs to be done twice, once forwards and once backwards). These are small drawbacks that are perfectly acceptable in practice, however a more problematic issue is the fact that the approach is numerically stable only for media that conserve total energy, that is, acoustic or elastic media, but not for media that have dissipation such as viscoelastic materials for instance (Liu & Tromp 2008; Ammari *et al.* 2013), the reason being that while amplifying the fields to restore energy when going backwards the numerical schemes will also amplify the numerical noise and thus quickly blow up.

In the case of a PML, the above limitation of approach (ii) is problematic because PML is a lossy medium, that is, total energy is not conserved and thus the numerical process of undoing a lossy numerical simulation backwards in time is unstable because numerical roundoff noise gets amplified as well when restoring energy by going backwards (e.g. Symes 2007; Anderson *et al.* 2012; Ammari *et al.* 2013), even if the mathematical viscoelastic inverse problem itself remains tractable (Tarantola 1988); thus a disk checkpointing strategy needs to be implemented (e.g. Symes 2007; Anderson *et al.* 2012). An obvious way of circumventing this problem is to store the wavefield inside the whole PMLs (or absorbing contour in the case of viscous boundary conditions) for all the time steps (e.g. Liu & Tromp 2006; Symes 2007; Clapp 2008), but this is then similar to approach (i) (inside the PMLs only) and thus one can encounter the same problem of large disk storage and input/output time (e.g. Symes 2007; Clapp 2008). However, keeping in mind that the role of a infinite-domain truncating method is only to provide the values of the field along the artificial boundaries, that is, one does not care about the values inside the absorbing layer as long as the waves are absorbed efficiently at the entrance of the layer, one can use the known boundary values as Dirichlet forcing source terms while performing wave propagation backwards, that is, one can use values obtained in forward runs along the entrance of the PML as sources for the backward run without having to store all field values inside the PML. Thus, under the condition that backward wave propagation of the classical wave equation inside the main domain be stable, that is, the medium not be viscoelastic, our improved computational strategy enables efficient on-the-fly computation of the sensitivity kernel required in time-domain adjoint methods by storing the field values along the inner edge of the PMLs only without having to store the forward wavefield in the whole PML.

4 NUMERICAL DISCRETIZATION OF THE WAVE PROPAGATION EQUATIONS

In this section, we will show how to use a displacement-based Legendre SEM for spatial discretization of the wave propagation equation. The SEM is an accurate and efficient technique to model acoustic or seismic wave propagation, as it combines the flexibility of FE techniques with the high accuracy of pseudospectral techniques (see, e.g. Komatitsch & Tromp 1999; Paolucci *et al.* 1999; De Basabe *et al.* 2008; Seriani & Oliveira 2008; Tromp *et al.* 2008; Oliveira & Seriani 2011; Peter *et al.* 2011; Melvin *et al.* 2012, and reference therein). Furthermore, it is highly efficient both on serial and on large parallel computers because it uses tensorized basis functions and a perfectly diagonal mass matrix (Tromp *et al.* 2008; Peter *et al.* 2011).

Regarding time integration, we will describe two schemes, one for the convolution formulation of CFS-UPML and one for its ADE formulation. In the former case, we will use the explicit Newmark scheme combined with a recursive convolution technique, while in the

latter case we will use a two-level storage, fourth-order, six-stage, LDDRK scheme. We will also show that in the case of low-order elements the SEM can be made equivalent to classical lumped FE methods. We will only describe how to implement the forward wave problem because the implementation of the adjoint problem is exactly the same.

4.1 Weak form and Legendre spectral-element discretization

Let us denote by Ω_u the CFS-UPMLs, Γ_u^F its free surface boundary, which is a continuous extension of the free-surface boundary in the main domain, Γ_u^I its interface with the main domain and Γ_u^O its outer boundaries except the free surface; One can then rewrite the convolution formulation of CFS-UPML (9) in a weak form by dotting it with an arbitrary test function $\mathbf{w}$ and integrating by part over the CFS-UPMLs to get:

$$\int_{\Omega_u} \rho a_0 \mathbf{w} \cdot \ddot{\mathbf{u}} \, d\Omega + \int_{\Omega_u} \rho a_1 \mathbf{w} \cdot \dot{\mathbf{u}} \, d\Omega + \int_{\Omega_u} \rho a_2 \mathbf{w} \cdot \mathbf{u} \, d\Omega + \int_{\Omega_u} \nabla \mathbf{w} : \bar{\boldsymbol{\sigma}} \, d\Omega = -\int_{\Omega_u} \mathbf{w} \cdot \mathbf{R}^u d\Omega + \int_{\Gamma_u^F + \Gamma_u^I + \Gamma_u^O} \mathbf{w} \cdot (\bar{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}) d\Gamma, \tag{50}$$

where $\mathbf{R}^u$ denotes all the convolution terms that arise when convolving $L(t)$ with the displacement field. The last term vanishes because: (i) $\bar{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}} = 0$ on Γ_u^F owing to the free boundary condition, (ii) $\mathbf{w} = 0$ on Γ_u^O owing to the fixed boundary condition imposed on the outer edges of CFS-UPML, (iii) the integral evaluated along Γ_u^I inside the CFS-UPML cancels out with its value along Γ_u^I inside the main domain because of the perfectly matched character of CFS-UPML, that is, the two integrals are exactly the same but with an opposite sign.

The weak form of the classical wave equation without CFS-UPML and its spatial discretization based on the Legendre SEM has been described in detail for instance in Komatitsch & Tromp (1999), Tromp *et al.* (2008) and Peter *et al.* (2011). Proceeding in a similar fashion, for the above weak form of the wave equation with the convolution formulation of CFS-UPML we can discretize all the terms of (50) using the Legendre SEM. The resulting ordinary differential equation can be written as:

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{C}\dot{\mathbf{U}} + \mathbf{M}^U \mathbf{U} + \sum_e (\mathbf{KU})^e = -\sum_e (\mathbf{R}^u)^e, \tag{51}$$

where $\mathbf{U}$ denotes the global displacement vector and $\mathbf{M}\ddot{\mathbf{U}}$, $\mathbf{C}\dot{\mathbf{U}}$ and $\mathbf{M}^U \mathbf{U}$ are, respectively, the assembly of element-level contributions of the first three terms on the left-hand side of (50). The $\sum_e (\mathbf{KU})^e$ and $\sum_e (\mathbf{R}^u)^e$ terms are summations of the element-level contributions of the fourth term on the left-hand side and the first term on the right-hand side of (50), without assembly. As mentioned above, in the case of the Legendre SEM the $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{M}^U$ matrices are perfectly diagonal.

Following the same procedure, the ordinary differential equations obtained after Legendre SEM discretization of the ADE formulation of CFS-UPML can be written as

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{C}\dot{\mathbf{U}} + \mathbf{M}^U \mathbf{U} + \sum_e (\mathbf{KU})^e = -\sum_e (\mathbf{R}^u)^e, \tag{52a}$$

$$\dot{\mathbf{R}}^e = -\mathbf{B}^e \mathbf{R}^e + \mathbf{G}^e(t), \tag{52b}$$

where $\mathbf{R}^e$ denotes all the convolution terms that arise when applying $L(t)$ to the displacement field, or $F^{-1}[s_1 s_2 s_3 / (s_i s_k)]$ to the strain field. The $\mathbf{R}^e$ are evaluated at the element level, that is, they are never assembled. Based on (20), the coefficient matrix $\mathbf{B}^e$ is diagonal, and $\mathbf{G}^e(t)$ is the assembled vector of $g(t)$.

4.2 Time integration

4.2.1 Second-order time scheme

Following Hughes (1987), the explicit conditionally stable second-order Newmark time scheme with $\beta = 0$ and $\gamma = 1/2$ for (51) can be written as

$$\mathbf{U}_{n+1} = \mathbf{U}_n + \Delta t \mathbf{V}_n + \frac{\Delta t^2}{2} \mathbf{a}_n, \tag{53a}$$

$$\tilde{\mathbf{V}}_{n+1} = \mathbf{V}_n + \frac{\Delta t}{2} \mathbf{a}_n, \tag{53b}$$

$$\left(\mathbf{M} + \frac{\mathbf{C}\Delta t}{2} \right) \mathbf{a}_{n+1} = \sum_e (\mathbf{R})_{n+1}^e - \mathbf{C}\tilde{\mathbf{V}}_{n+1} - \mathbf{M}^U \mathbf{U}_{n+1} - \sum_e (\mathbf{KU})_{n+1}^e, \tag{53c}$$

$$\mathbf{V}_{n+1} = \tilde{\mathbf{V}}_{n+1} + \frac{\Delta t}{2} \mathbf{a}_{n+1}, \tag{53d}$$

where $\mathbf{U}_n$, $\mathbf{V}_n$ and $\mathbf{a}_n$ denote, respectively, the displacement, velocity and acceleration vectors evaluated at time $t_n = n\Delta t$, where Δt is the discrete time step. $\tilde{\mathbf{V}}_{n+1}$ is the predictor of $\mathbf{V}_{n+1}$. Since matrices $\mathbf{M}$ and $\mathbf{C}$ are diagonal, the above time scheme is fully explicit and no linear system needs to be solved, which is one of the main advantages of using a SEM technique to discretize the problem. According to the

convolution formulation (9) of CFS-UPML, all the convolution terms can be written in the form $[\mathrm{e}^{-bt} H(t)] * g(t)$. Denoting its evaluation at t_{n+1} by Ψ_{n+1} and using the recursive convolution technique of Luebbers & Hunsberger (1992), we have

$$\Psi_{n+1} = \sum_{m=0}^{n} \int_{m}^{t_{m+1}} E^{n+1} g(\tau) \mathrm{d}\tau = \int_{t_n}^{t_{n+1}} E^{n+1} g(\tau) \mathrm{d}\tau + \mathrm{e}^{-b\Delta t} \sum_{m=0}^{n-1} \int_{t_m}^{t_{m+1}} E^{n} g(\tau) \mathrm{d}\tau, \tag{54}$$

where $E^n = \mathrm{e}^{-b(n\Delta t - \tau)} H(n\Delta t - \tau)$.

Assuming that $g(\tau)$ is approximately constant for $\tau \in [n\Delta t, (n+1)\Delta t]$, that is, the variation of $g(\tau)$ is small in that time interval, then by evaluating $g(\tau)$ either at $n\Delta t$ or at $(n+1)\Delta t$ the first-order recursive convolution scheme can be written as either

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t} \Psi_n + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t}) g_{n+1}, \tag{55}$$

or

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t} \Psi_n + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t}) g_n. \tag{56}$$

In terms of accuracy order, one can use either (55) or (56) together with (53) because they are both first-order accurate, but the choice of the recursive convolution scheme used combined with the Newmark scheme affects the stability of (53). In Section 5, we will show that long-time stability can only be achieved by using (55) together with (53), which is thus the formulation to use.

Matzen (2011) has shown that it is also possible to develop a second-order recursive convolution scheme that retains the accuracy order of the Newmark scheme, which thus improves the absorbing efficiency of CFS-UPML after discretization. To do so, based on the assumption that $g(t) = 0$ and the fact that $\mathrm{e}^{-bt} H(t) = 0$ when $t \le 0$, that is, the medium is at rest when the simulation starts, we have

$$\Psi_{n+1} = \sum_{m=0}^{n+1} \int_{t_{m-\frac{1}{2}}}^{t_{m+\frac{1}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau = \int_{t_{n+\frac{1}{2}}}^{t_{n+\frac{3}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau + \int_{t_{-\frac{1}{2}}}^{t_{\frac{1}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau + \sum_{m=1}^{n} \int_{t_{m-\frac{1}{2}}}^{t_{m+\frac{1}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau \tag{57}$$

with

$$\sum_{m=1}^{n} \int_{t_{m-\frac{1}{2}}}^{t_{m+\frac{1}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau = \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} E^{n+1} \mathbf{u}(\tau) \mathrm{d}\tau + \mathrm{e}^{-b\Delta t} \left[\Psi_n - \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} E^{n} g(\tau) \mathrm{d}\tau - \int_{t_{-\frac{1}{2}}}^{t_{\frac{1}{2}}} E^{n} g(\tau) \mathrm{d}\tau \right]. \tag{58}$$

Substituting (58) into (57), we get

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t} \Psi_n + \int_{t_{n+\frac{1}{2}}}^{t_{n+\frac{3}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau + \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau - \mathrm{e}^{-b\Delta t} \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} E^{n} g(\tau) \mathrm{d}\tau. \tag{59}$$

Making the same assumption as Matzen (2011) that the variation of $g(\tau)$ is small in the interval $[(n - \frac{1}{2})\Delta t, (n + \frac{1}{2})\Delta t]$, that is, we can assume that $g(\tau)$ is approximately constant in that interval, we set $g(\tau) = g_n$. Then, based on eq. (59) we get the second-order recursive convolution scheme

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t} \Psi_n + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2}) g_{n+1} + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2}) \mathrm{e}^{-b\Delta t/2} g_n. \tag{60}$$

Let us mention that (60) has already been given in Wang *et al.* (2006) but without derivation, and that a slightly different second-order recursive convolution scheme can be found in Rylander & Jin (2004). When b is large enough, the coefficients in (60) can be computed directly, however when b is very small we resort to a third-order Taylor expansion to evaluate $\frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})$ and $\frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2}) \mathrm{e}^{-b\Delta t/2}$ to avoid divisions between very small numbers, which can be numerically inaccurate. In practice, we switch to such a Taylor expansion when $b < 10^{-6}$.

4.2.2 Higher order LDDRK time scheme

In the case of long simulations with a very large number of time steps, it is necessary to resort to a high-order time scheme with low numerical dissipation and low numerical dispersion errors in order for the simulation to remain accurate. However, it is uneasy to extend a recursive convolution scheme to high order, but such difficulty can be avoided easily by introducing an ADE formulation (Gedney & Zhao 2010; Martin *et al.* 2010). Furthermore, trying to increase the accuracy of the simulation by simply decreasing the time step or increasing the accuracy order of the time scheme is inefficient, especially when both low dispersion and low dissipation errors are required to simulate waves over a wide frequency band. In such a situation, which is rather common in practice, it is better to resort to a time scheme designed specifically to reduce these sources of error. Thus, we select an explicit conditionally stable highly optimized two-level storage, fourth-order, six-stage, LDDRK scheme. As most Runge–Kutta methods, that scheme is designed for first-order differential equations and thus we rewrite (52) as

$$\dot{\mathbf{V}} = \mathbf{M}^{-1} \left[\sum_{e} \mathbf{R}^e - \mathbf{C}\mathbf{V} - \mathbf{M}^U \mathbf{U} - \sum_{e} (\mathbf{K}\mathbf{U})^e \right], \tag{61a}$$

Table 1. Coefficients α_i, β_i and γ_i used in the optimized LDDRK scheme, adapted from table 1 in Berland *et al.* (2006).

i	β_i	γ_i	c_i
1	0.0	0.032918605146	0.0
2	-0.737101392796	0.823256998200	0.032918605146
3	-1.634740794341	0.381530948900	0.249351723343
4	-0.744739003780	0.200092213184	0.466911705055
5	-1.469897351522	1.718581042715	0.582030414044
6	-2.813971388035	0.27	0.847252983783

$$\dot{\mathbf{U}} = \mathbf{V}, \tag{61b}$$

$$\dot{\mathbf{R}}^e = -\mathbf{B}^e \, \mathbf{R}^e + \mathbf{G}^e(t). \tag{61c}$$

The LDDRK scheme for that equation can then be written as

$$\mathbf{Z}_i^V = \beta_i \mathbf{Z}_{i-1}^V + \Delta t \mathbf{M}^{-1} \left[\sum_e \mathbf{R}_i^e - \mathbf{C}\mathbf{V}_i - \mathbf{M}^U \mathbf{U}_i - \sum_e (\mathbf{K}\mathbf{U})_i^e \right], \tag{62a}$$

$$\mathbf{Z}_i^U = \beta_i \mathbf{Z}_{i-1}^U + \Delta t \mathbf{V}_i, \tag{62b}$$

$$\mathbf{Z}_i^{R^e} = \beta_i \mathbf{Z}_{i-1}^{R^e} + \Delta t [-\mathbf{B}^e \, \mathbf{R}^e + \mathbf{G}^e(t_i)], \tag{62c}$$

$$\mathbf{V}_i = \mathbf{V}_{i-1} + \gamma_i \mathbf{Z}_i^V, \quad \mathbf{U}_i = \mathbf{U}_{i-1} + \gamma_i \mathbf{Z}_i^U, \quad \mathbf{R}_i^e = \mathbf{R}_{i-1}^e + \gamma_i \mathbf{Z}_i^{R^e}, \tag{62d}$$

where $i = 1, \ldots, 6$, $t_i = (n + c_i)\Delta t$ and $\mathbf{Z}_i^V$ and $\mathbf{Z}_i^U$ are intermediate variables. Six stages of computation are needed to go from time step n to $n + 1$. Thus, we denote variables $\mathbf{U}_n$, $\mathbf{V}_n$ and $\mathbf{R}_n^e$, that is, variables evaluated at starting time step n by $\mathbf{U}_0$, $\mathbf{V}_0$ and $\mathbf{R}_0^e$, since they correspond to stage 0 of the scheme, and similarly we denote $\mathbf{U}_{n+1}$, $\mathbf{V}_{n+1}$ and $\mathbf{R}_{n+1}^e$ by $\mathbf{V}_6$, $\mathbf{U}_6$ and $\mathbf{R}_6^e$, since they correspond to stage 6 of the scheme. When $\beta_0 = 0$ the scheme is fully explicit. Only two levels of storage are needed, which is very memory efficient compared to the classical fourth-order four-stage Runge–Kutta scheme (RK4). By optimizing the parameter sets β_i and γ_i based on a dispersion and dissipation analysis, nice low dispersion and low dissipation properties can be achieved and a high Courant–Friedrichs–Lewy (CFL) stability limit can be kept (Williamson 1980; Hu *et al.* 1996; Berland *et al.* 2006; Ketcheson 2008; Rajpoot *et al.* 2010; Tarrass *et al.* 2011; Toulorge & Desmet 2012). In what follows we use the parameters given in Table 1, which lead to a maximum CFL number that is 1.9 times that of the Newmark time scheme.

4.3 Approach introduced to stabilize the scheme when high-order spectral elements are used

We will show in Section 5 that when classical low-order FE methods are used, long-time stability is achieved numerically in all cases, that is, for the first-order convolution, second-order convolution and ADE formulations. However, when switching to high-order spectral elements for the spatial discretization and to using the second-order recursive convolution formula (60) together with the Newmark scheme (53) for the time integration of the CFS-UPML convolution (51), instabilities can sometimes appear. Although a detailed analysis of the instability mechanism is beyond the scope of this paper, the instabilities are caused by the incorrect match between the recursive convolution scheme and the Newmark scheme, and we can thus try to modify the convolution scheme to improve its match with the Newmark scheme and eliminate or drastically delay these instabilities.

To do so, we resort to the mean value theorem that is used to design the classical Newmark scheme (Hughes 1987):

$$\exists \tau \in [n\Delta t, (n + 1)\Delta t] \text{ such that } \dot{g}(\tau) = \frac{g((n + 1)\Delta t) - g(n\Delta t)}{\Delta t}, \tag{63}$$

when $g(t)$ is a continuous and differentiable function. The updated velocity in the Newmark scheme can thus be obtained by approximating

$$\dot{u}_{n+1} = \dot{u}_n + \int_{n\Delta t}^{(n+1)\Delta t} \ddot{u}(\tau)\mathrm{d}\tau \tag{64}$$

by

$$\dot{u}_{n+1} = \dot{u}_n + \Delta t \ddot{u}_\gamma, \tag{65}$$

where $\ddot{u}_\gamma = (1 - \gamma)\ddot{u}_n + \gamma \ddot{u}_{n+1}$. In a similar way, we can write the updated displacement as

$$\exists \beta \in \left[0, \frac{1}{2}\right] \text{ such that } u_{n+1} = u_n + \Delta t \dot{u}_n + \frac{\Delta t^2}{2}\ddot{u}_\beta, \tag{66}$$

where $\ddot{u}_\beta = (1 - 2\beta)\ddot{u}_n + 2\beta\ddot{u}_{n+1}$. Different choices of β and γ will result in different stability properties for the Newmark scheme (Hughes 1987). Following the same idea of using the mean value theorem for integrals, based on (59) we get

$$\exists \theta \in [t_{n-\frac{1}{2}}, t_{n+\frac{1}{2}}] \text{ such that } \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} E^{n+1} g(\tau) \mathrm{d}\tau = g_\theta \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} E^{n+1} \mathrm{d}\tau. \tag{67}$$

In accordance with the expression of $\ddot{u}_\gamma$ and $\ddot{u}_\beta$ introduced in the derivation of the classical Newmark scheme above, we define g_θ as $(1 - \theta)g_{n+\frac{1}{2}} + \theta g_{n-\frac{1}{2}}$, where $\theta \in [0, 1]$. Since $g_{n+\frac{1}{2}}$ and $g_{n-\frac{1}{2}}$ cannot be used directly because the discrete value of g is known only at time steps $n - 1$ and n, based on a Taylor expansion we get

$$g_\theta = (1 - \theta)g_{n+\frac{1}{2}} + \theta g_{n-\frac{1}{2}} = g_n + \frac{(1 - 2\theta)\Delta t}{2}\dot{g}_n + \frac{\Delta t^2}{8}\ddot{g}_n + O(\Delta t^3). \tag{68}$$

Substituting (68) into (60), we get

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t}\Psi_n + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\left[g_{n+1} + \frac{(1 - 2\theta)\Delta t}{2}\dot{g}_{n+1} + \frac{\Delta t^2}{8}\ddot{g}_{n+1}\right] + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\mathrm{e}^{-b\Delta t/2}\left[g_n + \frac{(1 - 2\theta)\Delta t}{2}\dot{g}_n + \frac{\Delta t^2}{8}\ddot{g}_n\right]$$
$$+ O(\Delta t^4), \tag{69}$$

where we used the fact that $\frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2}) \cdot O(\Delta t^3) = O(\Delta t^4)$ and $\frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\mathrm{e}^{-b\Delta t/2} \cdot O(\Delta t^3) = O(\Delta t^4)$. If we couple (69) with the Newmark scheme directly, that is, if we use (69) to compute convolution terms such as $[\mathrm{e}^{-\alpha_0 t} H(t)] * \mathbf{U}$ inside $\sum_e (\mathbf{R})_{\mathbf{n+1}}^\mathrm{e}$ in (53c), the resulting time scheme is implicit. This comes from the fact that the terms $\dot{g}_{n+1}$ used for instance in $\mathbf{V}_{n+1}$ and $\ddot{g}_{n+1}$ used for instance in $\mathbf{a}_{n+1}$ are unknown when they are being used, the only already-computed variables before the computation of the convolution terms being $\Psi_n, g_n, \dot{g}_n, \ddot{g}_n, g_{n+1}$ and $\tilde{g}_{n+1}$ computed based on $\tilde{g}_{n+1} = \dot{g}_n + (\Delta t/2)\ddot{g}_n$. Thus, in order to avoid having to use an implicit time scheme, we need to avoid using $\dot{g}_{n+1}$ and $\ddot{g}_{n+1}$ directly in (69). Following (53 d) we have:

$$\dot{g}_{n+1} = \tilde{g}_{n+1} + \frac{\Delta t}{2}\ddot{g}_{n+1}. \tag{70}$$

Substituting (70) in (69) and omitting the $O(\Delta t^4)$ term, we get

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t}\Psi_n + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\left[g_{n+1} + \frac{(1 - 2\theta)\Delta t}{2}\tilde{g}_{n+1} + \frac{(3 - 4\theta)\Delta t^2}{8}\ddot{g}_{n+1}\right] + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\mathrm{e}^{-b\Delta t/2}\left[g_n + \frac{(1 - 2\theta)\Delta t}{2}\dot{g}_n + \frac{\Delta t^2}{8}\ddot{g}_n\right]. \tag{71}$$

Then, based on a Taylor expansion we see that

$$\frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\frac{\Delta t^2}{8}\ddot{g}_{n+1} - \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\mathrm{e}^{-b\Delta t/2}\frac{\Delta t^2}{8}\ddot{g}_n = O(\Delta t^4) \tag{72}$$

and thus we can reformulate (71) as

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t}\Psi_n + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\left[g_{n+1} + \frac{(1 - 2\theta)\Delta t}{2}\tilde{g}_{n+1}\right] + \frac{1}{b}(1 - \mathrm{e}^{-b\Delta t/2})\mathrm{e}^{-b\Delta t/2}\left[g_n + \frac{(1 - 2\theta)\Delta t}{2}\dot{g}_n + \frac{(1 - \theta)\Delta t^2}{2}\ddot{g}_n\right] \tag{73}$$

inside which we have omitted the new $O(\Delta t^4)$ term that arose due to the substitution of (72) into (71). In a similar way, we can derive the new second-order convolution scheme for $[\mathrm{e}^{-\alpha_0 t} H(t)] * [t^p u_i]$, where p is an integer. Details can be found in Appendix D.

Analytically, determining the proper value of θ to use can be difficult, in particular due to the fact that the stability analysis of PML is not well established mathematically in the literature. We thus resort to numerical tests to find the experimental range of values of θ that can be used in practice to get stable results. We find that in the case of regular and non-oversampled meshes, using (73) with a value of θ in about $[0, 1/2]$ in the 2-D case and θ in about $[0, 1/4]$ in the 3-D case is sufficient to ensure long-time numerical stability in practice. It can seem unusual to obtain a different range in the 2-D and 3-D cases, however that is normal because when using (73) with the Newmark scheme different convolution schemes will result in a different structure of the two terms $\sum_e (\mathbf{KU})^\mathrm{e}$ and $\sum_e (\mathbf{R^u})^\mathrm{e}$ in (51). In our numerical tests in Section 5, we will thus use $\theta = 1/2$ in the 2-D case and $\theta = 1/4$ in the 3-D case.

A second kind of numerical instability can be found when the mesh contains mesh elements inside the PML that have very small edges compared to the wavelengths of the waves (which can also happen in the case of meshes that contain highly distorted elements inside the PML, because these elements often have one or more small edges). That kind of instability can appear in both the convolution formulation and the ADE formulation. Whenever possible, one way of avoiding it is to mesh the PMLs with a regular and non-oversampled mesh, if the geometry permits it. It very often does, because the mesh inside the PMLs can be built by regular extrusion of the mesh of the faces at the entrance of the PML. When this is not possible, Kreiss & Duru (2013) have shown based on an eigenvalue stability analysis of a discretized CFS-UPML in the context of a second-order FD scheme that the occurrence of these instabilities can be very significantly delayed by activating the scaling factor of CFS-UPML, that is, using $\kappa > 1$ in (6b), and by increasing the frequency-shifting factors α in (6b). This is usually enough to stabilize the simulation over a sufficient total duration in practice, as will be illustrated in the next section. An intuitive explanation of why increasing the value of the κ_{x_i} can help to delay instabilities very significantly is that these numerical instabilities are associated with the non-physical

amplification of high-frequency waves, and using a larger κ_{x_i} will reduce the mesh resolution for these waves, since κ is a coordinate scaling factor. Zhang & Shen (2010) have shown that κ values up to about 8 can be used in practice with only a slight decrease of the absorbing efficiency for normal or near-normal incidence waves and with more efficient absorption for grazing incidence waves.

5 NUMERICAL EXAMPLES

Let us now provide several examples to show numerically that our formulation is efficient at absorbing elastic body waves with normal to near-grazing incidence as well as surface waves. Unless otherwise specified, in all examples we use a polynomial degree $N = 4$ for the Lagrange interpolants inside each spectral element. Following Collino & Tsogka (2001), in the CFS coordinate stretching function (6b) we set d_{x_i} as

$$d_{x_i} = C_{x_i} d_0 \left(\frac{x_i}{L}\right)^{N_{x_i}}, \tag{74}$$

where L is the thickness of the CFS-UPML, x_i is the distance along the normal direction measured from the interface, N_{x_i} is a real number greater than 1, and

$$d_0 = -(N_{x_i} + 1)c_p^{\max} \log(R_c)/(2L). \tag{75}$$

The values of N_{x_i}, C_{x_i}, R_c can be optimized to obtain good absorption (Gedney 1998) and thus here we set $N_{x_i} = 2, C_{x_i} = 1, R_c = 0.001$ accordingly. Following Gedney (1998), we also set

$$\alpha_{x_i} = \pi f_0 \left(1 - \frac{x_i}{L}\right), \tag{76}$$

where f_0 is the dominant frequency of the source. The advantages and disadvantages of using a varying κ_{x_i} are discussed in detail in Zhang & Shen (2010). The main advantage is that it makes CFS-UPML more efficient at absorbing grazing incident waves, but a disadvantage is that it slightly decreases the absorbing efficiency for normal or near-normal incidence waves. Since to our knowledge no optimized profile of κ_{x_i} is discussed in the literature, we simply define κ_{x_i} as

$$\kappa_{x_i} = \kappa_0 + \kappa_1 \frac{x_i}{L}, \tag{77}$$

where κ_0 and κ_1 are positive real numbers, with $\kappa_0 \geq 1$. In the following 2-D examples for regular meshes, we will set $\kappa_0 = 1$ and $\kappa_1 = 0$ because for regular meshes the absorbing efficiency of CFS-UPML for grazing incident waves is very good even without activating the κ_{x_i} scaling factors of PML. However, in our realistic 3-D example with a distorted mesh that will contain small element edges we will set $\kappa_0 = 4$ and $\kappa_1 = 4$ to delay the occurrence of instabilities and to improve the absorbing efficiency of CFS-UPML for grazing incidence waves.

5.1 Simple homogeneous infinite 2-D case

Let us consider a square 2-D medium surrounded by four CFS-UPML with elastic properties $c_p = 3 \text{ km s}^{-1}$, $c_s = 1.732 \text{ km s}^{-1}$ and $\rho = 2.7 \text{ g cm}^{-3}$. An explosive source with a Ricker source time function is located at the centre of the model, with an onset time $t_0 = 0.15 \text{ s}$ and a dominant frequency $f_0 = 8 \text{ Hz}$. In the case of spectral elements of polynomial degree $N = 2$ or $N = 3$, we use a regular mesh containing 100×100 square elements of size $\Delta = 0.02 \text{ km}$ each and the CFS-UPML consists of the six outermost layers of the mesh, while in the case of spectral elements with $N = 4, 5$ or 6 we use a regular mesh with 50×50 square elements of size $\Delta = 0.04 \text{ km}$ and the CFS-UPML consists of the three outermost layers. We do not go beyond $N = 6$ because such high values are almost never used for elastic wave propagation in practice in the literature, even in the absence of ABCs. We use a time step $\Delta t = 0.0008 \text{ s}$ chosen close to the CFL stability limit of the scheme in the main domain without PML for spectral elements with $N = 6$, which is the most restrictive case in terms of the stability condition because it contains the smallest spacing between two adjacent Gauss–Lobatto–Legendre integration points. To facilitate comparison of the results, we use that time step for all the cases.

We study the decay of energy with time in the whole domain (including the CFS-UPMLs) to check the efficiency as well as the long-time stability of our formulation by evaluating the total energy, which is the sum of the kinetic and potential energies:

$$E = \frac{1}{2}\rho\|\mathbf{V}\|^2 + \frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D}\sigma_{ij}\varepsilon_{ij}. \tag{78}$$

We compute a total of 600 000 time steps in the case of $N = 2$ and 3, and 300 000 time steps in the case of $N = 4, 5$ and 6. In Figs 3 and 4, we see that long-time stability is achieved for both the convolution and the ADE formulations in the case of low-order spectral elements, that is, in the case of a classical low-order mass-lumped FE method. Note that once the spatial resolution is good enough to accurately sample all the waves radiated by the source, increasing the polynomial degree N in the SEM does not have a strong effect on the absorbing efficiency of CFS-UPML. However, when high-order spectral elements are used the type of recursive convolution scheme has a strong effect on long-time stability of the implementation, as illustrated in Fig. 3: when we use the original second-order recursive convolution scheme (60) coupled with the Newmark time scheme instabilities can occur, while when using the new second-order recursive convolution scheme (73) simulations are
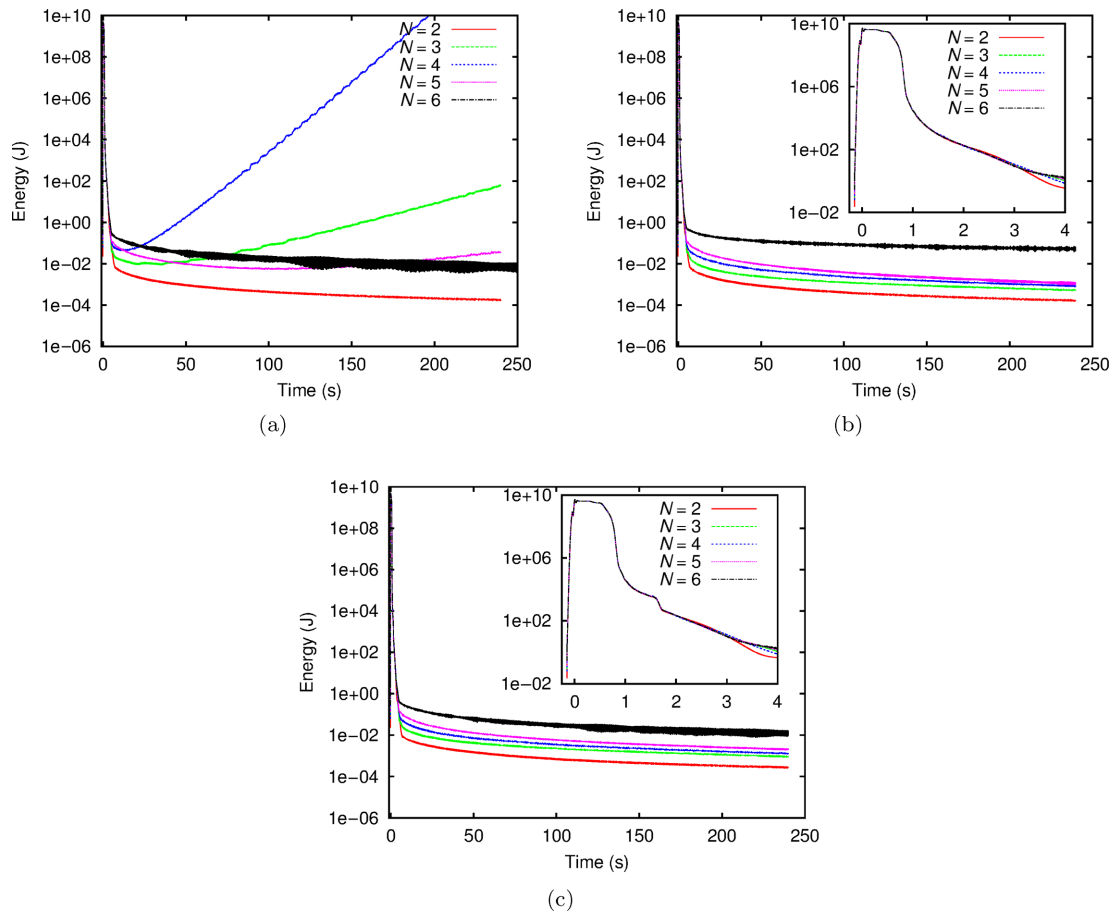
Figure 3. Time evolution of total energy for the SEM with different polynomial degrees *N* of the Lagrange interpolants when different types of recursive convolutions are used: (a) the original second-order recursive convolution scheme (60), (b) the new second-order recursive convolution scheme (73), (c) the first-order recursive convolution scheme (55). In (a), we do not have a clear understanding of the instability mechanism induced by the PML, which is still an open and difficult problem due to the current lack of complete mathematical tools to perform such analysis, as pointed out in Joly (2012). Duru (2014) also observes that using fourth-order polynomials leads to more unstable results compared to fifth- and sixth-order polynomials.

numerically stable. The same is true in the case of a first-order recursive convolution scheme: instabilities can occur when using (56) but do not occur when using (55). Let us also mention that when switching to the ADE formulation all simulations are numerically stable, as shown in Fig. 4(a). Furthermore, in the close-up of Fig. 4(b) it is clear that spurious reflections are smaller when a high-order time scheme is used.

5.2 Plane wave incidence in a semi-infinite 2-D medium

When the objects under study are located at large distance from the source, the assumption that incident waves are plane waves may be reasonable. However, implementing incident plane waves in a time-domain method based on a grid can be uneasy, in particular in the presence of absorbing conditions, due to the finite extension of the grid and to the contradiction between the absorbing conditions and the plane wave condition, which assumes an infinite or semi-infinite spatial extension. Let us thus show how to introduce a plane wave in a 2-D semi-infinite domain truncated by CFS-UPMLs in the context of time-domain simulations (more details can be found in Appendix E). We follow the method given in Bielak & Christiano (1984), Komatitsch *et al.* (1999), Chevrot *et al.* (2004), Delavaud (2007) and Godinho *et al.* (2009) in the case without PML and extend it to the case with PML. For simplicity, here we assume that outside the region under study the background semi-infinite domain is homogeneous (Bielak & Christiano 1984; Delavaud 2007). We also assume that the top surface is a flat free surface. It is important to mention that the same process can easily be used to handle the 3-D case, and that when the background model is heterogeneous (for instance, consisting of flat layers instead of a homogeneous half-space) the same technique can be employed, using inexpensive quasi-analytical or numerical techniques to compute the incident wavefield in that layered background model, for instance a discrete-wavenumber method (Bouchon 1981) used in conjunction with a reflectivity method (Müller 1985), or a Direct Solution Method (Monteiller *et al.* 2013).

The key idea is to use domain decomposition to decompose the truncated domain into two parts: one part dedicated to the computation of the outgoing wavefield diffracted from heterogeneities subjected to the incident plane wave, and another part dedicated to the computation of the total wavefield, that is, the sum of the incident and diffracted wavefields, as illustrated in Fig. 5(a). We label the first part Ω_1 and the second part Ω_2. A second key idea is that the incident wavefield is known analytically outside of the CFS-UPMLs, since it is a plane wave.
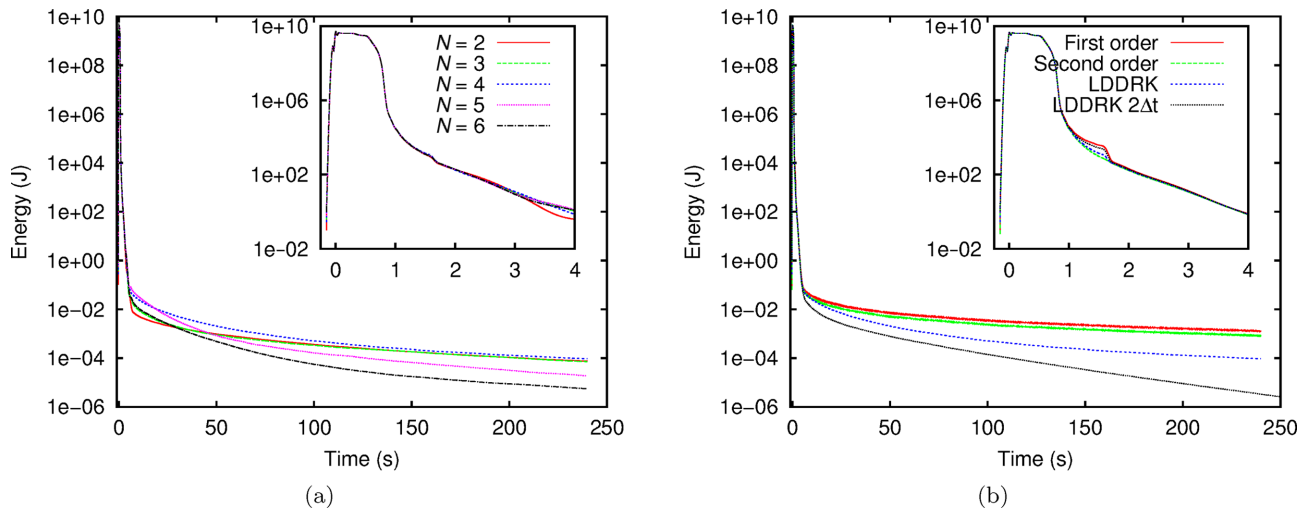
Figure 4. (a) For the SEM with different polynomial degrees N of the Lagrange interpolants, time evolution of total energy in the case of the ADE formulation with the fourth-order LDDRK time scheme of eq. (62); (b) For the same SEM mesh but with polynomial degree $N = 4$ only, time evolution of total energy for the convolution formulation with the first-order recursive convolution scheme of eq. (55) or the new second-order recursive convolution scheme of eq. (73), and for the ADE formulation with the fourth-order LDDRK time scheme of eq. (62).
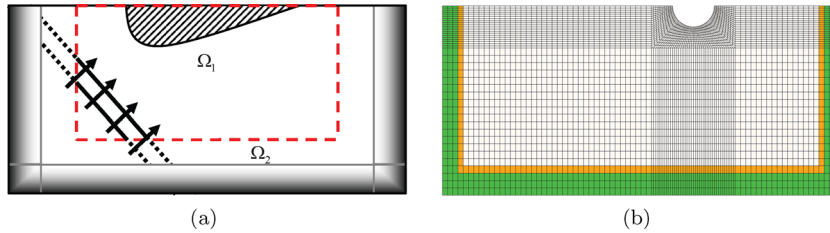
Figure 5. (a) Sketch of the implementation of an incident plane wave in the time domain in a model of finite size, where Ω_1 denotes the region in which the total wavefield is calculated and Ω_2 the region in which only the outgoing wavefield is calculated; (b) SEM mesh used for the 2-D canyon example, in which Ω_1 is the white region and Ω_2 is the union of the green and orange regions.

Intuitively, one obvious choice for Ω_1 would be the whole CFS-UPMLs (and only them). However, it turns out that such a choice would make the implementation much more complex because of the presence of additional terms such as $a_1\dot{\mathbf{u}}$ and $a_2\mathbf{u}$ that appear in the weak form of the CFS-UPML wave eq. (50), which would complicate the evaluation of the diffracted wavefield obtained by subtracting the incident wavefield from the total wavefield on the basis of eq. (51) compared with (E4). Such difficulty can be avoided by defining Ω_1 as the union of all the CFS-UPMLs and a small part of the interior domain, as shown in the SEM mesh of Fig. 5(b), in which we define Ω_1 as all the
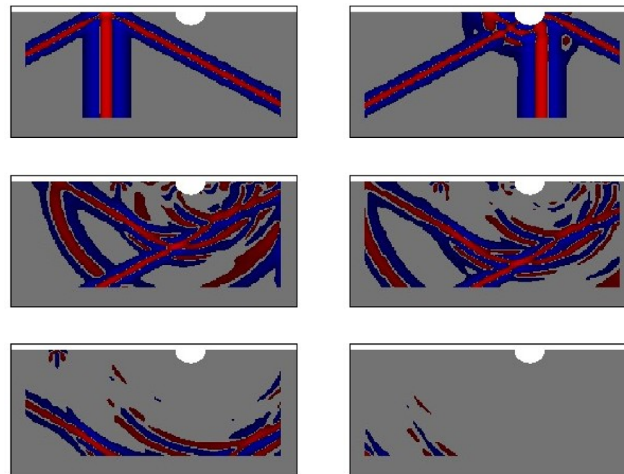

Figure 6. Snapshots of the horizontal component of the total displacement vector in the main domain represented by the white elements in the mesh shown in Fig. 5(b). The time steps represented go from 200 to 7200 from left to right and then top to bottom, with an interval of 1600 time steps. Positive values are represented in red and negative values in blue.

CFS-UPMLs plus one layer of elements inside, and Ω_2 as the rest of the domain. We then apply the incident wavefield on the interface between Ω_1 and Ω_2, that is, along the adjacent layer of elements located outside of the interface, which is shown by the yellow elements in Fig. 5(b).

Let us illustrate this technique by studying the response of a semi-circular canyon embedded in a homogeneous half-space subjected to a plane S wave with an incidence angle of 60°, as studied previously for instance in Komatitsch *et al.* (1999) and Godinho *et al.* (2009) but without ABCs. The shape of that incident plane wave is a Ricker wavelet with a dominant frequency $f_0 = 1$ Hz. The size of the model is 19 m × 9 m and its top edge is a free surface located at $x = 0$. The centre of the semi-circular canyon is located at (0, 0). The density is 1 kg m^{-3}, the P wave velocity is 2 m s^{-1} and the S wave velocity is 1 m s^{-1}, that is, the medium has a Poisson ratio of 0.25. The incident wavefield in the half-space, including the S wave reflected off the free surface and the converted S-to-P wave, can be computed analytically (table 5.1 of Aki & Richards 1980). Fig. 5(b) illustrates the SEM mesh used and the three layers of elements (in green) that are used to implement the CFS-UPML. We use a time step of $\Delta t = 0.002$ s. On the snapshots showing the total wavefield inside the main domain in Fig. 6, large diffraction as well as P-SV conversion due to presence of the canyon can be observed. It is clear that the CFS-UPML absorbs the outgoing wavefield, which consists of P, SV and Rayleigh waves, very efficiently, without creating artefacts in the incident plane wave system.

5.3 Banana–doughnut sensitivity kernel for a 2-D *P-SV* case

Let us now compute a slightly modified version of example 7.2.2 of Tromp *et al.* (2005) to show the significant improvement obtained when using CFS-UPML in the context of adjoint methods. Different from their example, the height of the model in our example is 60 km, that is, smaller than theirs, in order to make the configuration more difficult to compute accurately owing to the presence of more grazing incidence outgoing waves. The source–receiver geometry and the P-SV body wave arrivals are shown in Fig. 7. The model has a size of 200 km × 60 km, and is homogeneous with density $\rho = 2600$ kg m^{-3}, bulk modulus $\kappa = 5.20 \times 10^{10}$ Pa and shear modulus $\mu = 2.66 \times 10^{10}$ Pa. The source–time function used in the simulations is a Ricker wavelet:

$$h(t) = (2\alpha^3/\sqrt{\pi})(t - t_0)\exp[-\alpha^2(t - t_0)^2], \qquad (79)$$

where $t_0 = 8.0$ s, $\alpha = 2\tau_0/\tau$, $\tau_0 = 2.628$ s and $\tau = 4$ s is the approximate (apparent) duration of $h(t)$. The source is applied in the z-direction and we set $f_0 = 0.42$ Hz in the complex-shifting factor of eq. (76). The setup is shown in Fig. 7; the top surface is a free surface. The mesh contains 200 × 80 elements, with the three outer layers constituting the CFS-UPMLs. We use a time step $\Delta t = 0.002$ s.

In Fig. 8, for comparison we compute this example three times: using our CFS-UPML absorbing layers, using simpler viscous ABCs as in Tromp *et al.* (2005) and using a very large mesh used to provide a reference solution without any spurious reflection coming back into
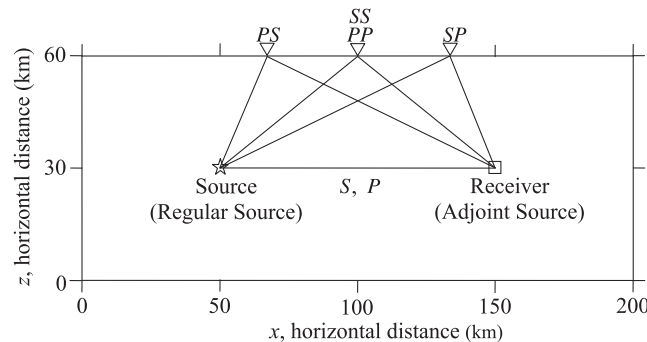
Figure 7. Sketch of the 2-D model dimensions and the source–receiver geometry as well as the possible body wave ray paths for the *P-SV* wavefield. The ray paths drawn are based on a homogeneous model. Modified from Tromp *et al.* (2005).
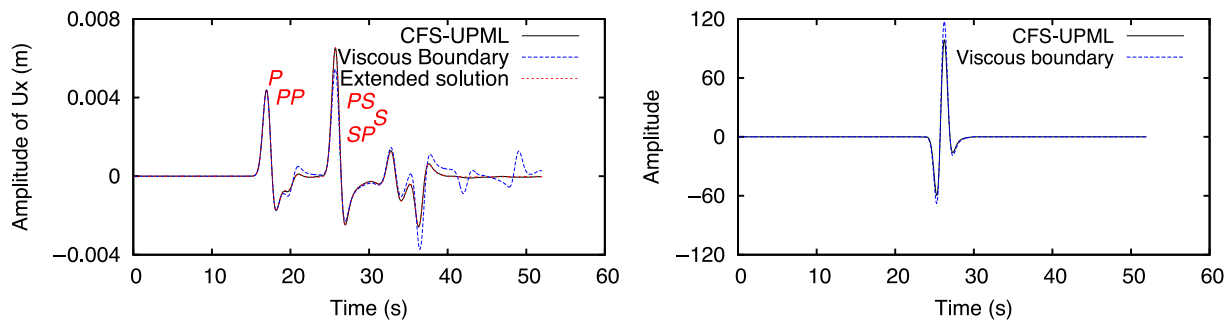
Figure 8. (a) Time history of the horizontal component of the displacement vector at the receiver shown in Fig. 7: with a simple viscous absorbing condition, with CFS-UPML and with a very large mesh used as a reference solution. The solution obtained with CFS-UPML, compared with the solution obtained with the viscous absorbing boundary condition, fits the reference solution almost perfectly; (b) Adjoint time source obtained for the *PS + SP* arrivals based on the horizontal component of the velocity vector recorded at the receiver shown in Fig. 7.
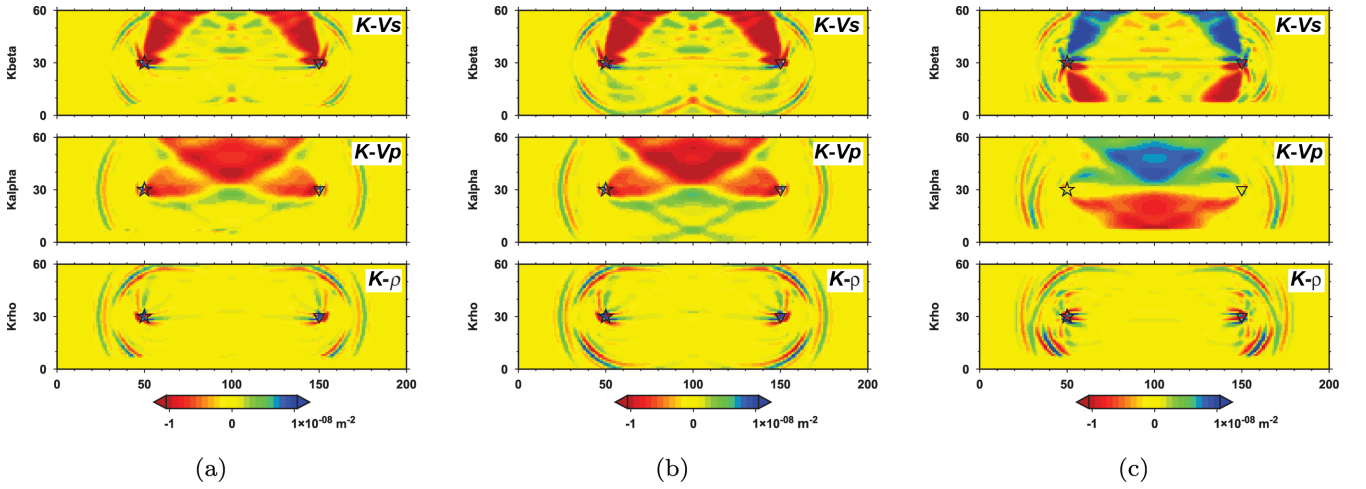
Figure 9. K'_ρ, K_{V_s} and K_{V_p} sensitivity kernels obtained from time reversing the $PS + SP$ arrivals (a) using CFS-UPML absorbing layers or (b) using a viscous absorbing boundary condition. In (c), we show the difference amplified by a factor of 5 for clarity.

the main domain. The much better accuracy obtained when using CFS-UPML is clear from Fig. 8(a), in which the numerical solution fits the reference solution almost perfectly, while the viscous ABC does not and is contaminated by large spurious reflections coming back from the edges of the domain. Thus, using the CFS-UPML technique allows for a much more accurate definition of the adjoint source, as shown in 8(b). Having cleaner forward seismograms also facilitates the use of automatic seismic phase picking software such as FlexWin (Maggi *et al.* 2009) that are often used to build the adjoint sources (see, e.g. Tape *et al.* 2010). As in Tromp *et al.* (2005), we have created that traveltime adjoint source by selecting the horizontal velocity component for the $PS + SP$ wave arrivals of the forward simulation at the receiver shown in Fig. 7:

$$f^\dagger(\mathbf{x}, t) = \frac{1}{M_r} w_r(T - t)\partial_t u_x(\mathbf{x}_r, T - t)\delta(\mathbf{x} - \mathbf{x}_r), \tag{80}$$

where $T = 60$ s is the total duration of the simulation, M_r is a normalization parameter equal to the L_2 norm of the horizontal velocity component at the receiver over the whole duration of the simulation, and w_r is the Welch window function:

$$w_r = 1 - \left[2\left(\frac{t - t_1}{t_2 - t_1}\right) - 1\right]^2, \tag{81}$$

where t_1 is the start time and t_2 the end time of the selected $PS + SP$ wave arrivals. In this example, $t_1 \simeq 23$ s and $t_2 \simeq 30$ s. We finally show the K'_ρ, K_{V_s} and K_{V_p} sensitivity kernels obtained using CFS-UPML in Fig. 9(a) and those obtained using the viscous absorbing condition in Fig. 9(b). It is clear that the use of the CFS-UPML absorbing layers leads to a much more accurate calculation of these kernels, as illustrated by the difference between the two results represented in Fig. 9(c). This difference is caused by the fact that different incident and adjoint wavefields are used in each sensitivity kernel computation. When using the viscous ABC the adjoint source cannot be defined as accurately as in the case of CFS-UPML because the incident wavefield is contaminated by spurious reflections generated on the artificial boundary, and in addition the adjoint wavefield itself is in turn contaminated by spurious reflections. Consequently, using an inaccurate artificial boundary condition not set sufficiently far away from the domain of interest can lead to inaccurate residuals computed from the incident field and used to define the adjoint source, and to inaccurate incident and adjoint fields that can lead to constructive spurious correlations in the gradient.

5.4 More realistic 3-D case

The main purpose of this example is to show how to tune parameters α_i, κ_i and d_i to achieve long-time numerical stability for the CFS-UPML when high-order spectral elements are used together with a non-structured FE mesh containing small element edges inside the CFS-UPMLs, following the stabilization techniques introduced in Section 4.3. In this example, we use the convolution formulation of CFS-UPML. We study a model consisting of a layer with significant surface topography over a half-space (Fig. 10 a). The elastic properties of the medium are $c_p = 4.0$ km s^{-1}, $c_s = 2.0$ km s^{-1} and $\rho = 2.6$ g cm^{-3} in the top layer and $c_p = 6.0$ km s^{-1}, $c_s = 3.436$ km s^{-1} and $\rho = 2.7$ g cm^{-3} in the half-space. The horizontal size of the model is 28 km along the x-direction and 30 km along the y-direction. It extends down to a depth of 25 km. CFS-UPML absorbing layers are implemented on all the sides of the model except the free surface. The non-structured mesh is coarsened in depth, where seismic velocities are higher and thus seismic wavelengths longer, as illustrated in Fig. 10(b), and is composed of a total of 185 600 spectral elements, which leads to a global grid that contains 13 107 116 points. We set an explosive source approximately in the middle of the model at a depth of 6.716 km, with a Ricker source time function

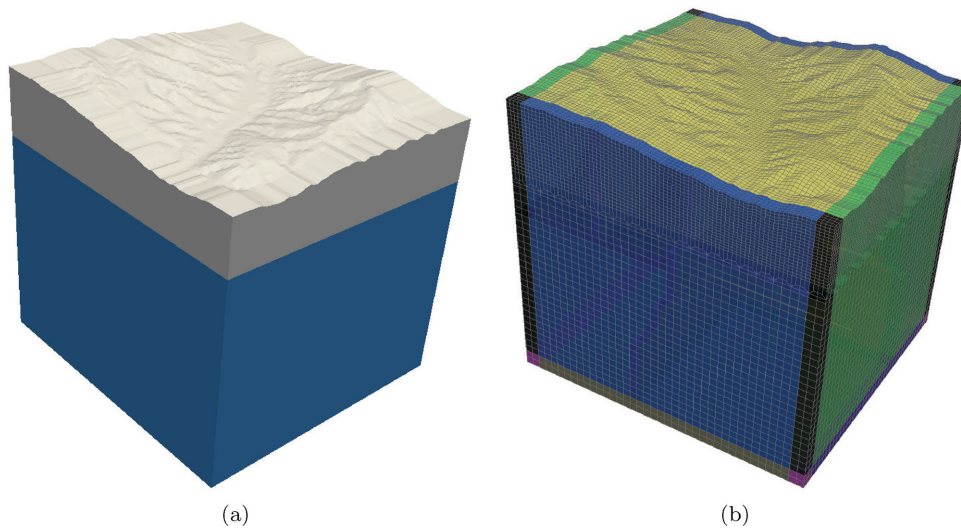$$h(t) = A\left(1 - 2[\pi f_0(t)]^2\right)\exp[-(\pi f_0)^2(t)^2] \tag{82}$$

(a)   (b)

Figure 10. (a) Three-dimensional two-layer geological model with significant topography on the top free surface and absorbing conditions that must be imposed along the five other faces. (b) Non-structured mesh created for that model, with higher mesh density used in the upper part of the model, where seismic wave speeds are smaller and thus seismic wavelengths shorter. In the mesh-density doubling layer, the elements are twice bigger at the bottom than at the top in each of the two horizontal spatial directions. Dark blue elements at the centre correspond to the main domain, and elements in all other colours correspond to the different CFS-UPML absorbing layers.
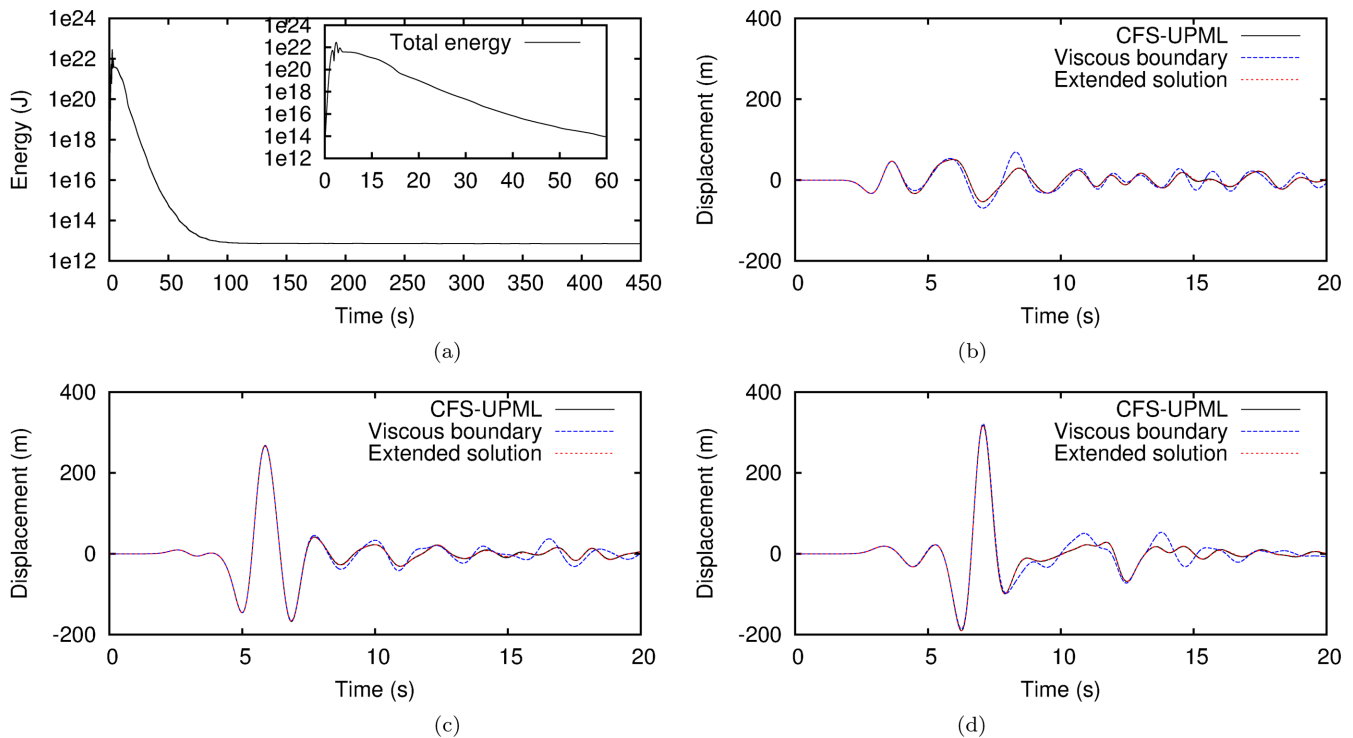

(a)   (b)

(c)   (d)

Figure 11. (a) Time evolution of total energy inside the main domain when CFS-UPML is used in the 3-D mesh of Fig. 10. We then show the *x*-component of the displacement vector recorded at receiver # 1 in (b), receiver # 2 in (c) and receiver # 3 in (d) for the three simulations performed: with CFS-UPML absorbing layers, with viscous absorbing boundary conditions (BC) and with a very large mesh used to provide a reference solution with no spurious waves coming back into the main domain from the edges of the mesh.

of amplitude $A = 1 \times 10^{18}$ and dominant frequency $f_0 = 0.5$ Hz. We use a time step of 0.003 s and propagate the signal for 150 000 time steps, that is, 450 s. Three receivers located on the free surface at (4.0 km, 25.0 km), (14.0 km, 25.0 km) and (24.0 km, 25.0 km), respectively, record the three components of the displacement vector. In order to delay the late-time instability, we activate the scaling factor by using $\kappa_0 = 4$ and $\kappa_1 = 4$ in (77) for the CFS-UPML elements located in the top layer.

In Fig. 11(a), we show the time evolution of total energy in the main domain, which keeps decaying up to about 150 000 time steps, thus showing that numerical stability has been achieved for such a total duration, keeping in mind that most numerical simulations of elastic wave propagation require far less than 150 000 time steps in practice and thus that such a typical number will be sufficient in most applications. In

Figs 11(b)–(d), we show the seismograms recorded at the three receivers and compare them to a reference numerical solution computed on a much larger model in order to evaluate the efficiency of CFS-UPML compared to the efficiency of viscous ABCs for the same mesh. *P* and *S* waves can be clearly observed, as well as multiples generated by transmission and reflection between the two media, and it is clear that the efficiency of CFS-UPML is much higher, since the seismograms obtained fit the reference solution far better.

6 CONCLUSIONS AND FUTURE WORK

We have derived both a convolution and an ADE formulation of CFS-UPML as an infinite-domain truncating method for both forward and adjoint wave propagation problems governed by the elastic wave equation formulated as a second-order system in displacement. This can help to significantly reduce the cost of numerical simulations by using smaller models and/or thinner mesh slices, which can be useful both in the case of forward simulations of a known model as well as in the case of adjoint-based simulations in the case of inverse imaging problems. We have implemented the CFS-UPML with both classical low-order mass-lumped FE and high-order SEM. In the case of classical low-order mass-lumped FEs, we have achieved numerical long-time stability. In the case of high-order Legendre spectral elements, we have shown that weak instabilities can in principle appear but that in practice long-time numerical stability can be achieved by using a stretching factor in the formulation to delay the occurrence of the weak instabilities as much as needed.

We have also rigorously derived the CFS-UPML formulation for time-domain adjoint elastic wave problems, which to our knowledge had never been done before. To facilitate its implementation, we have introduced an efficient computational approach that avoids having to solve a backward wave propagation problem inside the CFS-UPML, which is known to be highly ill-posed, and that enables on-the-fly calculation of sensitivity kernels for imaging without having to store the wavefield in the whole PMLs.

We have then demonstrated the absorbing efficiency of this improved CFS-UPML based on several numerical examples, both for forward and for adjoint problems. We have shown that by replacing the viscous boundary condition or tapered ALM that are still currently widely used in time-domain adjoint-based techniques for tomographic imaging with our CFS-UPML the quality of the sensitivity kernels built improves significantly. This should in turn improve the convergence rate of imaging techniques based on adjoint methods.

We have implemented our technique into our widely used open-source software packages called SPECFEM2D and SPECFEM3D_ Cartesian, which can be freely downloaded from the website of the Computational Infrastructure for Geodynamics (www.geodynamics.org).

It is worth mentioning that because we did not change the basic idea behind the CFS-UPML, for some (uncommon) kinds of anisotropic materials our CFS-PML formulation will also be ill-posed, as the classical PML and CFS-PML are for the same anisotropic media (Bécache *et al.* 2003).

It is also worth noting that the results presented in this paper can also be applied to time-domain forward or adjoint simulations implemented based on other displacement-based numerical schemes.

Due to the weak numerical instability found when implementing CFS-UPML in SEMs, we believe that in future work it would be worth trying to: (i) develop a mathematical well-posedness analysis of CFS-UPML written for the second-order displacement-based form of the elastic wave equation, since the numerical results presented in this paper tend to indicate that the proposed CFS-UPML formulation is weakly ill-posed in that case (but not in the case of classical lower order FEs), (ii) perform a detailed stability analysis to find the cause of the weak high-frequency instability that then appears. Replacing the Dirichlet (fixed) boundary condition with a viscous ABC at the end of the CFS-UPML (Collino & Monk 1998) to improve its absorbing efficiency and more importantly to avoid the growth of unstable cavity modes that may exist inside the CFS-UPML (Jiao *et al.* 2003) could also be of interest. Another topic of interest would be to extend this work to fluid–solid and/or viscoelastic media.

ACKNOWLEDGEMENTS

REFERENCES

Akçelik, V. *et al.*, 2003. High-resolution forward and inverse earthquake modeling on terascale computers, in *Proceedings of the SC'03 ACM/IEEE Conference on Supercomputing,* Phoenix, Arizona, USA, pp. 52–72.

Aki, K. & Richards, P.G., 1980. *Quantitative Seismology, Theory and Methods,* W. H. Freeman.

Alpert, B., Greengard, L. & Hagstrom, T., 2002. Nonreflecting boundary conditions for the time-dependent wave equation, *J. Comput. Phys.,* **180**(1), 270–296.

Ammari, H., Bretin, E., Garnier, J. & Wahab, A., 2013. Time-reversal algorithms in viscoelastic media, *Eur. J. appl. Math.,* **24**(4), 565–600.

Anderson, J.E., Tan, L. & Wang, D., 2012. Time-reversal checkpointing methods for RTM and FWI, *Geophysics,* **77**(4), S93–S103.

Arora, J.S. & Haug, E.J., 1979. Methods of design sensitivity analysis in structural optimization, *AIAA J.,* **17**(9), 970–974.

Astley, R.J. & Hamilton, J.A., 2006. The stability of infinite element schemes for transient wave problems, *Comput. Methods appl. Mech. Eng.,* **195**(29), 3553–3571.

Baffet, D., Bielak, J., Givoli, D., Hagstrom, T. & Rabinovich, D., 2012. Long-time stable high-order absorbing boundary conditions for elastodynamics, *Comput. Methods appl. Mech. Eng.,* **241–244**(1), 20–37.

Bamberger, A., Chavent, G., Hemons, C. & Lailly, P., 1982. Inversion of normal incidence seismograms, *Geophysics,* **47**(5), 757–770.

Bao, H., Bielak, J., Ghattas, O., Kallivokas, L.F., O'Hallaron, D.R., Shewchuk, J.R. & Xu, J., 1998. Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers, *Comput. Methods appl. Mech. Eng.,* **152**, 85–102.

Basu, U. & Chopra, A.K., 2004. Perfectly matched layers for transient elastodynamics of unbounded domains, *Int. J. Numer. Methods Eng.,* **59**, 1039–1074.

Bayliss, A. & Turkel, E., 1980. Radiation boundary conditions for wave-like equations, *Commun. Pure appl. Math.,* **33**, 707–725.

Bécache, E., Joly, P. & Tsogka, C., 2001. Fictitious domains, mixed finite elements and perfectly matched layers for 2-D elastic wave propagation, *J. Comput. Acoust.,* **9**(3), 1175–1203.

Bécache, E., Fauqueux, S. & Joly, P., 2003. Stability of perfectly matched layers, group velocities and anisotropic waves, *J. Comput. Phys.,* **188**(2), 399–433.

Bérenger, J.P., 1994. A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.,* **114**, 185–200.

Berland, J., Bogey, C. & Bailly, C., 2006. Low-dissipation and low-dispersion fourth-order Runge-Kutta algorithm, *Comput. Fluids,* **35**, 1459–1463.

Bielak, J. & Christiano, P., 1984. On the effective seismic input for non-linear soil-structure interaction systems, *Earthq. Eng. Struct. Dyn.,* **12**, 107–119.

Bouchon, M., 1981. A simple method to calculate Green's functions for elastic layered media, *Bull. seism. Soc. Am.,* **71**, 959–971.

Bozdağ, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.,* **185**(2), 845–870.

Cerjan, C., Kosloff, D., Kosloff, R. & Reshef, M., 1985. A nonreflecting boundary condition for discrete acoustic and elastic wave equation, *Geophysics,* **50**, 705–708.

Chavent, G., Dupuy, M. & Lemonnier, P., 1975. History matching by use of optimal control theory, *Soc. Petrol. Eng. J.,* **15**, 74–86.

Chen, Q., Tang, T. & Teng, Z., 2004. A fast numerical method for integral equations of the first kind with logarithmic kernel using mesh grading, *J. Comput. Math.,* **22**(2), 287–298.

Chevrot, S., Favier, N. & Komatitsch, D., 2004. Shear wave splitting in three-dimensional anisotropic media, *Geophys. J. Int.,* **159**(2), 711–720.

Chew, W.C. & Liu, Q., 1996. Perfectly matched layers for elastodynamics: a new absorbing boundary condition, *J. Comput. Acoust.,* **4**(4), 341–359.

Chew, W.C. & Weedon, W.H., 1994. A 3-D perfectly matched medium from modified Maxwell's equations with stretched coordinates, *Microw. Opt. Technol. Lett.,* **7**(13), 599–604.

Chung, Y.-S., Cheon, C., Park, I.-H. & Hahn, S.-Y., 2000. Optimal shape design of microwave device using FDTD and design sensitivity analysis, *IEEE Trans. Microw. Theory Tech.,* **48**(12), 2289–2296.

Clapp, R., 2008. Reverse time migration: saving the boundaries, Tech. Rep. SEP-136, Stanford Exploration Project, Stanford University, Stanford, California, USA, Unpublished.

Clayton, R. & Engquist, B., 1977. Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. seism. Soc. Am.,* **67**, 1529–1540.

Cohen, G. & Fauqueux, S., 2005. Mixed spectral finite elements for the linear elasticity system in unbounded domains, *SIAM J. Sci. Comput.,* **26**(3), 864–884.

Collino, F. & Monk, P., 1998. Optimizing the perfectly matched layer, *Comput. Methods appl. Mech. Eng.,* **164**, 157–171.

Collino, F. & Tsogka, C., 2001. Application of the PML absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media, *Geophysics,* **66**(1), 294–307.

De Basabe, J.D., Sen, M.K. & Wheeler, M.F., 2008. The interior penalty discontinuous Galerkin method for elastic wave propagation: grid dispersion, *Geophys. J. Int.,* **175**(1), 83–93.

Deeks, A.J. & Randolph, M.F., 1994. Axisymmetric time-domain transmitting boundaries, *J. Eng. Mech.,* **120**(1), 25–42.

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 (Venezuela), *PhD thesis,* Institut de Physique du Globe, Paris, France.

Dmitriev, M.N. & Lisitsa, V.V., 2011. Application of M-PML reflectionless boundary conditions to the numerical simulation of wave propagation in anisotropic media. Part I: reflectivity, *Sib. Zh. Vych. Mat.,* **14**(4), 333–344.

Douma, H., Yingst, D., Vasconcelos, I. & Tromp, J., 2010. On the connection between artifact filtering in reverse-time migration and adjoint tomography, *Geophysics,* **75**(6), S219–S223.

Drossaert, F.H. & Giannopoulos, A., 2007a. A nonsplit complex frequency-shifted PML based on recursive integration for FDTD modeling of elastic waves, *Geophysics,* **72**(2), T9–T17.

Drossaert, F.H. & Giannopoulos, A., 2007b. Complex frequency shifted convolution PML for FDTD modelling of elastic waves, *Wave Motion,* **44**(7–8), 593–604.

Du, X. & Zhao, M., 2010. A local time-domain transmitting boundary for simulating cylindrical elastic wave propagation in infinite media, *Soil Dyn. Earthq. Eng.,* **30**(10), 937–946.

Duru, K., 2014. The role of numerical boundary procedures in the stability of perfectly matched layers, in press.

Festa, G. & Nielsen, S., 2003. PML absorbing boundaries, *Bull. seism. Soc. Am.,* **93**, 891–903.

Festa, G. & Vilotte, J.P., 2005. The Newmark scheme as velocity-stress time-staggering: an efficient PML implementation for spectral-element simulations of elastodynamics, *Geophys. J. Int.,* **161**, 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**(20), L20306, doi:10.1029/2005GL024091.

Fichtner, A., 2010. *Full Seismic Waveform Modelling and Inversion, Advances in Geophysical and Environmental Mechanics and Mathematics,* Springer Verlag.

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**(1–2), 11–22.

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.P., 2009b. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.,* **179**(3), 1703–1725.

Gedney, S.D., 1998. The perfectly matched layer absorbing medium, in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain method,* Chap. 5, pp. 263–343, ed. Taflove, A., Artech House.

Gedney, S.D. & Zhao, B., 2010. An auxiliary differential equation formulation for the complex-frequency shifted PML, *IEEE Trans. Antennas Propag.,* **58**(3), 838–847.

Givoli, D., 2004. High-order local non-reflecting boundary conditions: a review, *Wave Motion,* **39**, 319–326.

Givoli, D. & Keller, J.B., 1990. Non-reflecting boundary conditions for elastic waves, *Wave motion,* **12**, 261–279.

Givoli, D. & Neta, B., 2003. High-order non-reflecting boundary scheme for time-dependent waves, *J. Comput. Phys.,* **186**(1), 24–46.

Givoli, D. & Patlashenko, I., 2004. Dirichlet-to-Neumann boundary condition for time-dependent dispersive waves in three-dimensional guides, *J. Comput. Phys.,* **199**(1), 339–354.

Godinho, L., Mendes, P.A., Tadeu, A., Cadena-Isaza, A., Smerzini, C., Sánchez-Sesma, F.J., Madec, R. & Komatitsch, D., 2009. Numerical simulation of ground rotations along 2D topographical profiles under the incidence of elastic plane waves, *Bull. seism. Soc. Am.,* **99**(2B), 1147–1161.

Greiner, W. & Reinhardt, J., 1996. *Field Quantization,* Springer.

Grote, M.J. & Keller, J.B., 1995. Exact nonreflecting boundary conditions for the time dependent wave equation, *SIAM J. appl. Math.,* **55**(2), 280–297.

Grote, M.J. & Kirsch, C., 2007. Nonreflecting boundary condition for time-dependent multiple scattering, *J. Comput. Phys.,* **221**(1), 41–62.

Grote, M.J. & Sim, I., 2011. Local nonreflecting boundary condition for time-dependent multiple scattering, *J. Comput. Phys.,* **230**(8), 3135–3154.

Guddati, M.N. & Tassoulas, J.L., 2000. Continued-fraction absorbing boundary conditions for the wave equation, *J. Comput. Acoust.,* **8**(1), 139–156.

Hagstrom, T. & Warburton, T., 2004. A new auxiliary variable formulation of high-order local radiation boundary conditions: corner compatibility conditions and extensions to first-order systems, *Wave Motion,* **39**(4), 327–338.

Hagstrom, T. & Warburton, T., 2009. Complete radiation boundary conditions: minimizing the long time error growth of local methods, *SIAM J. Numer. Anal.,* **47**(5), 3678–3704.

Hagstrom, T., Mar-Or, A. & Givoli, D., 2008. High-order local absorbing conditions for the wave equation: extensions and improvements, *J. Comput. Phys.,* **227**(6), 3322–3357.

Higdon, R.L., 1986. Absorbing boundary conditions for difference approximations to the multi-dimensional wave equation, *Math. Comput.,* **47**(176), 437–459.

Higdon, R.L., 1990. Radiation boundary conditions for elastic wave propagation, *SIAM J. Numer. Anal.,* **27,** 831–870.

Hu, F.Q., Hussaini, M.Y. & Manthey, J.L., 1996. Low-dissipation and low-dispersion Runge-Kutta schemes for computational acoustics, *J. Comput. Phys.,* **124**(1), 177–191.

Hughes, T.J.R., 1987. *The Finite Element Method, Linear Static and Dynamic Finite Element Analysis,* Prentice-Hall International.

Israeli, M. & Orszag, S.A., 1981. Approximation of radiation boundary conditions, *J. Comput. Phys.,* **41**(1), 115–135.

Jiao, D., Jin, J.-M., Michielssen, E. & Riley, D.J., 2003. Time-domain finite-element simulation of three-dimensional scattering and radiation problems using perfectly matched layers, *IEEE Trans. Antennas Propag.,* **51**(2), 296–305.

Joly, P., 2012. An elementary introduction to the construction and the analysis of perfectly matched layers for time domain wave propagation, *SeMA J.,* **57**(1), 5–48.

Keller, J. & Grote, M., 2000. Exact nonreflecting boundary condition for elastic waves, *SIAM J. appl. Math.,* **60**(3), 803–819.

Ketcheson, D.I., 2008. Highly efficient strong stability-preserving Runge-Kutta methods with low-storage implementations, *SIAM J. Sci. Comput.,* **30**(4), 2113–2136.

Komatitsch, D. & Martin, R., 2007. An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation, *Geophysics,* **72**(5), SM155–SM167.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral-element method for 3-D seismic wave propagation, *Geophys. J. Int.,* **139**(3), 806–822.

Komatitsch, D. & Tromp, J., 2003. A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation, *Geophys. J. Int.,* **154**(1), 146–153.

Komatitsch, D., Vilotte, J.P., Vai, R., Castillo-Covarrubias, J.M. & Sánchez-Sesma, F.J., 1999. The spectral-element method for elastic wave equations: application to 2D and 3D seismic problems, *Int. J. Numer. Methods Eng.,* **45**(9), 1139–1164.

Kosloff, R. & Kosloff, D., 1986. Absorbing boundaries for wave propagation problems, *J. Comput. Phys.,* **63**(2), 363–376.

Kreiss, G. & Duru, K., 2013. Discrete stability of perfectly matched layers for anisotropic wave equations in first and second order formulation, *BIT Numer. Math.,* **53**(3), 641–663.

Kreiss, H.-O., Petersson, N.A. & Yström, J., 2002. Difference approximations for the second-order wave equation, *SIAM J. Numer. Anal.,* **40**(5), 1940–1967.

Kucukcoban, S. & Kallivokas, L.F., 2011. Mixed perfectly matched layers for direct transient analysis in 2D elastic heterogeneous media, *Comput. Methods appl. Mech. Eng.,* **200**(1), 57–76.

Kucukcoban, S. & Kallivokas, L.F., 2013. A symmetric hybrid formulation for transient wave simulations in PML-truncated heterogeneous media, *Wave Motion,* **50**(1), 57–79.

Kuzuoglu, M. & Mittra, R., 1996. Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers, *IEEE Microw. Guid. Wave Lett.,* **6**(12), 447–449.

Lailly, P., 1983. The seismic inverse problem as a sequence of before-stack migrations, in *Proceedings of the Conference on Inverse Scattering, Theory and Application Expanded Abstracts,* Society of Industrial and Applied Mathematics, Philadelphia, PA, USA, pp. 206–220.

Le Dimet, F.-X. & Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus,* **38A,** 97–110.

Li, Y.F. & Matar, O.B., 2010. Convolutional perfectly matched layer for elastic second-order wave equation, *J. acoust. Soc. Am.,* **127,** 1318–1327.

Liao, Z.P. & Wong, H.L., 1984. A transmitting boundary for the numerical simulation of elastic wave propagation, *Int. J. Soil Dyn. Earthq. Eng.,* **3**(4), 174–183.

Liao, Z.P., Wong, H.L., Yang, B.P. & Yuan, Y.F., 1984. A transmitting boundary for transient wave analysis, *Sci. Sin.,* **27**(10), 1063–1076.

Liu, J. & Li, B., 2005. A unified viscous-spring artificial boundary for 3-D static and dynamic applications, *Sci. China E,* **48**(5), 570–584.

Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.,* **96**(6), 2383–2397.

Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.,* **174**(1), 265–286.

Luebbers, R.J. & Hunsberger, F., 1992. FDTD for Nth-order dispersive media, *IEEE Trans. Antennas Propag.,* **40**(11), 1297–1301.

Luo, Y., Modrak, R. & Tromp, J., 2014. Full waveform inversion strategies using adjoint methods, *Handbook of Geomathematics,* in press.

Lysmer, J. & Kuhlemeyer, R.L., 1969. Finite dynamic model for infinite media, *J. Eng. Mech. Div.,* **4,** 859–877.

Ma, S. & Liu, P., 2006. Modeling of the perfectly matched layer absorbing boundaries and intrinsic attenuation in explicit finite-element methods, *Bull. seism. Soc. Am.,* **96**(5), 1779–1794.

Maggi, A., Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.,* **178,** 257–281.

Marcinkovich, C. & Olsen, K.B., 2003. On the implementation of perfectly matched layers in a three-dimensional fourth-order velocity-stress finite-difference scheme, *J. geophys. Res.,* **108**(B5), 2276, doi:10.1029/2002JB002235.

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**(1), 333–344.

Martin, R., Komatitsch, D. & Gedney, S., 2008. A variational formulation of a stabilized unsplit convolutional perfectly matched layer for the isotropic or anisotropic seismic wave equation, *Comput. Model. Eng. Sci.,* **37**(3), 274–304.

Martin, R., Komatitsch, D., Gedney, S.D. & Bruthiaux, E., 2010. A high-order time and space formulation of the unsplit perfectly matched layer for the seismic wave equation using auxiliary differential equations (ADE-PML), *Comput. Model. Eng. Sci.,* **56**(1), 17–42.

Matzen, R., 2011. An efficient finite element time-domain formulation for the elastic second-order wave equation: a non-split complex frequency shifted convolutional PML, *Int. J. Numer. Methods Eng.,* **88**(10), 951–973.

Melvin, T., Staniforth, A. & Thuburn, J., 2012. Dispersion analysis of the spectral-element method, *Q. J. R. Meteorol. Soc.,* **138**(668), 1934–1947.

Meza-Fajardo, K.C. & Papageorgiou, A.S., 2008. A nonconvolutional, split-field, perfectly matched layer for wave propagation in isotropic and anisotropic elastic media: stability analysis, *Bull. seism. Soc. Am.,* **98**(4), 1811–1836.

Monteiller, V., Chevrot, S., Komatitsch, D. & Fuji, N., 2013. A hybrid method to compute short period synthetic seismograms of teleseismic body waves in a 3-D regional model, *Geophys. J. Int.,* **192**(1), 230–247.

Müller, G., 1985. The reflectivity method: a tutorial, *J. Geophys.,* **58,** 153–174.

Oliveira, S.P. & Seriani, G., 2011. Effect of element distortion on the numerical dispersion of spectral-element methods, *Commun. Comput. Phys.,* **9**(4), 937–958.

Paolucci, R., Faccioli, E. & Maggio, F., 1999. 3D response analysis of an instrumented hill at Matsuzaki, Japan, by a spectral method, *J. Seismol.,* **3**, 191–209.

Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.,* **186**(2), 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.

Rabinovich, D., Givoli, D., Bielak, J. & Hagstrom, T., 2011. A finite element scheme with a high-order absorbing boundary condition for elastodynamics, *Comput. Methods appl. Mech. Eng.,* **200**(23), 2048–2066.

Rabinovich, D., Givoli, D., Hagstrom, T. & Bielak, J., 2013. Stress-velocity complete radiation boundary conditions, *J. Comput. Acoust.,* **21**(2), 1–38.

Rajpoot, M.K., Sengupta, T.K. & Dutt, P.K., 2010. Optimal time-advancing dispersion-relation-preserving schemes, *J. Comput. Phys.,* **229**(10), 3623–3651.

Rickard, Y.S., Georgieva, N.K. & Tam, H.W., 2003. Absorbing boundary conditions for adjoint problems in the design sensitivity analysis with the FDTD method, *IEEE Trans. Microw. Theory Tech.,* **51**(2), 526–529.

Roden, J.A. & Gedney, S.D., 2000. Convolution PML (CPML): an efficient FDTD implementation of the CFS-PML for arbitrary media, *Microw. Opt. Technol. Lett.,* **27**(5), 334–339.

Rylander, T. & Jin, J.-M., 2004. Perfectly matched layer for the time domain finite element method, *J. Comput. Phys.,* **200**, 238–250.

Sarma, G.S., Mallick, K. & Gadhinglajkar, V.R., 1998. Nonreflecting boundary condition in finite-element formulation for an elastic wave equation, *Geophysics,* **63**(3), 1006–1016.

Semblat, J.F., Lenti, L. & Gandomzadeh, A., 2011. A simple multi-directional absorbing layer method to simulate elastic wave propagation in unbounded domains, *Int. J. Numer. Methods Eng.,* **85**(12), 1543–1563.

Seriani, G. & Oliveira, S.P., 2008. Dispersion analysis of spectral-element methods for elastic wave propagation, *Wave Motion,* **45**, 729–744.

Smith, W.D., 1974. A nonreflecting plane boundary for wave propagation problems, *J. Comput. Phys.,* **15**, 492–503.

Sochacki, J., Kubichek, R., George, J., Fletcher, W.R. & Smithson, S., 1987. Absorbing boundary conditions and surface waves, *Geophysics,* **52**(1), 60–71.

Stacey, R., 1988. Improved transparent boundary formulations for the elastic wave equation, *Bull. seism. Soc. Am.,* **78**(6), 2089–2097.

Symes, W.W., 2007. Reverse time migration with optimal checkpointing, *Geophysics,* **72**(5), SM213–SM221.

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., 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation,* Elsevier Science Publishers.

Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure appl. Geophys.,* **128**, 365–399.

Tarantola, A. & Valette, B., 1982. Inverse problem = quest for information, *J. Geophys.,* **50**, 159–170.

Tarrass, I., Giraud, L. & Thore, P., 2011. New curvilinear scheme for elastic wave propagation in presence of curved topography, *Geophys. Prospect.,* **59**(5), 889–906.

Teixeira, F.L. & Chew, W.C., 1997. Systematic derivation of anisotropic PML absorbing media in cylindrical and spherical coordinates, *IEEE Microw. Guid. Wave Lett.,* **7**(11), 371–373.

Teixeira, F.L. & Chew, W.C., 1998. General closed-form PML constitutive tensors to match arbitrary bianisotropic and dispersive linear media, *IEEE Microw. Guid. Wave Lett.,* **8**(6), 223–225.

Teixeira, F.L. & Chew, W.C., 1999. On causality and dynamic stability of perfectly matched layers for FDTD simulations, *IEEE Trans. Microw. Theory and Tech.,* **47**(6), 775–785.

Teng, Z.H., 2003. Exact boundary condition for time-dependent wave equation based on a boundary integral, *J. Comput. Phys.,* **190**(2), 398–418.

Ting, L. & Miksis, M.J., 1986. Exact boundary conditions for scattering problems, *J. acoust. Soc. Am.,* **80**(6), 1825–1827.

Toulorge, T. & Desmet, W., 2012. Optimal Runge-Kutta schemes for discontinuous Galerkin space discretizations applied to wave propagation problems, *J. Comput. Phys.,* **231**(4), 2067–2091.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160**(1), 195–216.

Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.,* **3**(1), 1–32.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.

Wang, S., Lee, R. & Teixeira, F.L., 2006. Anisotropic-medium PML for vector FETD with modified basis functions, *IEEE Trans. Antennas Propag.,* **54**(1), 20–27.

Williamson, J.H., 1980. Low-storage Runge-Kutta schemes, *J. Comput. Phys.,* **35**(1), 48–56.

Zhang, W. & Shen, Y., 2010. Unsplit complex frequency-shifted PML implementation using auxiliary differential equations for seismic wave modeling, *Geophysics,* **75**(4), T141–T154.

Zhu, H. & Tromp, J., 2013. Mapping tectonic deformation in the crust and upper mantle beneath Europe and the North Atlantic ocean, *Science,* **341**(6148), 871–875.

Zhu, H., Luo, Y., Nissen-Meyer, T., Morency, C. & Tromp, J., 2009. Elastic imaging and time-lapse migration based on adjoint methods, *Geophysics,* **74**, WCA167–WCA177.

Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nature Geosci.,* **5**(7), 493–498.

APPENDIX A: TIME-DOMAIN COMPLEX-FREQUENCY-SHIFTED UNSPLIT-FIELD PERFECTLY MATCHED LAYER (CFS-UPML) IN THE 3-D CASE

In vector components, eq. (8) can be written as

$$-\rho s_x s_y s_z \omega^2 u_x = \partial_x \left((\lambda + 2\mu) \frac{s_y s_z}{s_x} \partial_x u_x + \lambda s_z \partial_y u_y + \lambda s_y \partial_z u_z \right) + \partial_y \left(\mu s_z \partial_x u_y + \mu \frac{s_x s_z}{s_y} \partial_y u_x \right) + \partial_z \left(\mu s_y \partial_x u_z + \mu \frac{s_x s_y}{s_z} \partial_z u_x \right), \quad \text{(A1a)}$$

$$-\rho s_x s_y s_z \omega^2 u_y = \partial_x \left(\mu \frac{s_y s_z}{s_x} \partial_x u_y + \mu s_z \partial_y u_x \right) + \partial_y \left(\lambda s_z \partial_x u_x + (\lambda + 2\mu) \frac{s_x s_z}{s_y} \partial_y u_y + \lambda s_x \partial_z u_z \right) + \partial_z \left(\mu s_x \partial_y u_z + \mu \frac{s_x s_y}{s_z} \partial_z u_y \right), \quad \text{(A1b)}$$

$$-\rho s_x s_y s_z \omega^2 u_z = \partial_x \left(\mu \frac{s_y s_z}{s_x} \partial_x u_z + \mu s_y \partial_z u_x \right) + \partial_y \left(\mu \frac{s_x s_z}{s_y} \partial_y u_z + \mu s_x \partial_z u_y \right) + \partial_z \left(\lambda s_y \partial_x u_x + \lambda s_x \partial_y u_y + (\lambda + 2\mu) \frac{s_x s_y}{s_z} \partial_z u_z \right), \quad \text{(A1c)}$$

where

$$s_{x_i} = \kappa_{x_i} \frac{\beta_{x_i} + i\omega}{\alpha_{x_i} + i\omega} \tag{A2}$$

for $i = 1, 2, 3$. When the notation x_i is used, we use x_1 to represent x, x_2 to represent y and x_3 to represent z.

A1 Convolution formulation

With the help of symbolic calculation software such as Mathematica or Maple, the inverse Fourier transform of eq. (A1), that is, the convolution formulation of CFS-UPML, can be computed analytically

$$\rho L(t) * u_x = \partial_x \left[(\lambda + 2\mu)L_{231}(t) * \partial_x u_x + \lambda L_3(t) * \partial_y u_y + \lambda L_2(t) * \partial_z u_z \right] + \partial_y \left[\mu L_3(t) * \partial_x u_y + \mu L_{132}(t) * \partial_y u_x \right]$$
$$+ \partial_z \left[\mu L_2(t) * \partial_x u_z + \mu L_{123}(t) * \partial_z u_x \right], \tag{A3a}$$

$$\rho L(t) * u_y = \partial_x \left[\mu L_{231}(t) * \partial_x u_y + \mu L_3(t) * \partial_y u_x \right] + \partial_y \left[\lambda L_3(t) * \partial_x u_x + (\lambda + 2\mu)L_{132}(t) * \partial_y u_y + \lambda L_1(t) * \partial_z u_z \right]$$
$$+ \partial_z \left[\mu L_1(t) * \partial_y u_z + \mu L_{123}(t) * \partial_z u_y \right], \tag{A3b}$$

$$\rho L(t) * u_z = \partial_x \left[\mu L_{231}(t) * \partial_x u_z + \mu L_2(t) * \partial_z u_x \right] + \partial_y \left[\mu L_{132}(t) * \partial_y u_z + \mu L_1(t) * \partial_z u_y \right]$$
$$+ \partial_z \left[\lambda L_2(t) * \partial_x u_x + \lambda L_1(t) * \partial_y u_y + (\lambda + 2\mu)L_{123}(t) * \partial_z u_z \right], \tag{A3c}$$

where

$$L(t) = F^{-1} \left[-\omega^2 s_{x_1} s_{x_2} s_{x_3} \right], \quad L_i(t) = F^{-1} \left[s_{x_i} \right], \quad L_{ijk}(t) = F^{-1} \left[\frac{s_{x_i} s_{x_j}}{s_{x_k}} \right]. \tag{A4}$$

For $L(t)$, we have

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_{x_1} t} H(t) + \bar{a}_4 e^{-\alpha_{x_2} t} H(t) + \bar{a}_5 e^{-\alpha_{x_3} t} H(t), \tag{A5a}$$

$$\bar{a}_0 = \kappa_{x_1} \kappa_{x_2} \kappa_{x_3}, \tag{A5b}$$

$$\bar{a}_1 = \bar{a}_0 (\beta_{x_1} - \alpha_{x_1}) + \bar{a}_0 (\beta_{x_2} - \alpha_{x_2}) + \bar{a}_0 (\beta_{x_3} - \alpha_{x_3}), \tag{A5c}$$

$$\bar{a}_2 = \bar{a}_0 \left(\beta_{x_1} - \alpha_{x_1} \right) \left(\beta_{x_2} - \alpha_{x_2} - \alpha_{x_1} \right) + \bar{a}_0 \left(\beta_{x_2} - \alpha_{x_2} \right) \left(\beta_{x_3} - \alpha_{x_3} - \alpha_{x_2} \right) + \bar{a}_0 \left(\beta_{x_3} - \alpha_{x_3} \right) \left(\beta_{x_1} - \alpha_{x_1} - \alpha_{x_3} \right), \tag{A5d}$$

$$\bar{a}_3 = \bar{a}_{123}, a_4 = \bar{a}_{213}, a_5 = \bar{a}_{312}, \tag{A5e}$$

where

$$\bar{a}_{x_i x_j x_k} = \kappa_{x_i} \kappa_{x_j} \kappa_{x_k} \alpha_{x_i}^2 \frac{(\beta_{x_i} - \alpha_{x_i})(\beta_{x_j} - \alpha_{x_i})(\beta_{x_k} - \alpha_{x_i})}{(\alpha_{x_j} - \alpha_{x_i})(\alpha_{x_k} - \alpha_{x_i})}. \tag{A6}$$

The above equations for $\bar{a}_3$, $\bar{a}_4$ and $\bar{a}_4$ are only valid for $\alpha_{x_1} \neq \alpha_{x_2} \neq \alpha_{x_3}$ inside the CFS-UPML region. When it is not the case, $\bar{a}_{x_i x_j x_k}$ may hold singularities up to first order when (1) $\alpha_{x_1} = \alpha_{x_2}$, (2) $\alpha_{x_1} = \alpha_{x_3}$, (3) $\alpha_{x_2} = \alpha_{x_3}$ or up to second order when (4) $\alpha_{x_1} = \alpha_{x_2} = \alpha_{x_3}$. In these cases, we have:
(1) $\alpha_{x_1} = \alpha_{x_2} = \alpha_0$

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_0 t} H(t) + \bar{a}_4 e^{-\alpha_0 t} t H(t) + \bar{a}_5 e^{-\alpha_{x_3} t} H(t), \tag{A7a}$$

$$\bar{a}_3 = \bar{a}_{123} + \bar{a}_{213}, \quad \bar{a}_4 = \bar{a}_{213}(\alpha_{x_1} - \alpha_{x_2}), \quad \bar{a}_5 = \bar{a}_{312}, \tag{A7b}$$

where $\bar{a}_0$, $\bar{a}_1$ and $\bar{a}_2$ are given by (A5b)–(A5d) with $\alpha_{x_1} = \alpha_{x_2} = \alpha_0$. It is worth mentioning that although a_{123} and a_{213} are singular, their sum is non-singular because after summation the singular term $(\alpha_{x_1} - \alpha_{x_2})$ in the denominator is cancelled by the exact same term in the numerator, and thus in the final expression of $\bar{a}_3$ there are no singularities. This is true for all the other expressions that involve the summation of singular coefficients.
(2) $\alpha_{x_1} = \alpha_{x_3} = \alpha_0$

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + a_2 \delta(t) + \bar{a}_3 e^{-\alpha_0 t} H(t) + \bar{a}_4 e^{-\alpha_{x_2} t} H(t) + \bar{a}_5 e^{-\alpha_0 t} t H(t), \tag{A8a}$$

$$\bar{a}_3 = \bar{a}_{123} + \bar{a}_{312}, \quad \bar{a}_4 = \bar{a}_{213}, \quad \bar{a}_5 = \bar{a}_{312}(\alpha_{x_1} - \alpha_{x_3}), \tag{A8b}$$

where $\bar{a}_0$, $\bar{a}_1$ and $\bar{a}_2$ are given by (A5b)–(A5d) with $\alpha_{x_1} = \alpha_{x_3} = \alpha_0$.

(3) $\alpha_{x_2} = \alpha_{x_3} = \alpha_0$

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_{x_1} t} H(t) + \bar{a}_4 e^{-\alpha_0 t} H(t) + \bar{a}_5 e^{-\alpha_0 t} t H(t),$$ (A9a)

$$\bar{a}_3 = \bar{a}_{123}, \quad \bar{a}_4 = \bar{a}_{213} + \bar{a}_{312}, \quad \bar{a}_5 = \bar{a}_{312}(\alpha_{x_2} - \alpha_{x_3}),$$ (A9b)

where $\bar{a}_0$, $\bar{a}_1$ and $\bar{a}_2$ are given by (A5b)–(A5d) with $\alpha_{x_2} = \alpha_{x_3} = \alpha_0$.

(4) $\alpha_{x_i} = \alpha_{x_j} = \alpha_{x_k} = \alpha_0$

$$L(t) = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_0 t} H(t) + \bar{a}_4 e^{-\alpha_0 t} t H(t) + \bar{a}_5 e^{-\alpha_0 t} t^2 H(t),$$ (A10a)

$$\bar{a}_5 = \frac{1}{2} \bar{a}_{x_i x_j x_k}(\alpha_{x_j} - \alpha_{x_i})(\alpha_{x_k} - \alpha_{x_i}), \quad \bar{a}_4 = -2 \frac{\partial \bar{a}_5}{\partial \alpha_0}, \quad \bar{a}_3 = -\frac{1}{2} \frac{\partial \bar{a}_4}{\partial \alpha_0},$$ (A10b)

where a_0, a_1 and a_2 are given by (A5b)–(A5d) with $\alpha_{x_1} = \alpha_{x_2} = \alpha_{x_3} = \alpha_0$.

Applying the $L(t)$ operator to the displacement field u_{x_i} will result in convolution terms like

$$\left[e^{-\alpha_0 t} t H(t) \right] * u_{x_i}, \quad \left[e^{-\alpha_0 t} t^2 H(t) \right] * u_{x_i}.$$ (A11)

Such terms are difficult to implement because they cannot be efficiently computed based on a recursive convolution technique such as that of Luebbers & Hunsberger (1992), thus one would need to store all the past states (past time steps) of the variables involved in their computation. Such a difficulty can be avoided by using the linear property of the convolution operation to decompose these convolution terms as

$$\left[e^{-\alpha_0 t} t H(t) \right] * u_{x_i} = t \left[e^{-\alpha_0 t} H(t) \right] * u_{x_i} + \left[e^{-\alpha_0 t} H(t) \right] * \left[t u_{x_i} \right]$$ (A12)

and

$$\left[e^{-\alpha_0 t} t^2 H(t) \right] * u_{x_i} = t^2 \left[e^{-\alpha_0 t} H(t) \right] * u_{x_i} - 2t \left[e^{-\alpha_0 t} H(t) \right] * \left[t u_{x_i} \right] + \left[e^{-\alpha_0 t} H(t) \right] * \left[t^2 u_{x_i} \right].$$ (A13)

Based on (A12) and (A13), the application of the $L(t)$ operator to the displacement field thus becomes

(1) $\alpha_{x_1} = \alpha_{x_2} = \alpha_0$

$$L(t) * u_{x_i} = a_0 \ddot{u}_{x_i} + a_1 \dot{u}_{x_i} + a_2 u_{x_i} + a_3 \left[e^{-\alpha_0 t} H(t) \right] * u_{x_i} + a_4 \left[e^{-\alpha_0 t} H(t) \right] * \left[t u_{x_i} \right] + a_5 \left[e^{-\alpha_{x_3} t} H(t) \right] * u_{x_i},$$ (A14a)

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3 + t \bar{a}_4, \quad a_4 = -\bar{a}_4, \quad a_5 = \bar{a}_5.$$ (A14b)

(2) $\alpha_{x_1} = \alpha_{x_3} = \alpha_0$

$$L(t) * u_{x_i} = a_0 \ddot{u}_{x_i} + a_1 \dot{u}_{x_i} + a_2 u_{x_i} + a_3 \left[e^{-\alpha_0 t} H(t) \right] * u_{x_i} + a_4 \left[e^{-\alpha_{x_2} t} H(t) \right] * u_{x_i} + a_5 \left[e^{-\alpha_0 t} H(t) \right] * \left[t u_{x_i} \right],$$ (A15a)

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3 + t \bar{a}_5, \quad a_4 = \bar{a}_4, \quad a_5 = -\bar{a}_5.$$ (A15b)

(3) $\alpha_{x_2} = \alpha_{x_3} = \alpha_0$

$$L(t) * u_{x_i} = a_0 \ddot{u}_{x_i} + a_1 \dot{u}_{x_i} + a_2 u_{x_i} + a_3 \left[e^{-\alpha_{x_1} t} H(t) \right] * u_{x_i} + a_4 \left[e^{-\alpha_0 t} H(t) \right] * u_{x_i} + a_5 \left[e^{-\alpha_0 t} H(t) \right] * \left[t u_{x_i} \right],$$ (A16a)

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3, \quad a_4 = \bar{a}_4 + t \bar{a}_5, \quad a_5 = -\bar{a}_5.$$ (A16b)

(4) $\alpha_{x_i} = \alpha_{x_j} = \alpha_{x_k} = \alpha_0$

$$L(t) * u_{x_i} = a_0 \ddot{u}_{x_i} + a_1 \dot{u}_{x_i} + a_2 u_{x_i} + a_3 \left[e^{-\alpha_0 t} H(t) \right] * u_{x_i} + a_4 \left[e^{-\alpha_0 t} H(t) \right] * \left[t u_{x_i} \right] + a_5 \left[e^{-\alpha_0 t} H(t) \right] * \left[t^2 u_{x_i} \right],$$ (A17a)

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3 + t \bar{a}_4 + t^2 \bar{a}_5, \quad a_4 = -\bar{a}_4 - 2t \bar{a}_5, \quad a_5 = \bar{a}_5.$$ (A17b)

For the PML-stretched stress tensor the time-domain counterpart of s_{x_i} and $s_{x_i} s_{x_j} / s_{x_k}$ for $i, j, k = 1, 2, 3$ (these indices should not be mistaken for those used in the constitutive relation) is required, that is, $L_i(t)$ and $L_{ijk}(t)$. For $L_i(t)$, we have

$$L_i(t) = \kappa_{x_i} \delta(t) + d_{x_i} e^{-\alpha_{x_i} t} H(t)$$ (A18)

and the application of $L_i(t)$ operator to the strain field leads to

$$L_i(t) = \kappa_{x_i} \partial_j u_k + d_{x_i} \left[e^{-\alpha_{x_i} t} H(t) \right] * \partial_j u_k, \quad \text{with } j \neq i \text{ and } k \neq i \text{ but } j \text{ can be equal to } k.$$ (A19)

For $L_{ijk}(t)$, we have

$$L_{ijk}(t) = \bar{b}_0^{ijk} \delta(t) + \bar{b}_1^{ijk} e^{-\alpha_{x_i} t} H(t) + \bar{b}_2^{ijk} e^{-\alpha_{x_j} t} H(t) + \bar{b}_3^{ijk} e^{-\beta_{x_k} t} H(t),$$ (A20a)

$$\bar{b}_0^{ijk} = \gamma_0^{ijk} = \frac{\kappa_{x_i} \kappa_{x_j}}{\kappa_k},$$ (A20b)

$$\bar{b}_1^{ijk} = \gamma_1^{ijk} = -b_0^{ijk} \frac{(\alpha_{x_i} - \alpha_{x_k})(\alpha_{x_i} - \beta_{x_i})(\alpha_{x_i} - \beta_{x_j})}{(\alpha_{x_i} - \alpha_{x_j})(\alpha_{x_i} - \beta_{x_k})}, \tag{A20c}$$

$$\bar{b}_2^{ijk} = \gamma_2^{ijk} = -b_0^{ijk} \frac{(\alpha_{x_j} - \alpha_{x_k})(\alpha_{x_j} - \beta_{x_i})(\alpha_{x_j} - \beta_{x_j})}{(\alpha_{x_j} - \alpha_{x_i})(\alpha_{x_j} - \beta_{x_k})}, \tag{A20d}$$

$$\bar{b}_3^{ijk} = \gamma_3^{ijk} = -b_0^{ijk} \frac{(\beta_{x_k} - \alpha_{x_k})(\beta_{x_k} - \beta_{x_i})(\beta_{x_k} - \beta_{x_j})}{(\beta_{x_k} - \alpha_{x_i})(\beta_{x_k} - \alpha_{x_j})}. \tag{A20e}$$

According to equations (A20c)–(A20e), first-order singularities will arise in case (1) $\alpha_{x_i} = \alpha_{x_j}$, (2) $\alpha_{x_i} = \beta_{x_k}$ or (3) $\alpha_{x_j} = \beta_{x_k}$; and second-order singularities will arise in case (4) $\alpha_{x_i} = \alpha_{x_j} = \beta_{x_k}$. In these cases, we have

(1) $\alpha_{x_i} = \alpha_{x_j} = \alpha_0$

$$L_{ijk}(t) = \bar{b}_0^{ijk} \delta(t) + \bar{b}_1^{ijk} e^{-\alpha_0 t} H(t) + \bar{b}_2^{ijk} e^{-\alpha_0 t} t H(t) + \bar{b}_3^{ijk} e^{-\beta_{x_k} t} H(t), \tag{A21a}$$

$$\bar{b}_0^{ijk} = \gamma_0^{ijk}, \; \bar{b}_1^{ijk} = \gamma_1^{ijk} + \gamma_2^{ijk}, \; \bar{b}_2^{ijk} = \gamma_2^{ijk}(\alpha_{x_i} - \alpha_{x_j}), \; \bar{b}_3^{ijk} = \gamma_3^{ijk}. \tag{A21b}$$

(2) $\alpha_{x_i} = \beta_{x_k} = \alpha_0$

$$L_{ijk}(t) = \bar{b}_0^{ijk} \delta(t) + \bar{b}_1^{ijk} e^{-\alpha_0 t} H(t) + \bar{b}_2^{ijk} e^{-\alpha_{x_j} t} H(t) + \bar{b}_3^{ijk} e^{-\alpha_0 t} t H(t), \tag{A22a}$$

$$\bar{b}_0^{ijk} = \gamma_0^{ijk}, \; \bar{b}_1^{ijk} = \gamma_1^{ijk} + \gamma_3^{ijk}, \; \bar{b}_2^{ijk} = \gamma_2^{ijk}, \; \bar{b}_3^{ijk} = \gamma_3^{ijk}(\alpha_{x_i} - \beta_{x_k}). \tag{A22b}$$

(3) $\alpha_{x_j} = \beta_{x_k} = \alpha_0$

$$L_{ijk}(t) = \bar{b}_0^{ijk} \delta(t) + \bar{b}_1^{ijk} e^{-\alpha_{x_i} t} H(t) + \bar{b}_2^{ijk} e^{-\alpha_0 t} H(t) + \bar{b}_3^{ijk} e^{-\alpha_0 t} t H(t), \tag{A23a}$$

$$\bar{b}_0^{ijk} = \gamma_0^{ijk}, \; \bar{b}_1^{ijk} = \gamma_1^{ijk}, \; \bar{b}_2^{ijk} = \gamma_2^{ijk} + \gamma_3^{ijk}, \; \bar{b}_3^{ijk} = \gamma_3^{ijk}(\alpha_{x_j} - \beta_{x_k}). \tag{A23b}$$

(4) $\alpha_{x_i} = \alpha_{x_j} = \beta_{x_k} = \alpha_0$

$$L_{ijk}(t) = \bar{b}_0^{ijk} \delta(t) + \bar{b}_1^{ijk} e^{-\alpha_0 t} H(t) + \bar{b}_2^{ijk} e^{-\alpha_0 t} t H(t) + \bar{b}_3^{ijk} e^{-\alpha_0 t} t^2 H(t), \tag{A24a}$$

$$\bar{b}_3^{ijk} = \frac{1}{2} \gamma_3^{ijk}(\beta_{x_k} - \alpha_{x_i})(\beta_{x_k} - \alpha_{x_j}), \; \bar{b}_2^{ijk} = -2 \frac{\partial \bar{b}_3^{ijk}}{\partial \alpha_0}, \; \bar{b}_1^{ijk} = -\frac{1}{2} \frac{\partial \bar{b}_2^{ijk}}{\partial \alpha_0}. \tag{A24b}$$

Similarly, based on (A12) and (A13), the application of the $L_{ijk}(t)$ operator to the strain field results in

(1) $\alpha_{x_i} = \alpha_{x_j} = \alpha_0$

$$L_{ijk}(t) * \partial_k u_l = b_0^{ijk} \partial_k u_l + b_1^{ijk} \left[e^{-\alpha_0 t} H(t) \right] \partial_k u_l + b_2^{ijk} \left[e^{-\alpha_0 t} H(t) \right] [t \, \partial_k u_l] + b_3^{ijk} \left[e^{-\alpha_{x_k} t} H(t) \right] \partial_k u_l, \tag{A25a}$$

$$b_0^{ijk} = \bar{b}_0^{ijk}, \; b_1^{ijk} = \bar{b}_1^{ijk} + t \, \bar{b}_2^{ijk}, \; b_2^{ijk} = -\bar{b}_2^{ijk}, \; b_3^{ijk} = \bar{b}_3^{ijk}. \tag{A25b}$$

(2) $\alpha_{x_i} = \beta_{x_k} = \alpha_0$

$$L_{ijk}(t) * \partial_k u_l = b_0^{ijk} \partial_k u_l + b_1^{ijk} \left[e^{-\alpha_0 t} H(t) \right] \partial_k u_l + b_2^{ijk} \left[e^{-\alpha_{x_j} t} H(t) \right] \partial_k u_l + b_3^{ijk} \left[e^{-\alpha_0 t} H(t) \right] [t \, \partial_k u_l], \tag{A26a}$$

$$b_0^{ijk} = \bar{b}_0^{ijk}, \; b_1^{ijk} = \bar{b}_1^{ijk} + t \, \bar{b}_3^{ijk}, \; b_2^{ijk} = \bar{b}_2^{ijk}, \; b_3^{ijk} = -\bar{b}_3^{ijk}. \tag{A26b}$$

(3) $\alpha_{x_j} = \beta_{x_k} = \alpha_0$

$$L_{ijk}(t) * \partial_k u_l = b_0^{ijk} \partial_k u_l + b_1^{ijk} \left[e^{-\alpha_{x_i} t} H(t) \right] \partial_k u_l + b_2^{ijk} \left[e^{-\alpha_0 t} H(t) \right] \partial_k u_l + b_3^{ijk} \left[e^{-\alpha_0 t} H(t) \right] [t \, \partial_k u_l], \tag{A27a}$$

$$b_0^{ijk} = \bar{b}_0^{ijk}, \; b_1^{ijk} = \bar{b}_1^{ijk}, \; b_2^{ijk} = \bar{b}_2^{ijk} + t \, \bar{b}_3^{ijk}, \; b_3^{ijk} = -\bar{b}_3^{ijk}. \tag{A27b}$$

(4) $\alpha_{x_i} = \alpha_{x_j} = \beta_{x_k} = \alpha_0$

$$L_{ijk}(t) * \partial_k u_l = b_0^{ijk} \partial_k u_l + b_1^{ijk} \left[e^{-\alpha_{x_i} t} H(t) \right] \partial_k u_l + b_2^{ijk} \left[e^{-\alpha_0 t} H(t) \right] [t \, \partial_k u_l] + b_3^{ijk} \left[e^{-\alpha_0 t} H(t) \right] [t^2 \, \partial_k u_l], \tag{A28a}$$

$$b_0^{ijk} = \bar{b}_0^{ijk}, \; b_1^{ijk} = \bar{b}_1^{ijk} + t \, \bar{b}_2^{ijk} + t^2 \, \bar{b}_3^{ijk}, \; b_2^{ijk} = -\bar{b}_2^{ijk} - 2t \, \bar{b}_3^{ijk}, \; b_3^{ijk} = \bar{b}_3^{ijk}. \tag{A28b}$$

A2 Auxiliary differential equation (ADE) formulation

The goal of reformulating the convolution form of CFS-UPML into an ADE formulation is to facilitate the use of higher order time schemes and to be able to use the same time and space numerical schemes in the CFS-UPMLs as in the main domain. The main idea is to avoid having to handle convolution terms and introducing ADEs for these terms instead.

To do so, let us first derive the ADEs for convolution terms in $L(t) * u_{x_i}(t)$. According to (A14)–(A17), in all cases, three kinds of ADEs need to be derived; using $R_L^{u_{x_i},1}(t)$, $R_L^{u_{x_i},2}$ and $R_L^{u_{x_i},3}(t)$ to represent the first, second and third convolution terms, respectively, we have

(1) $\alpha_{x_1} \neq \alpha_{x_2} \neq \alpha_{x_3}$

$$\frac{dR_L^{u_{x_i},1}}{dt} = -\alpha_{x_1} R_L^{u_{x_i},1} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},2}}{dt} = -\alpha_{x_2} R_L^{u_{x_i},2} + tu_{x_i}, \quad \frac{dR_L^{u_{x_i},3}}{dt} = -\alpha_{x_3} R_L^{u_{x_i},3} + u_{x_i}. \tag{A29}$$

(2) $\alpha_{x_1} = \alpha_{x_2} = \alpha_0$

$$\frac{dR_L^{u_{x_i},1}}{dt} = -\alpha_0 R_L^{u_{x_i},1} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},2}}{dt} = -\alpha_0 R_L^{u_{x_i},2} + tu_{x_i}, \quad \frac{dR_L^{u_{x_i},3}}{dt} = -\alpha_{x_3} R_L^{u_{x_i},3} + u_{x_i}. \tag{A30}$$

(3) $\alpha_{x_1} = \alpha_{x_3} = \alpha_0$

$$\frac{dR_L^{u_{x_i},1}}{dt} = -\alpha_0 R_L^{u_{x_i},1} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},2}}{dt} = -\alpha_{x_2} R_L^{u_{x_i},2} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},3}}{dt} = -\alpha_0 R_L^{u_{x_i},3} + tu_{x_i}. \tag{A31}$$

(4) $\alpha_{x_2} = \alpha_{x_3} = \alpha_0$

$$\frac{dR_L^{u_{x_i},1}}{dt} = -\alpha_{x_1} R_L^{u_{x_i},1} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},2}}{dt} = -\alpha_0 R_L^{u_{x_i},2} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},3}}{dt} = -\alpha_0 R_L^{u_{x_i},3} + tu_{x_i} \tag{A32}$$

5) $\alpha_{x_i} = \alpha_{x_j} = \alpha_{x_k} = \alpha_0$

$$\frac{dR_L^{u_{x_i},1}}{dt} = -\alpha_0 R_L^{u_{x_i},1} + u_{x_i}, \quad \frac{dR_L^{u_{x_i},2}}{dt} = -\alpha_0 R_L^{u_{x_i},2} + tu_{x_i}, \quad \frac{dR_L^{u_{x_i},3}}{dt} = -\alpha_0 R_L^{u_{x_i},3} + t^2 u_{x_i}. \tag{A33}$$

In a similar way, we can derive the corresponding ADEs governing convolution terms when applying $L_i(t)$ to $\partial_j u_k$ and $L_{ijk}(t)$ to $\partial_k u_l$.

APPENDIX B: TIME-DOMAIN CFS-UPML IN THE 2-D P-SV PLANE STRAIN CASE

In the 2-D plane-strain case, in component form the elastic wave equation can be written as

$$\rho \ddot{u}_x = \partial_x[(\lambda + 2\mu)\, \partial_x u_x + \lambda\, \partial_y u_y] + \partial_y[\mu\, (\partial_x u_y + \partial_y u_x)], \tag{B1a}$$

$$\rho \ddot{u}_y = \partial_x[\mu\, (\partial_x u_y + \partial_y u_x)] + \partial_y[\lambda\, \partial_x u_x + (\lambda + 2\mu)\, \partial_y u_y]. \tag{B1b}$$

B1 Convolution formulation

Following the same procedure as that used to derive the CFS-UPML in 3-D in Section 2.2, we first transform the above equations into the frequency domain:

$$-\rho\omega^2\hat{u}_x = \partial_x[(\lambda + 2\mu)\, \partial_x \hat{u}_x + \lambda\, \partial_y \hat{u}_y] + \partial_y[\mu\, (\partial_x \hat{u}_y + \partial_y \hat{u}_x)], \tag{B2a}$$

$$-\rho\omega^2\hat{u}_y = \partial_x[\mu\, (\partial_x \hat{u}_y + \partial_y \hat{u}_x)] + \partial_y[\lambda\, \partial_x \hat{u}_x + (\lambda + 2\mu)\, \partial_y \hat{u}_y], \tag{B2b}$$

where '^' over a symbol represents its Fourier transform. In the PML, the initial displacement and velocity vectors as well as the source term are assumed to be zero, that is, the medium is initially at rest. We then introduce the new complex coordinate:

$$\tilde{x}(x) = \int_0^x s_x(x')\, dx'\, ; \; \tilde{y}(y) = \int_0^y s_y(y')\, dy'. \tag{B3}$$

Using the chain rule, we then have

$$\partial_{\tilde{x}} = \frac{1}{s_x}\, \partial_x, \; \partial_{\tilde{y}} = \frac{1}{s_y}\, \partial_y, \tag{B4}$$

and we obtain the CFS-UPML formulation in the frequency domain:

$$-\rho\omega^2 s_x s_y \hat{u}_x = \partial_x \left((\lambda + 2\mu)\frac{s_y}{s_x}\, \partial_x \hat{u}_x + \lambda\, \partial_y \hat{u}_y\right) + \partial_y \left(\mu\, \partial_x \hat{u}_y + \mu\frac{s_x}{s_y}\, \partial_y \hat{u}_x\right), \tag{B5a}$$

$$-\rho\omega^2 s_x s_y \hat{u}_y = \partial_x \left(\mu\frac{s_y}{s_x}\, \partial_x \hat{u}_y + \mu\, \partial_y \hat{u}_x\right) + \partial_y \left(\lambda\, \partial_x \hat{u}_x + (\lambda + 2\mu)\frac{s_x}{s_y}\, \partial_y \hat{u}_y\right), \tag{B5b}$$

in which we have already multiplied both sides of the equations by $s_x s_y$ in order to be able to obtain a formulation in the time domain based on the displacement vector only. This enables us to write the 2-D convolution formulation of CFS-UPML as:

$$\rho F^{-1}[-\omega^2 s_x s_y] * u_x = \partial_x \left((\lambda + 2\mu)F^{-1}\left[\frac{s_y}{s_x}\right] * \partial_x u_x + \lambda\, \partial_y u_y\right) + \partial_y \left(\mu\, \partial_x u_y + \mu\, F^{-1}\left[\frac{s_x}{s_y}\right] * \partial_y u_x\right) \tag{B6a}$$

$$\rho F^{-1}[-\omega^2 s_x s_y] * u_y = \partial_x \left(\mu \, F^{-1} \left[\frac{s_y}{s_x} \right] * \partial_x u_y + \mu \, \partial_y u_x \right) + \partial_y \left(\lambda \, \partial_x u_x + (\lambda + 2\mu) \, F^{-1} \left[\frac{s_x}{s_y} \right] * \partial_y u_y \right), \tag{B6b}$$

where we define s_x and s_y as CFS stretching functions:

$$s_x = \kappa_x + \frac{d_x}{\alpha_x + \mathbf{i}\omega}, \quad s_y = \kappa_y + \frac{d_y}{\alpha_y + \mathbf{i}\omega}. \tag{B7}$$

As in Section 2.2, we reformulate s_x and s_y as

$$s_x = \kappa_x \frac{\beta_x + \mathbf{i}\omega}{\alpha_x + \mathbf{i}\omega}, \quad s_y = \kappa_y \frac{\beta_y + \mathbf{i}\omega}{\alpha_y + \mathbf{i}\omega}, \tag{B8}$$

where $\beta_x = \alpha_x + d_x/\kappa_x$, $\beta_y = \alpha_y + d_y/\kappa_y$. With the help of symbolic calculation software such as Mathematica or Maple we obtain the analytical expression of the inverse Fourier transform of $-\omega^2 s_x s_y$ as:

(1) $\alpha_x \neq \alpha_y$

$$F^{-1}[-\omega^2 s_x s_y] = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_x t} H(t) + \bar{a}_4 e^{-\alpha_y t} H(t), \tag{B9a}$$

$$\bar{a}_0 = \kappa_x \kappa_y, \quad \bar{a}_1 = \bar{a}_0[(\beta_x - \alpha_x) + (\beta_y - \alpha_y)], \tag{B9b}$$

$$\bar{a}_2 = \bar{a}_0[(\beta_x - \alpha_x)(\beta_y - \alpha_y) - \alpha_x(\beta_x - \alpha_x) - \alpha_y(\beta_y - \alpha_y)], \tag{B9c}$$

$$\bar{a}_3 = \bar{a}_0 \alpha_x^2 \frac{(\alpha_x - \beta_x)(\alpha_x - \beta_y)}{\alpha_y - \alpha_x}, \quad \bar{a}_4 = \bar{a}_0 \alpha_y^2 \frac{(\alpha_y - \beta_x)(\alpha_y - \beta_y)}{\alpha_x - \alpha_y}. \tag{B9d}$$

Thus, applying $F^{-1}[-\omega^2 s_x s_y]$ to u_x leads to

$$L(t) * u_x = a_0 \ddot{u}_x + a_1 \dot{u}_x + a_2 u_x + a_3 \left[e^{-\alpha_x t} H(t) \right] * u_x + a_4 \left[e^{-\alpha_y t} H(t) \right] * u_x, \tag{B10a}$$

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3, \quad a_4 = \bar{a}_4. \tag{B10b}$$

(2) $\alpha_x = \alpha_y = \alpha_0$

$$F^{-1}[-\omega^2 s_x s_y] = \bar{a}_0 \ddot{\delta}(t) + \bar{a}_1 \dot{\delta}(t) + \bar{a}_2 \delta(t) + \bar{a}_3 e^{-\alpha_0 t} H(t) + \bar{a}_4 e^{-\alpha_0 t} t H(t), \tag{B11a}$$

$$\bar{a}_0 = \kappa_x \kappa_y, \quad \bar{a}_1 = \bar{a}_0(\beta_x + \beta_y - 2\alpha_0), \tag{B11b}$$

$$\bar{a}_2 = \bar{a}_0[(\beta_x - \alpha_0)(\beta_y - \alpha_0) - \alpha_0(\beta_x - \alpha_0) - \alpha_0(\beta_y - \alpha_0)], \tag{B11c}$$

$$\bar{a}_3 = \bar{a}_0 \left[3\alpha_0^2(\beta_x + \beta_x) - 2\alpha_0 \beta_x \beta_y - 4\alpha_0^3 \right], \quad \bar{a}_4 = \bar{a}_0 \left[\alpha_0^2 (\beta_x - \alpha_0)(\beta_y - \alpha_0) \right]. \tag{B11d}$$

Thus, using the linear property of the convolution operation, applying $F^{-1}[-\omega^2 s_x s_y]$ to u_x results in

$$L(t) * u_x = a_0 \ddot{u}_x + a_1 \dot{u}_x + a_2 u_x + a_3 \left[e^{-\alpha_0 t} H(t) \right] * u_x + a_4 \left[e^{-\alpha_0 t} H(t) \right] * [t u_x], \tag{B12a}$$

$$a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1, \quad a_2 = \bar{a}_2, \quad a_3 = \bar{a}_3 + t\bar{a}_4, \quad a_4 = -\bar{a}_4 \tag{B12b}$$

and similarly we can express the result of applying $F^{-1}[-\omega^2 s_x s_y]$ to u_y.

Regarding the s_x/s_y term, its inverse Fourier transform $F^{-1}[s_x/s_y]$ can be written as

(1) $\alpha_x \neq \beta_y$

$$F^{-1}\left[\frac{s_x}{s_y} \right] = \bar{b}_0 \delta(t) + \bar{b}_1 e^{-\alpha_x t} H(t) + \bar{b}_2 e^{-\beta_y t} H(t), \tag{B13a}$$

$$\bar{b}_0 = \frac{\kappa_x}{\kappa_y}, \quad \bar{b}_1 = -\bar{b}_0 \frac{(\alpha_x - \alpha_y)(\alpha_x - \beta_x)}{\alpha_x - \beta_y}, \quad \bar{b}_2 = -\bar{b}_0 \frac{(\beta_y - \alpha_y)(\beta_y - \beta_x)}{\beta_y - \alpha_x}. \tag{B13b}$$

Applying $F^{-1}[\frac{s_x}{s_y}]$ to $\partial_x u_x$, we get:

$$F^{-1}\left[\frac{s_x}{s_y} \right] * \partial_x u_x = \bar{b}_0 \partial_x u_x + \bar{b}_1 \left[e^{-\alpha_x t} H(t) \right] * \partial_x u_x + \bar{b}_2 \left[e^{-\beta_y t} H(t) \right] * \partial_x u_x, \tag{B14a}$$

$$b_0 = \bar{b}_0, \quad b_1 = \bar{b}_1, \quad b_2 = \bar{b}_2. \tag{B14b}$$

(2) $\alpha_x = \beta_y = \alpha_0$

$$F^{-1}\left[\frac{s_x E}{s_y} \right] = \bar{b}_0 \delta(t) + \bar{b}_1 e^{-\alpha_0 t} H(t) + \bar{b}_2 e^{-\alpha_0 t} t H(t), \tag{B15a}$$

$$\bar{b}_0 = \kappa_x/\kappa_y, \ \bar{b}_1 = \bar{b}_0[(\alpha_y - \alpha_0) - (\alpha_0 - \beta_x)], \ \bar{b}_2 = \bar{b}_0(\alpha_0 - \alpha_y)(\alpha_0 - \beta_x). \tag{B15b}$$

Similarly, in this case applying $F^{-1}[\frac{s_x}{s_y}]$ to $\partial_x u_x$ results in

$$F^{-1}\left[\frac{s_x}{s_y}\right] * \partial_x u_x = \bar{b}_0 \partial_x u_x + \bar{b}_1 \left[e^{-\alpha_0 t} H(t)\right] * \partial_x u_x + \bar{b}_2 \left[e^{-\alpha_0 t} H(t)\right] * [t \partial_x u_x], \tag{B16a}$$

$$b_0 = \bar{b}_0, \ b_1 = \bar{b}_1 + t\bar{b}_2, \ b_2 = -\bar{b}_2. \tag{B16b}$$

Using the same procedure, we can obtain $F^{-1}[\frac{s_x}{s_y}] * \partial_x u_y$.

Similarly, the inverse Fourier transform of s_y/s_x is

(1) $\alpha_y \neq \beta_x$

$$F^{-1}\left[\frac{s_y}{s_x}\right] = \bar{b}_0 \delta(t) + \bar{b}_1 e^{-\alpha_y t} H(t) + \bar{b}_1 e^{-\beta_x t} H(t), \tag{B17a}$$

$$\bar{b}_0 = \frac{\kappa_y}{\kappa_x}, \ \bar{b}_1 = -\bar{b}_0 \frac{(\alpha_y - \alpha_x)(\alpha_y - \beta_y)}{\alpha_y - \beta_x}, \ \bar{b}_2 = -\bar{b}_0 \frac{(\beta_x - \alpha_x)(\beta_x - \beta_y)}{\beta_x - \alpha_y}. \tag{B17b}$$

Applying $F^{-1}[\frac{s_y}{s_x}]$ to $\partial_y u_x$, we get

$$F^{-1}\left[\frac{s_y}{s_x}\right] * \partial_y u_x = \bar{b}_0 \partial_y u_x + \bar{b}_1 \left[e^{-\alpha_y t} H(t)\right] * \partial_y u_x + \bar{b}_1 \left[e^{-\beta_x t} H(t)\right] * \partial_y u_x, \tag{B18a}$$

$$b_0 = \bar{b}_0, \ b_1 = \bar{b}_1, \ b_2 = \bar{b}_2. \tag{B18b}$$

(2) $\alpha_y = \beta_x = \alpha_0$

$$F^{-1}\left[\frac{s_y}{s_x}\right] = \bar{b}_0 \delta(t) + \bar{b}_1 e^{-\alpha_0 t} H(t) + \bar{b}_2 e^{-\alpha_0 t} t H(t), \tag{B19a}$$

$$\bar{b}_0 = \kappa_y/\kappa_x, \ \bar{b}_1 = \bar{b}_0[(\alpha_x - \alpha_0) - (\alpha_0 - \beta_y)], \ \bar{b}_2 = \bar{b}_0(\alpha_0 - \alpha_x)(\alpha_0 - \beta_y). \tag{B19b}$$

Similarly, applying $F^{-1}[\frac{s_y}{s_x}]$ to $\partial_y u_x$ leads to

$$F^{-1}\left[\frac{s_y}{s_x}\right] * \partial_y u_x = \bar{b}_0 \partial_y u_x + \bar{b}_1 \left[e^{-\alpha_0 t} H(t)\right] * \partial_y u_x + \bar{b}_1 \left[e^{-\alpha_0 t} H(t)\right] * [t \partial_y u_x], \tag{B20a}$$

$$b_0 = \bar{b}_0, \ b_1 = \bar{b}_1 + t\bar{b}_2, \ b_2 = -\bar{b}_2 \tag{B20b}$$

and using the same procedure we can obtain $F^{-1}[\frac{s_y}{s_x}] * \partial_y u_y$.

B2 ADE formulation

The goal of reformulating the convolution form of CFS-UPML into an ADE formulation is to facilitate the use of higher order time schemes and to be able to use the same time and space numerical schemes in the CFS-UPMLs as in the main domain. The main idea is to avoid having to handle convolution terms and introducing ADEs for these terms instead.

To do so, let us first derive the ADEs for convolution terms in $L(t)^* u_x(t)$. According to (B10) and (B11), in all cases two kinds of ADEs need to be derived. Using $R_L^{u_x,1}$, $R_L^{u_x,2}$ to denote the first and second convolution terms respectively, we have

(1) $\alpha_x \neq \alpha_y$

$$\frac{dR_L^{u_x,1}}{dt} = -\alpha_x R_L^{u_x,1} + u_x, \ \frac{dR_L^{u_x,2}}{dt} = -\alpha_y R_L^{u_x,1} + u_x. \tag{B21}$$

(2) $\alpha_x = \alpha_y = \alpha_0$

$$\frac{dR_L^{u_x,1}}{dt} = -\alpha_0 R_L^{u_x,1} + u_x, \ \frac{dR_L^{u_x,2}}{dt} = -\alpha_0 R_L^{u_x,1} + t u_x. \tag{B22}$$

In a similar way, we can derive the corresponding ADEs governing convolution terms when applying $L(t)$ to u_y, $F^{-1}[s_x/s_y]$ to $\partial_x u_x$ or $\partial_x u_y$ and $F^{-1}[s_y/s_x]$ to $\partial_y u_x$ or $\partial_y u_y$.

APPENDIX C: INTERMEDIATE STEPS TO DERIVE EQUATION (32) FROM (29)

Let us first focus on the term involving $\rho L(t) * \lambda_i u_i$:

$$\int_0^T \int_{\Omega_u} \lambda_i \rho L(t) * u_i^s d^3\mathbf{x} dt = \int_0^T \int_{\Omega_u} \lambda_i \rho \left[a_0 \ddot{u}_i^s + a_1 \dot{u}_i^s + a_2 u_i + L^c * u_i^s \right] d^3\mathbf{x} dt$$

$$= \int_{\Omega_u} \left[\left(\rho a_1 \lambda_i - \rho a_0 \dot{\lambda}_i \right) u_i^s + \rho a_0 \lambda_i \dot{u}_i^s \right]_0^T d^3\mathbf{x} + \int_0^T \int_{\Omega_u} \rho \left[a_0 \ddot{\lambda}_i - a_1 \dot{\lambda}_i + a_2 \lambda_i \right] u_i^s d^3\mathbf{x} dt$$

$$+ \int_{\Omega_u} \int_{t'=0}^{t'=T} \left[\int_{t=t'}^{t=T} L^c(t-t')\lambda_i(t) dt \right] u_i^s(t') dt' d^3\mathbf{x}$$

$$= \int_{\Omega_u} \left[\left(\rho a_1 \lambda_i - \rho a_0 \dot{\lambda}_i \right) u_i^s + \rho a_0 \lambda_i \dot{u}_i^s \right]_0^T d^3\mathbf{x} + \int_0^T \int_{\Omega_u} \rho \left[a_0 \ddot{\lambda}_i - a_1 \dot{\lambda}_i + a_2 \lambda_i + L^c * \lambda_i \right] u_i^s d^3\mathbf{x} dt, \tag{C1}$$

where L^c comes from eq. (24). We have introduced the time-shift $\tau \equiv T - t$, thus using the that $dt = -d\tau'$ as well as the fact that

$$\int_{t'=t}^{t'=T} L^c(t'-t)\lambda_i(t') dt' = - \int_{\tau'=\tau}^{\tau'=0} L^c(\tau - \tau')\lambda(T-\tau') d\tau'$$

$$= \int_{\tau'=0}^{\tau'=\tau} L^c(\tau - \tau')\lambda(T-\tau') d\tau'. \tag{C2}$$

Secondly, we rewrite the term involving $\lambda_i \partial_j (c_{ijkl} * \partial_k u_l^s)$ using integration by parts, invoking Gauss' theorem, using the symmetry properties of the elastic tensor and swapping the order of integration over t and t':

$$\int_0^T \int_{\Omega_u} \lambda_i \partial_j \left(\bar{c}_{ijkl} * \partial_k u_l^s \right) = \int_0^T \int_{\Gamma_u} \hat{n}_j \left(\bar{c}_{ijkl} * \partial_k u_l^s \right) \lambda_i d^2\mathbf{x} dt - \int_{\Omega_u} \int_{t=0}^{t=T} \partial_i \lambda_j(t) \left[\int_{t'=0}^{t'=t} \bar{c}_{ijkl}(t-t') \partial_k u_l^s(t') dt' \right] dt d^3\mathbf{x}$$

$$= - \int_0^T \int_{\Gamma_u^F} \hat{n}_j (\bar{c}_{ijkl} * \partial_k \lambda_l) u_i^s d^2\mathbf{x} dt + \int_{\Omega_u} \int_{t'=0}^{t'=T} \partial_j \left[\int_{t=t'}^{t=T} \bar{c}_{ijkl}(t-t') \partial_k \lambda_l(t) dt \right] u_i^s(t') dt' d^3\mathbf{x}, \tag{C3}$$

where the boundary terms vanish owing to the traction free surface condition $\hat{n}_j \bar{\sigma}_{ij}^\lambda = 0$ on Γ_u^F and the constraint condition that $\lambda = 0$ on Γ_u^D for the forward wavefield, and the similar constraint for the Lagrange multiplier wavefield $\hat{n}_j (\bar{c}_{ijkl} * \partial_k \lambda_l) = 0$ on Γ_u^F. If we define the second-order tensor

$$\bar{\sigma}_{ij}^\lambda(t) = \int_t^T \bar{c}_{ijkl}(t'-t) \partial_k \lambda_l(t') dt', \tag{C4}$$

it follows upon taking the variation of (C3) that the Lagrange multiplier field λ can be determined by the wave eq. (22), replacing the elastic stress tensor $\bar{\sigma}_{ij}$ in eq. (22) with the definition in (C4). Using the time-shifts $\tau = T - t$ then leads to

$$\bar{\sigma}_{ij}^\dagger(\tau) = \bar{\sigma}_{ij}^\lambda(T-\tau) = - \int_\tau^0 \bar{c}_{ijkl}(\tau - \tau') \partial_k \lambda_l(T-\tau') d\tau' = \int_0^\tau \bar{c}_{ijkl}(\tau - \tau') \partial_k \lambda_l(T-\tau') d\tau', \tag{C5}$$

where we have used the fact that $dt = -d\tau'$ and $\bar{c}_{ijkl}(t'-t) = \bar{c}_{ijkl}(\tau - \tau')$.

APPENDIX D: DERIVATION OF THE RECURSIVE CONVOLUTION SCHEME

Besides the derivation of the second-order recursive convolution scheme in (57)–(59), the CFS-PML formulation that we have introduced in this paper gives rise to 'higher order' convolution integrals (see Section 2.2.2 or Appendix A):

$$\left[e^{-bt} H(t) \right] * \left[t g(t) \right] = \int_0^t E(t-\tau)\tau g(\tau) d\tau, \tag{D1}$$

$$\left[e^{-bt} H(t) \right] * \left[t^2 g(t) \right] = \int_0^t E(t-\tau)\tau^2 g(\tau) d\tau, \tag{D2}$$

where $E(t) = e^{-bt} H(t)$, which comes from eliminating potential singularities in the CFS-PML parameter space as we have seen above. These higher order convolution integrals can be handled in a similar fashion as the ordinary one presented in eq. (73). However, the presence of τ and τ^2 modifies the coefficients in the recursive update scheme. Fortunately, these coefficients still have a closed-form expression, but they will now depend on time.

After applying the mean value theorem for integrals, the recursive update scheme of the convolution integrals can be summarized in the following generic form for any integer value $p = 0, 1, 2$:

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t}\Psi_n + g_{\theta_{n+1}}\int_{t_{n+\frac{1}{2}}}^{t_{n+\frac{3}{2}}} E^{n+1}\tau^p\mathrm{d}\tau + g_{\theta_n}\int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} \left[E^{n+1} - \mathrm{e}^{-b\Delta t}E^n\right]\tau^p\mathrm{d}\tau, \tag{D3}$$

where $\theta_k \in [t_{k-\frac{1}{2}}, t_{k+\frac{1}{2}}]$. Thus, in short form the recursive convolution scheme can be expressed as

$$\Psi_{n+1} = \mathrm{e}^{-b\Delta t}\Psi_n + \xi_p^{n+1}g_{\theta_{n+1}} + \xi_p^n g_{\theta_n} \tag{D4}$$

with

$$\xi_p^{n+1} = \int_{t_{n+\frac{1}{2}}}^{t_{n+\frac{3}{2}}} E^{n+1}\tau^p\mathrm{d}\tau, \quad \xi_p^n = \int_{t_{n-\frac{1}{2}}}^{t_{n+\frac{1}{2}}} \left[E^{n+1} - \mathrm{e}^{-b\Delta t}E^n\right]\tau^p\mathrm{d}\tau. \tag{D5}$$

For the sake of completeness, let us now write ξ_p^{n+1} and ξ_p^n for $p = 0, 1, 2$:

(1) $p = 0$

$$\xi_0^n = \frac{1}{b}\left(1 - \mathrm{e}^{-b\Delta t/2}\right)\mathrm{e}^{-b\Delta t/2}, \tag{D6a}$$

$$\xi_0^{n+1} = \frac{1}{b}\left(1 - \mathrm{e}^{-b\Delta t/2}\right). \tag{D6b}$$

(2) $p = 1$

$$\xi_1^n = \left(t_n - \frac{1}{b}\right)\xi_0^n + \frac{1}{b}\frac{\Delta t}{2}\mathrm{e}^{-b\Delta t/2}, \tag{D7a}$$

$$\xi_1^{n+1} = \left(t_{n+1} - \frac{1}{b}\right)\xi_0^{n+1} + \frac{1}{b}\frac{\Delta t}{2}\mathrm{e}^{-b\Delta t/2}. \tag{D7b}$$

(3) $p = 2$

$$\xi_2^n = \left(t_n^2 - \frac{2}{b}t_n + \frac{2}{b^2}\right)\xi_0^n + \frac{1}{b}\left[\Delta t\left(t_n - \frac{1}{b}\right) + \left(\frac{\Delta t}{2}\right)^2\right]\mathrm{e}^{-b\Delta t/2}, \tag{D8a}$$

$$\xi_2^{n+1} = \left(t_{n+1}^2 - \frac{2}{b}t_{n+1} + \frac{2}{b^2}\right)\xi_0^{n+1} + \frac{1}{b}\left[\Delta t\left(t_{n+1} - \frac{1}{b}\right) - \left(\frac{\Delta t}{2}\right)^2\right]\mathrm{e}^{-b\Delta t/2}. \tag{D8b}$$

Note that from (D6)–(D8) we can readily deduce that $\xi_{p-1}^k = \frac{1}{2}\partial\xi_p^k/\partial t_k$; Thus, if coefficients of the recursive convolution scheme are needed for $p > 2$ it is advantageous to calculate them in a descending order.

For small values of b, for instance in practice when $b < 10^{-6}$, terms proportional to $1/b^2$ and $1/b^3$ are highly sensitive to truncation errors with respect to b. Thus, to circumvent that we perform a Taylor expansion of (D6)–(D8) with respect to b to fourth-order accuracy. Furthermore, in order to be consistent with what we have done in the derivation of the new second-order convolution scheme of eq. (73), we also omit the new $O(\Delta t^4)$ terms that arise in the Taylor expansion of (D6)–(D8).

APPENDIX E: PLANE WAVE INCIDENCE IN THE 2-D CASE

In order to introduce a plane wave inside the domain Ω_1 shown in Fig. 5(b), the weak form of the wave eq. (1) can be written as:

$$\int_{\Omega_1} \rho\mathbf{w}\cdot\ddot{\mathbf{u}}_{\mathrm{d}}\mathrm{d}\Omega + \int_{\Omega_1} \nabla\mathbf{w} : \bar{\boldsymbol{\sigma}}_{\mathrm{d}}\mathrm{d}\Omega = \int_{\Gamma^\mathrm{I}} \mathbf{w}\cdot(\bar{\boldsymbol{\sigma}}_{\mathrm{d}}\cdot\hat{\mathbf{n}})\mathrm{d}\Gamma, \tag{E1}$$

where $\mathbf{u}_{\mathrm{d}}$ and $\bar{\boldsymbol{\sigma}}_{\mathrm{d}}$ are the displacement and stress diffracted wavefields, respectively. Similarly, along the adjacent layer of elements located inside of the interface the equation can be written as

$$\int_{\Omega_2} \rho\mathbf{w}\cdot\ddot{\mathbf{u}}\mathrm{d}\Omega + \int_{\Omega_2} \nabla\mathbf{w} : \bar{\boldsymbol{\sigma}}\mathrm{d}\Omega = -\int_{\Gamma^\mathrm{I}} \mathbf{w}\cdot(\bar{\boldsymbol{\sigma}}\cdot\hat{\mathbf{n}})\mathrm{d}\Gamma, \tag{E2}$$

where $\mathbf{u}$ and $\bar{\boldsymbol{\sigma}}$ are the displacement and stress total wavefields, respectively. Using the SEM technique for spatial discretization, for (E1) the resulting ordinary differential equation is

$$\mathbf{M}_{\mathrm{d}}\ddot{\mathbf{U}}_{\mathrm{d}} + \sum_e (\mathbf{K}_{\mathrm{d}}\mathbf{U}_{\mathrm{d}})^e = \mathbf{F}(\bar{\boldsymbol{\sigma}}_{\mathrm{d}}), \tag{E3}$$

while for (E2) it is:

$$\mathbf{M}\ddot{\mathbf{U}} + \sum_e (\mathbf{K}\mathbf{U})^e = -\mathbf{F}(\bar{\boldsymbol{\sigma}}). \tag{E4}$$

Furthermore, we have $\mathbf{F}(\bar{\boldsymbol{\sigma}}) = \mathbf{F}(\bar{\boldsymbol{\sigma}}_d) + \mathbf{F}(\bar{\boldsymbol{\sigma}}_i)$, where $\mathbf{F}(\bar{\boldsymbol{\sigma}}_i)$ are incident stress fields along Γ^I. The only remaining undefined value is then $\mathbf{F}(\bar{\boldsymbol{\sigma}}_d)$ along Γ^I, which can then be computed based on the diffracted-field expression $\mathbf{U}_d = \mathbf{U} - \mathbf{U}_i$ and $\mathbf{V}_d = \mathbf{V} - \mathbf{V}_i$. In the context of the Newmark scheme, based on (E3) we then have:

$$(\mathbf{V_d})_{n+1} = (\tilde{\mathbf{V}}_d)_{n+1} + \frac{\Delta t}{4} \mathbf{M}_d^{-1} \left(\mathbf{F}(\bar{\boldsymbol{\sigma}}_d) - \mathbf{M}_d(\mathbf{U}_d)_{n+1} - \sum_e (\mathbf{K}_d \mathbf{U}_d)_{n+1}^e \right) \tag{E5}$$

and for (E4) we have:

$$(\mathbf{V})_{n+1} = (\tilde{\mathbf{V}})_{n+1} + \frac{\Delta t}{4} \mathbf{M}^{-1} \left(-\mathbf{F}(\bar{\boldsymbol{\sigma}}_d) - \mathbf{F}(\bar{\boldsymbol{\sigma}}_i) - \mathbf{M}(\mathbf{U})_{n+1} - \sum_e (\mathbf{KU})_{n+1}^e \right). \tag{E6}$$

Thus, based on $(\mathbf{V_d})_{n+1} = (\mathbf{V})_{n+1} - (\mathbf{V_i})_{n+1}$ we have:

$$\left(\frac{\Delta t}{4} \mathbf{M}^{-1} + \frac{\Delta t}{4} \mathbf{M}_d^{-1} \right) \mathbf{F}(\bar{\boldsymbol{\sigma}}_d) = -(\mathbf{V_i})_{n+1} + \left[(\tilde{\mathbf{V}})_{n+1} + \frac{\Delta t}{4} \mathbf{M}^{-1} \left(-\mathbf{F}(\bar{\boldsymbol{\sigma}}_i) - \mathbf{M}(\mathbf{U})_{n+1} - \sum_e (\mathbf{KU})_{n+1}^e \right) \right]$$

$$- \left[(\tilde{\mathbf{V}}_d)_{n+1} + \frac{\Delta t}{4} \mathbf{M}_d^{-1} \left(-\mathbf{M}_d(\mathbf{U_d})_{n+1} - \sum_e (\mathbf{K}_d \mathbf{U}_d)_{n+1}^e \right) \right].$$

After obtaining $\mathbf{F}(\bar{\boldsymbol{\sigma}}_d)$ we can then finish the time integration of (E3) and (E4) based on the Newmark scheme.