

# Comparison of expansion-based explicit time-integration schemes for acoustic wave propagation

Carlos Spa<sup>1</sup>, Otilio Rojas<sup>2</sup>, and Josep de la Puente<sup>3</sup>

## ABSTRACT

We have developed a von Neumann stability and dispersion analysis of two time-integration techniques in the framework of Fourier pseudospectral (PS) discretizations of the second-order wave equation. The first technique is a rapid expansion method (REM) that uses Chebyshev matrix polynomials to approximate the continuous solution operator of the discrete wave equation. The second technique is a Lax-Wendroff method (LWM) that replaces time derivatives in the Taylor expansion of the solution wavefield with their equivalent spatial PS differentiations. In both time-integration schemes, each expansion term  $J$  results in an extra application of the spatial differentiation operator; thus, both methods are similar in terms of their implementation and the freedom to arbitrarily increase accuracy by using more expansion terms. Nevertheless, their limiting Courant-Friedrichs-Lewy

stability number  $S$  and dispersion inaccuracies behave differently as  $J$  varies. We establish the  $S$  bounds for both methods in cases of practical use,  $J \leq 10$ , and we confirm the results by numerical simulations. For both schemes, we explore the dispersion dependence on modeling parameters  $J$  and  $S$  on the wavenumber domain, through a new error metric. This norm weights errors by the source spectrum to adequately measure the accuracy differences. Then, we compare the theoretical computational costs of LWM and REM simulations to attain the same accuracy target by using the efficiency metric  $J/S$ . In particular, we find optimal  $(J, S)$  pairs that ensure a certain accuracy at a minimal computational cost. We also extend our dispersion analysis to heterogeneous media and find the LWM accuracy to be significantly better for representative  $J$  values. Moreover, we perform 2D wave simulations on the SEG/EAGE Salt Model, in which larger REM inaccuracies are clearly observed on waveform comparisons in the range  $J \leq 3$ .

## INTRODUCTION

The accuracy and efficiency of geophysical imaging techniques such as full-waveform inversion (FWI) or reverse time migration (RTM) depend strongly on the numerical techniques used to model seismic wave propagation. In the acoustic approximation, phases can be correctly reproduced by means of the scalar wave equation. Many methods exist to model wave propagation in acoustic media. The most common methods are the explicit in time and high order in space finite differences (FDs) (for a thorough review, see [Moczo et al., 2014](#)). Despite this method's broad success, the accurate and efficient approximation of the time and space derivatives remain a topic open for study. On one hand, time integration is classically attained using low-order FD schemes. On the other hand, the

spatial accuracy could be improved by using spectral differentiation. The Fourier pseudospectral time-domain (PSTD) method is based on such spectral differentiation by means of Fourier transforms (FTs). Early PSTD methods, similar to FD methods of their time, make use of explicit second-order time discretizations ([Gazdag, 1981](#); [Kosloff and Baysal, 1982](#); [Kosloff et al., 1984](#); [Reshef et al., 1988](#)). Such PSTD methods present stricter stability constraints compared with their FD counterparts, in addition to suffering from stringent accuracy constraints related to time discretization errors. As a consequence, efforts were made to explore time-integration schemes for PSTD based on polynomial expansions, hoping to mitigate such errors and relax stability and accuracy constraints resulting from oversimplified time integrators.

Manuscript received by the Editor 10 July 2019; revised manuscript received 27 December 2019; published ahead of production 4 February 2020.

<sup>1</sup>Universidad Técnica Federico Santa María, Mathematics Department, Av. España 1680, Valparaíso, Chile. E-mail: carlos.spa@usm.cl (corresponding author).

<sup>2</sup>Barcelona Supercomputing Center, Jordi Girona 29, 08034 Barcelona, Spain and Universidad Central de Venezuela, Faculty of Sciences, Caracas 1040, Venezuela. E-mail: otilio.rojas@bsc.es.

<sup>3</sup>Barcelona Supercomputing Center, Jordi Girona 29, 08034 Barcelona, Spain. E-mail: josep.delapuate@bsc.es.

© 2020 Society of Exploration Geophysicists. All rights reserved.

Tal-Ezer (1986) and Tal-Ezer et al. (1987) introduce the Chebyshev polynomial expansion in time integration, which results in the rapid-expansion method (REM) for PSTD wave-propagation problems (Kosloff et al., 1989; Stoffa and Pestana, 2009; Pestana and Stoffa, 2010), a technique that approximates the continuous solution operator of the semidiscrete wave equation. The approximation is given by a truncated Chebyshev polynomial expansion of such operator up to a term  $J$  that ensures sufficient accuracy for the time integration interval  $\Delta t$  under consideration. Tal-Ezer (1986) uses the convergence properties of the Chebyshev expansion to lower bound  $J$  and also to study the stability and time resolution of the REM algorithm. Subsequently, Tal-Ezer et al. (1987) use the  $J$  estimates from their previous work to model 2D acoustic waves to highlight the stability of the algorithm for large  $J$  values. They find that, in this case,  $J$  must be slightly higher than  $R\Delta t$  with  $R$  being an appropriate scaling of the Chebyshev terms. The estimation of  $R$  in Tal-Ezer et al. (1987) is not clearly outlined but might result from the von Neumann dispersion analysis for PSTD for a second-order time discretization. Moreover, they find that  $R$  is directly proportional to the largest wavenumber in the simulated domain but independent of  $\Delta t$ , which in turn means that  $J$  grows proportionally with  $\Delta t$ . This result motivated the development of the so-called one-step REM algorithm, in which scaled Chebyshev terms are iteratively constructed according to a known recurrence formula and are then linearly combined by the expansion Bessel coefficients with nonnegligible magnitude for the desired time span  $\Delta t$ . Given that Bessel coefficients are narrow banded, this REM implementation is adequate for large  $\Delta t$  (in the order of seconds), but it is restricted to separable sources in space and time that allow decoupling the computations of matrix terms and coefficients. One-step REM is used in Kosloff et al. (1989), Stoffa and Pestana (2009), and Pestana and Stoffa (2010) for 2D acoustic and elastic wave simulations, with efficient computational implementations for equations in second-order form.

The more conventional recursive REM is used as an explicit time-integration algorithm for shorter  $\Delta t$  values (in the order of milliseconds). Recursive REM is discussed in Pestana and Stoffa (2010), and then it is explored by Tessmer (2011) on acoustic modeling and RTM. In RTM, sources have different time histories and each source must be modeled independently by a time-stepping evolution algorithm. As pointed out by Pestana and Stoffa (2010) and Tessmer (2011), the truncation criterion  $J \approx R\Delta t$  is no longer precise for short  $\Delta t$  and more terms are actually needed to ensure stability and achieve accurate simulation results. Given the  $R$  estimation in Tal-Ezer et al. (1987),  $J$  can be related to the Courant-Friedrichs-Lewy (CFL) number, hereafter also referred to as  $S$ , and the aforementioned works show that the limiting  $S$  increases with  $J$  in multiple numerical experiments. In particular, Tessmer (2011) compares recursive REM solutions for  $J = 4, 5, 6$  and fixed  $S = 0.4$  against analytical solutions, and dispersion errors clearly decay as  $J$  grows. He also presents low-dispersive results on simple heterogeneous media for  $S = 0.2$ . Undoubtedly, all previous results provide insight into the numerical properties of REM. However, a formal bound on  $S$  for a given  $J$ , as established from a stability and dispersion von Neumann analysis is absent on all the works mentioned above. Furthermore, the literature only hints at what a minimum  $J$  value should be to ensure spectral convergence and accuracy, but no detailed study is shown of the relation of dispersion errors with  $J$ . These deficiencies motivate the present paper in

which, for REM, we quantify the dependence of stability and dispersion on modeling parameters  $S$  and  $J$ .

An alternative time-integration method also coupled to PSTD is the classic Lax-Wendroff method (LWM), in which high-order terms in the temporal Taylor expansion of the solution wavefield are replaced by their equivalent spatial derivatives by means of the differential equation. Thus, computations of high-order LWM terms result in additional spatial differentiation, in a similar way to REMs expansion terms. The LWM stability and dispersion properties for FD-based wave propagation have been explored in early works, where only a few expansion terms were sufficient to match the spatial accuracy, limited to fourth order in Blanch and Robertsson (1997) and to tenth order in Dablain (1986). Alternatively, to overcome the stability limitation of PS methods under second-order time integration, few LWM-based schemes have been developed with optimal expansion coefficients. In Soubaras and Zhang (2008), these coefficients result from a constrained Remez optimization, and the new method allows for large integration steps up to the Nyquist time sampling, as given by the frequency content of input data in RTM applications. The work in Chu et al. (2009) proposes two LWM-type methods with modified coefficients of fourth- and sixth-order accuracy that exhibit larger stability bounds than their counterparts, based on standard Taylor coefficients. They demonstrate the larger time stepping of modified methods on numerical examples. Recently, Amundsen and Pedersen (2017) use the classic second-order Leapfrog integration for a reference time step  $\Delta t$ , as the basis for the mathematical formulation of new arbitrarily high-order schemes. Such schemes allow stable wavefield time marching for steps of sizes  $2\Delta t$ ,  $3\Delta t$ , and in general,  $n\Delta t$ . At the implementation level, these schemes also rely on optimally weighted averages of nested applications of the PS Laplacian to the wavefield, so the computational burden per time step is nearly the same as that for standard LWM. However, schemes in Amundsen and Pedersen (2017) present larger stability limits, as formally established by these authors. They also highlight that their first two candidates, i.e., time advancing schemes for steps  $2\Delta t$  and  $3\Delta t$ , correspond to the modified LWM methods of fourth- and sixth-order accuracy given in Chu et al. (2009).

Chu et al. (2009) also give the limiting  $S$  values of standard LWM up to the 20th order, and these values fluctuate for increasing  $J$ . More precisely, the maximum CFL splits into two increasing sequences, one defined for even  $J$  values, and the other sequence for odd  $J$  values. The variations of the limiting  $S$  value of standard LWM on high-order FD spatial discretizations can also be deduced from results in Crase (1990) for arbitrary  $J$ , and explicitly observed in Rojas et al. (2017) for  $J < 10$ . Alternatively, Chen (2009) lists maximum CFL values that are nearly constant above the eighth-order correction for a standard LWM-PSTD, also developed for the second-order wave equation. It is worth noting that these highly similar CFL limits come from a numerical bounding of the von Neumann dispersion relation, potentially sensitive to roundoff errors as  $J$  grows, as discussed in the body of this paper. The theoretical study of stability and dispersion applied to standard LWM and Nyström methods in Chen (2009) is classic, as is the study in Amundsen and Pedersen (2017). However, a thorough assessment of dispersion errors and practical implications of the method's accuracy is missing in those references. Such evaluations demand error metrics as introduced in Kristeková et al. (2006), and any additional frequency-dependent misfit with physical relevance, as the one introduced in this work.

Thus, driven by a lack of formal stability and dispersion analyses of REM-PSTD methods, and the need of complementary accuracy studies of LWM-PSTD dispersion of relevance to geophysical applications, here we present a comparative assessment of their numerical properties. The implementation of both methods results in further applications of the spatial differentiation operator for each additional temporal correction term. This implementation similarity is exploited by making a full comparison of such REM and LWM properties, under free variations of  $J$  and  $S$ , in homogeneous and heterogeneous media. In addition, their global computational efficiency is analyzed in practical terms. For concreteness, we only focus on standard LWM in this work, but our study methodology can be easily applied to the aforementioned nonclassic LWM to understand their accuracy dependence on  $(J, S)$ . Such an analysis remains open for future work.

Multisteping integration strategies have also been used on seismic wave-propagation modeling to reduce time discretization errors. In the framework of high-order spatial FD methods, numerical analyses of a variety of Adams, backward differentiation, and Runge-Kutta integration strategies can be found in [Ghrist et al. \(2000\)](#), [Zhang et al. \(2012, 2014\)](#), and [Bohlen and Wittkamp \(2016\)](#). Although possible, the implementation of multisteping techniques into PSTD methods yield formulations whose von Neumann analysis is order dependent and less flexible than the recursive formulations of REM and LWM. A good example is given by the fourth-order Nyström integration, also developed in [Chen \(2009\)](#). Alternatively, a novel strategy to remove time dispersion consists of filtering such errors from either wavelets or recorded waveforms, with no direct impact on the cost of numerical simulations ([Stork, 2013](#)). On acoustic or elastic propagation models, these filters can be applied in combination with classic low-order time discretization schemes, as used in RTM and FWI applications by [Stork \(2013\)](#) and [Koene et al. \(2017\)](#). If the propagation models include nonlinearities (e.g., the source terms in rupture problems), these strategies would not be straightforward. In this work, we omit these two complementary strategies, i.e., the aforementioned multisteping and filtering techniques.

The rest of this paper is divided into eight sections. Numerical algorithms presents the formulation of REM and LWM time-integration schemes for a Fourier PS spatial discretization of the second-order wave equation. In the two sections that follow, we perform a von Neumann stability and dispersion analysis of these methods, revealing their dependence on parameters  $J$  and  $S$ . In these sections, we also validate the limiting  $S$  values through comparisons against 1D numerical simulations and then we extend the analytical bounds to 2D and 3D domains. In section computational cost on homogeneous simulations, we discuss the trade-offs between the accuracy and computational cost of the methods presented, by using the cost indicator  $C = J/S$ . In addition, we study the impact of material heterogeneities on REM and LWM dispersion in one and two dimensions in section dispersion assessment in heterogeneous media, and we provide a practical example in section numerical experiment: SEG/EAGE salt model. Finally, we summarize our conclusions and point out future works in the last section.

## NUMERICAL ALGORITHMS

Let us consider the acoustic wave equation:

$$\frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} + \mathcal{L}^2 u(\mathbf{x}, t) = 0, \quad (1)$$

where  $u(\mathbf{x}, t)$  is the acoustic pressure at time  $t$  and the spatial operator is defined according to  $\mathcal{L} = \iota c \nabla$ , where  $\iota = \sqrt{-1}$ . We take

$\mathbf{x} \in \Omega$ , a 3D spatial domain across which the speed of sound  $c(\mathbf{x})$  may vary. To numerically solve equation 1, we use the Fourier pseudospectral (PS) method for the spatial term, combined with either the REM or the LWM for integration of the temporal term. All three methods are described in the following.

### The Fourier PS method

The continuous spatial term in equation 1, i.e.,  $\mathcal{L}^2$ , can be expanded as

$$\begin{aligned} \mathcal{L}^2 u(\mathbf{x}, t) = & -c^2(\mathbf{x}) \{ \mathcal{F}_x^{-1} [k_x^2 \mathcal{F}_x [u(\mathbf{x}, t)]] \\ & + \mathcal{F}_y^{-1} [k_y^2 \mathcal{F}_y [u(\mathbf{x}, t)]] + \mathcal{F}_z^{-1} [k_z^2 \mathcal{F}_z [u(\mathbf{x}, t)]] \}, \end{aligned} \quad (2)$$

where  $k_x$ ,  $k_y$ , and  $k_z$  are the wavenumber in each Cartesian dimension and  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  denote the continuous 1D FT and its inverse, respectively, in the direction of the subscript. The formulation of Fourier PS methods is based on a discrete analog of equation 2, in which 1D derivatives are computed in the wavenumber domain by means of discrete FT (DFT) and brought back to the spatial domain by means of an inverse FT. Then, we consider a rectangular medium  $\Omega$  discretized by a regular grid of equal spacing  $\delta$ , i.e.,  $x = i\delta$ ,  $y = j\delta$ , and  $z = l\delta$ , and we denote grid points by  $\mathbf{x}$ . In addition, we denote as  $\mathbf{L}^2$  the PS approximation of  $\mathcal{L}^2$  on a given grid sampling of  $u$ , according to equation 2. Thus,  $\mathbf{L}^2$  is a discrete matrix operator with same dimensions as the rectangular mesh. It is important to remark that the DFT is a good approximation to the FT as long as the Nyquist-Shannon sampling theorem is fulfilled in each dimension, e.g.,  $k_{\max} \leq \pi/\delta$ , and the periodic extension of the spatial distribution is continuous. In this case, the PS scheme has a spatial accuracy of  $\mathcal{O}(\delta^N)$ . For details, extensions, and applications of the PS method, we refer to [Gazdag \(1981\)](#), [Kosloff and Baysal \(1982\)](#), [Kosloff et al. \(1984\)](#), and [Reshef et al. \(1988\)](#).

### The REM

Here, we adopt the one-step REM formulation for second-order wave equations used in [Kosloff et al. \(1989\)](#), [Pestana and Stoffa \(2010\)](#), and [Tessmer \(2011\)](#). Upon PS discretization of equation 1, we can obtain the following semidiscrete solution in terms of the initial condition  $u(\mathbf{x}, 0)$ :

$$u(\mathbf{x}, t) = -u(\mathbf{x}, -t) + 2 \cos(\mathbf{L}t)u(\mathbf{x}, 0) \quad (3)$$

where time is still continuous. The term  $\cos(\mathbf{L}t)$  represents the cosine matrix operator. Then, we take discrete times  $t_n = n\Delta t$ , where  $\Delta t$  is the temporal updating step and define  $u^n(\mathbf{x}) = u(\mathbf{x}, t_n)$ . A fully discrete version of equation 3, at the time level  $t_n$  may read

$$u^{n+1}(\mathbf{x}) = -u^{n-1}(\mathbf{x}) + 2 \cos(\mathbf{L}\Delta t)u^n(\mathbf{x}). \quad (4)$$

The method in equation 4 is an explicit temporal updating scheme. Nevertheless, evaluation of the global operator  $\cos(\mathbf{L}\Delta t)$  is cumbersome; hence, it is of practical importance to use the polynomial expansion:

$$\cos(\mathbf{L}\Delta t) = \sum_{j=0}^{\infty} \mu_{2j} \mathcal{J}_{2j}(R\Delta t) \mathcal{Q}_{2j} \left( \frac{\mathbf{L}}{R} \right). \quad (5)$$

Note that  $\mathcal{J}$  are the Bessel functions of the first kind,  $\mathcal{Q}$  are the Chebyshev polynomials, and coefficients  $\mu_0 = 1$  and  $\mu_j = 2$  for  $j \neq 0$ . The scaling parameter  $R$  is typically chosen as the upper bound of the eigenvalue spectrum of the matrix operator  $\mathbf{L}^2$ . A bound for  $R$  that is explicitly given in aforementioned references, and also previously used in [Tal-Ezer et al. \(1987\)](#), reads

$$R \geq R_0 \equiv \frac{\pi c \sqrt{D}}{\delta} = \frac{\pi S \sqrt{D}}{\Delta t}, \quad (6)$$

where  $D$  is the dimension of the space and  $S$  is defined as the stability Courant number that relates the wave speed  $c$  to the temporal and spatial steps

$$S = c \frac{\Delta t}{\delta}. \quad (7)$$

The computation of high-order Chebyshev polynomials appearing in equation 5 is calculated by the recursive formula:

$$\mathcal{Q}_{2j+2}\left(\frac{\mathbf{L}}{R}\right) = 2\left(1 + 2\left(\frac{\mathbf{L}}{R}\right)^2\right)\mathcal{Q}_{2j}\left(\frac{\mathbf{L}}{R}\right) - \mathcal{Q}_{2j-2}\left(\frac{\mathbf{L}}{R}\right), \quad (8)$$

while the two initial terms are  $\mathcal{Q}_0 = I$  and  $\mathcal{Q}_2 = 2(\mathbf{L}/R)^2 + 1$ . The sum in equation 5, in practical terms, must be truncated up to a maximum number of terms  $J + 1$ . Finding suitable values of  $J$  is part of the work that we will pursue in the following sections. As reference, we consider two lower bounds recommended in the literature. Based on the work of [Tal-Ezer \(1986\)](#), the convergence of the REM underlying series is guaranteed by the condition

$$J > R\Delta t \geq \pi S \sqrt{D}, \quad (9)$$

as used in [Kosloff et al. \(1989\)](#) and [Stoffa and Pestana \(2009\)](#). The second bound is given by [Pestana and Stoffa \(2010\)](#) and refers to the magnitude of the Bessel functions in equation 5. They suggest that

$$\begin{aligned} J &> \max_j (|\mathcal{J}_{2j}(R\Delta t)| > 10^{-3}) - 1 \\ &= \max_j (|\mathcal{J}_{2j}(\pi S \sqrt{D})| > 10^{-3}) - 1. \end{aligned} \quad (10)$$

The largest of the values in equations 9 and 10 should be used to ensure convergence and accuracy.

### The LWM

Alternatively, the Taylor expansion of  $u(\mathbf{x}, t)$  at discrete times  $t_{n+1}$  and  $t_{n-1}$  results in

$$u^{n+1}(\mathbf{x}) = 2u^n(\mathbf{x}) - u^{n-1}(\mathbf{x}) + 2 \sum_{j=1}^{\infty} \frac{(\Delta t)^{2j}}{(2j)!} \frac{\partial^{2j} u^n(\mathbf{x})}{\partial t^{2j}}, \quad (11)$$

which is an explicit temporal updating scheme of increasing accuracy with additional terms in the sum. To avoid evaluating high-order temporal derivatives, the following recursive formula can be used for  $j \geq 1$ :

$$\frac{\partial^{2j} u^n(\mathbf{x})}{\partial t^{2j}} = -\mathbf{L}^2 \frac{\partial^{2j-2} u^n(\mathbf{x})}{\partial t^{2j-2}}. \quad (12)$$

In this way, all of the temporal derivatives can be replaced by spatial derivatives. Being based upon a Taylor expansion, this scheme has a time accuracy of order  $\mathcal{O}(\Delta t^{2j})$ . Applications and analyses of LWM in the frameworks of FD and PS schemes can be found in [Dablain \(1986\)](#), [Crase \(1990\)](#), [Rojas et al. \(2017\)](#), [Chen \(2009\)](#), and [Chu et al. \(2009\)](#).

### STABILITY

Let us consider a 3D plane wave

$$u(\mathbf{x}, t) = u_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{num}} t)}, \quad (13)$$

where  $u_0$  is the amplitude;  $i = \sqrt{-1}$  is the imaginary unit as defined before;  $\mathbf{k}$  is the wavenumber vector, i.e.,  $\mathbf{k} = (k_x, k_y, k_z)$ ; and  $\omega_{\text{num}}$  is the numerical frequency. Let us assume a discrete space-time grid such as  $\mathbf{x} = (i, j, l)\delta$  and  $t = n\Delta t$ . In the case of PSTD-REM, we can insert equation 13 into equations 4 and 5 to obtain the von-Neumann relation:

$$\cos(\omega_{\text{num}} \Delta t) = \sum_{j=0}^J \mu_{2j} \mathcal{J}_{2j}(R\Delta t) \mathcal{Q}_{2j}\left(\frac{ic|\mathbf{k}|}{R}\right). \quad (14)$$

where  $|\mathbf{k}| = \sqrt{k_x^2 + k_y^2 + k_z^2}$  is the modulus of the wavenumber.

Similarly, we can reproduce the same analysis with LWM using equations 11 and 12 resulting in

$$\cos(\omega_{\text{num}} \Delta t) = 1 + \sum_{j=1}^J (-1)^j \frac{(c|\mathbf{k}|\Delta t)^{2j}}{(2j)!}. \quad (15)$$

In general, an algorithm is stable when the value of the magnitude of the right side of the von Neumann relation is smaller than one. In the case of REM, we obtain

$$\left| \mathcal{J}_0(R\Delta t) + 2 \sum_{j=1}^J \mathcal{J}_{2j}(R\Delta t) \mathcal{Q}_{2j}\left(\frac{ic|\mathbf{k}|}{R}\right) \right| \leq 1, \quad (16)$$

whereas for LWM, we obtain

$$\left| 1 + \sum_{j=1}^J (-1)^j \frac{(c|\mathbf{k}|\Delta t)^{2j}}{(2j)!} \right| \leq 1. \quad (17)$$

It is useful to present both stability relations in a dimensionless form. Thus, we rewrite equations 16 and 17, in the case of  $R = R_0$ , for wavenumbers attainable by spectral differentiation  $\mathbf{k} = \hat{\mathbf{k}}\pi/\delta$  with  $\hat{\mathbf{k}}$  defined at each component in the range  $[0, 1]$ . In the case of REM, we obtain

$$\left| \mathcal{J}_0(\pi S \sqrt{D}) + 2 \sum_{j=1}^J \mathcal{J}_{2j}(\pi S \sqrt{D}) \mathcal{Q}_{2j}\left(i \frac{|\hat{\mathbf{k}}|}{\sqrt{D}}\right) \right| \leq 1, \quad (18)$$

whereas for LWM, we can write

$$\left| 1 + \sum_{j=1}^J (-1)^j \frac{(S\pi|\hat{\mathbf{k}}|)^{2j}}{(2j)!} \right| \leq 1. \quad (19)$$

In Figure 1, we plot the left side of equation 18 as a function of the stability number for  $J \in [1, 4]$  and  $D = 1$ . The curves for  $J = 1$  and  $J = 2$  cross the stability limit, i.e., the value one, rather sharply. However, for  $J = 3$  and  $J = 4$ , there are  $S$  values between one and two, where the inequality 18 does not hold strictly, but the values are not large. We refer to these as regions of mild instability because practical simulations could still be carried out, if the number of iterations is small. In particular, if we relax the inequalities in equations 18 and 19 to  $|\cdot| \leq 1 + \tau$ , simulations should be stable for iteration amounts roughly smaller than  $\tau^{-1}$ . To visualize the impact of the mild stability zones in 1-D, we evaluate the left side of the von-Neumann equations in the complete  $k \in [0, 1]$  range, i.e., equations 18 and 19, for  $J \in [0, 10]$ , and stability numbers  $S \in [0.1, 4.0]$ . The results are shown in Figure 2, in which the white areas are stable regions (i.e.,  $|\cdot| \leq 1$ ), the black areas are unstable regions (i.e.,  $|\cdot| > 1 + \tau$ ), and the gray areas are mildly stable regions.

In Table 1, we present the maximum stability number  $S_{\max}$  obtained for each method, using  $\tau = 10^{-4}$ . The results show that, for REM, the values of  $S_{\max}$  monotonically increase with the number of terms  $J$ . However, for LWM, we observe alternating  $S_{\max}$  values as  $J$  grows, which hints at two increasing sequences of  $S_{\max}$  according to the parity of  $J$ . This nonmonotonic behavior of  $S_{\max}$  has been described in [Chu et al. \(2009\)](#), in the case of LWM with modified coefficients. Compared to 1D results, 2D and 3D  $S_{\max}$  values scale exactly by the  $\sqrt{D}$  factor in the case of LWM, whereas REM scaling is less regular. We would like to remark that, for LWM (equation 11), the Taylor-expansion term  $(\Delta t)^{2j}/(2j)!$  goes to zero rapidly, which can result in round-off errors when the summation is not performed carefully. In [Chen \(2009\)](#), a von Neumann analysis of a PSTD-LWM renders results very different from those in Table 1 at larger  $J$  values. Such results seem affected by the aforementioned round-off errors.

Our analytical results in Table 1 have been further corroborated with 1D and 2D numerical simulations using meshes with  $N = 500$  and  $N = 500 \times 500$  nodes, respectively, and performing  $n_t = 10000$  temporal iterations. The results are not shown here, but the differences between numerical and analytical  $S_{\max}$  are within 3%.

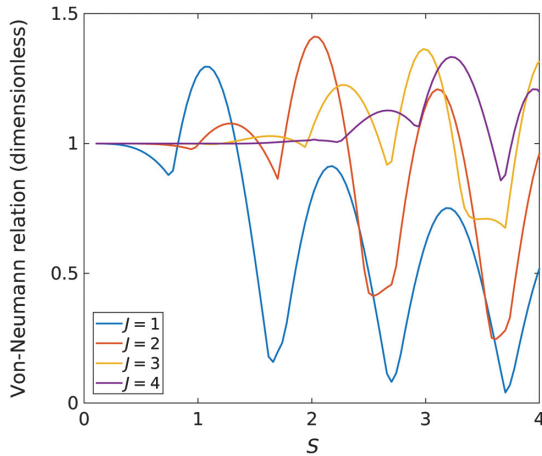


Figure 1. The left side of the REM stability condition for  $D = 1$  and different  $S$  values using equation 18 from  $J = 1$  to  $J = 4$ .

In summary, we have found that stability strongly increases with  $J$  in both methods, albeit only monotonically in REM. Moreover, mild stability regions enable larger  $S$  values in simulations having moderate numbers of time iterations. We confirm these observations with simulations having  $n_t < 1/\tau$ . Overall, the ratio between the REM and LWM stability limit increases with  $J$ , although LWM is less sensitive to mild instabilities than REM.

## DISPERSION ERROR

In this section, we present a complete analysis of the dispersion error for REM and LWM using a PSTD method. Each algorithm results in a numerical wave velocity  $c_{\text{num}}$  which, in general, is different from the physical and expected velocity  $c$ . For the two methods at hand,  $c_{\text{num}}$  depends, at each wavenumber  $\mathbf{k}$ , on the stability number  $S$ . The dispersion analysis can be performed analytically for homogeneous media if we use  $c = \omega/|\mathbf{k}|$  and  $c_{\text{num}} = \omega_{\text{num}}/|\mathbf{k}|$  in equations 14 and 15. In dimensionless form and using  $R = R_0$ , we obtain the following expression for REM:

$$\frac{c_{\text{num}}}{c} = \frac{1}{S\pi|\hat{\mathbf{k}}|} \arccos \left( \mathcal{J}_0(\pi S\sqrt{D}) + 2 \sum_{j=1}^J \mathcal{J}_{2j}(\pi S\sqrt{D}) \mathcal{Q}_{2j} \left( i \frac{|\hat{\mathbf{k}}|}{\sqrt{D}} \right) \right), \quad (20)$$

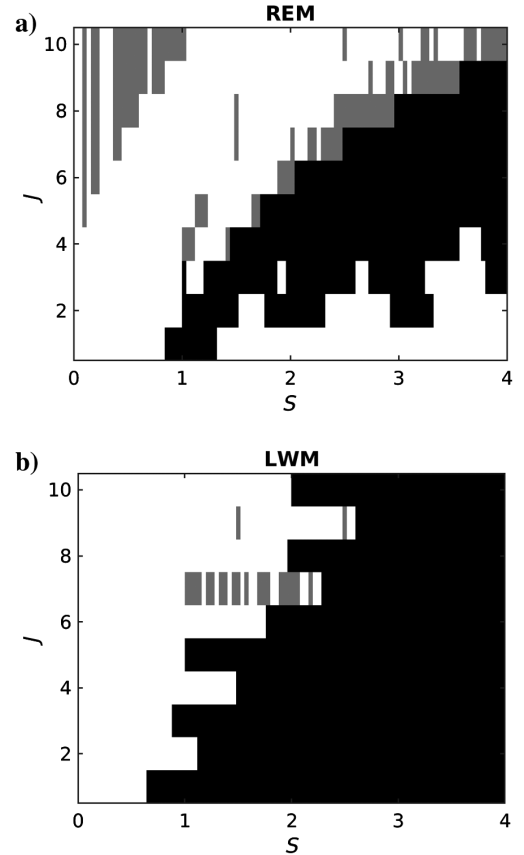


Figure 2. Analytical dependence of  $S_{\max}$  on  $J$  for (a) REM (equation 18) and (b) LWM (equation 19) for the 1D case. The white areas are stable, the black areas are unstable, and the gray areas are mildly stable ( $\tau = 10^{-4}$ ).



whereas for LWM, we obtain

$$\frac{c_{\text{num}}}{c} = \frac{1}{S\pi|\hat{\mathbf{k}}|} \arccos \left( 1 + \sum_{j=1}^J (-1)^j \frac{(S\pi|\hat{\mathbf{k}}|)^{2j}}{(2j)!} \right). \quad (21)$$

In this case, we restrict our study to  $|\mathbf{k}| \in [0, \pi/\delta]$ , i.e.,  $|\hat{\mathbf{k}}| \in [0, 1]$ . We remark that PS methods possess isotropy in their dispersion error; i.e., dispersion depends on the modulus of the wavenumber  $|\mathbf{k}|$ , but not on the plane-wave propagation angle (see, e.g., [Fornberg, 1998](#) and [Spa et al., 2009](#)).

In Figure 3, we show examples of dispersion curves generated by REM-PSTD (Figure 3a) and LWM-PSTD (Figure 3b). These examples are obtained using  $S = 0.62$  and considering only the first three polynomial terms. As expected, errors decrease with  $J$  for both methods. Nevertheless, the distribution of the errors is very different in REM and LWM. REM displays high dispersion errors at low wavenumbers. On the other hand, LWM is inaccurate only at high wavenumbers. As a corollary, REM performs poorly at modeling quasistatic behaviors, but it excels close to  $|\mathbf{k}| = \pi/\delta$ . LWM works exactly in the opposite way. These fundamentally different behaviors make it difficult to quantify, and fairly compare, REM and LWM dispersion for wave-propagation problems. For instance, if we study the maximum wavenumber at which  $c_{\text{num}}$  is close to  $c$ , as is customary in FDTD, REM will perform poorly because of inaccuracies at the low  $k$  range, whereas computing the average error along the  $k$  axis might not be useful if the signals under study do not have a flat spectrum.

To obtain geophysically relevant conclusions, we introduce the following metric that we will refer to as the weighted dispersion error:

$$\epsilon_{S,J} = \frac{1}{N_k} \sum_{k=0}^{\pi/\delta} \left| 1 - \frac{c_{\text{num}}(|\mathbf{k}|, S, J)}{c} \right| r(|\mathbf{k}|), \quad (22)$$

where  $N_k$  is the total number of samples in the  $k$  space. Moreover, we explicitly write  $c_{\text{num}}(|\mathbf{k}|, S, J)$  to highlight that the numerical velocity depends on the wavenumber modulus  $|\mathbf{k}|$ , the stability number  $S$ , and the number of polynomial terms  $J$ . The weighting function  $r(|\mathbf{k}|)$  is taken as the spectrum of the impulse signal used.

In Figure 4, we show snapshots from REM and LWM 1D simulations, and we explore the error behavior of both methods. In this case, we use a Ricker wavelet with a minimum wavenumber discretized with two points. We choose combinations of  $S$  and  $J$  that yield similar errors in terms of the norm, equation 22. Then, we

analyze the results by comparing them to the analytical solution in terms of the adimensional phase misfit (PM) and the envelope misfit (EM) from [Kristeková et al. \(2006\)](#). The PM and EM metrics have been originally developed as space-time metrics of errors in seismic signals — EM being closely related to differences in amplitude and less sensitive to phase, and PM having a complementary behavior. Two signals in phase with differing amplitude would have  $\text{PM} = 0$  but nonzero EM (which would coincide with the rms misfit), and two signals differing in phase but not in amplitude would display  $\text{EM} = 0$  and nonzero PM. A value of  $\text{PM} = 100$  is obtained from two signals out of phase (or with a phase difference of  $\pi$ ) and  $\text{EM} = 100$  from a signal having double the amplitude of the other. We note, however, that we compare spatial signals on a 1D domain, instead of the temporal traces used in the original publication. Hence, we refer to the misfits as space-wavenumber misfits instead of time-frequency misfits. The space-wavenumber misfits for the first simulation case (Figure 4a) are given in Figure 5, and for the second simulation case (Figure 4b), they are given in Figure 6. We can clearly see that PM distinguishes clearly between subsonic and supersonic dispersion. It is also clear, especially for the higher error case (Figure 5) that REM is far worse at low wavenumbers than LWM, consistent with the observations above. It is also clear that, in broad terms, the space-wavenumber misfits are comparable for LWM and REM under the same error value from equation 22, except for the PM of REM being remarkably better than the PM of LWM at the highest error threshold, hence the error norm, equation 22, seems a reasonable misfit function to quantify dispersion.

In Figure 7, we show the logarithm of  $\epsilon_{S,J}$  for REM (Figure 7a) and LWM (Figure 7b). Just like in the stability analysis, we focus on the first 10 polynomial terms and on the stability numbers from  $S = 0.1$  to  $S = 4$ . In this figure, we clip values of  $\epsilon_{S,J} > 1$ . Overall, at low  $J$  values both methods perform similarly, but the REM results improve faster with increasing  $J$  values than LWM results. In the case of REM, we have also plotted the minimum  $J$  computed using equations 9 and 10. The limit comes from the accuracy and convergence criteria used by [Pestana and Stoffa \(2010\)](#). In terms of dispersion error, the limit seems to result in errors below  $10^{-2}$  if  $J$  is chosen greater than the suggested value.

Nevertheless, the most striking feature of Figure 7 is that, for both methods, there is a sharp accuracy limit at  $S \leq 1$ . Above such values, the dispersion accuracy breaks down for all  $J$  values. This is explained by considering equation 7 with a spatial sampling of  $\Delta x = \lambda_{\text{min}}/2$ , which is the maximum sampling allowed by the Nyquist theorem. Moreover, if the velocity  $c$  is written in terms of the minimum wavelength, we have that  $c = \lambda_{\text{min}}/T_{\text{min}}$ , for  $T_{\text{min}}$  the

**Table 1. The 1D, 2D, and 3D stability limits  $S_{\text{max}}$  of REM and LWM using the first 10 expansion terms, found analytically from equations 18 and 19. We use  $\tau = 10^{-4}$  in the analytical results.**

Method	$J = 1$	$J = 2$	$J = 3$	$J = 4$	$J = 5$	$J = 6$	$J = 7$	$J = 8$	$J = 9$	$J = 10$
REM 1D	0.820	1.004	0.984	1.432	1.704	2.056	2.476	2.952	3.484	3.998
LWM 1D	0.636	1.100	0.872	1.472	0.980	1.764	2.256	1.936	2.608	1.992
REM 2D	0.580	0.712	0.696	1.000	1.204	1.460	1.756	2.095	2.460	2.824
LWM 2D	0.449	0.778	0.616	1.040	0.692	1.247	1.595	1.368	1.844	1.408
REM 3D	0.472	0.580	0.568	0.816	0.998	1.192	1.428	1.712	2.008	2.300
LWM 3D	0.367	0.635	0.503	0.849	0.565	1.018	1.392	1.117	1.505	1.150

period related to the minimum wavelength  $\lambda_{\min}$ . Finally, let us consider a stability number  $S \geq 1$  and introduce all these terms into equation 7. The result is  $\Delta t \geq T_{\min}/2$ , which means that, for  $S \geq 1$ , the temporal sampling is greater than half of the minimum period involved in simulations; therefore, the Nyquist theorem is not fulfilled.

We remark, however, that  $S \geq 1$  should only be dismissed for homogeneous cases. In heterogeneous materials, such limitation can be relaxed and large  $S$  values can fulfill the minimum Nyquist criterion for temporal resolution, and it can then lead to low dispersive results. For instance, let us consider  $c_{\min}$  and  $c_{\max}$  as the minimum and the maximum values of wave speed on a heterogeneous domain, respectively. In the case of uniform time-space meshing, the minimum Nyquist resolution requires

$$\Delta t \leq \frac{1}{2f_{\max}}, \quad \delta \leq \frac{c_{\min}}{2f_{\max}}, \quad (23)$$

$f_{\max}$  being the maximum frequency of interest of the source spectrum. Now, using equation 7 and assuming the limiting spatial resolution for PSTD, the maximum stability number is given by

$$S_{\max} = c_{\max} \frac{2\Delta t f_{\max}}{c_{\min}}. \quad (24)$$

Finally, after considering the maximum time step allowed in equation 23, we write

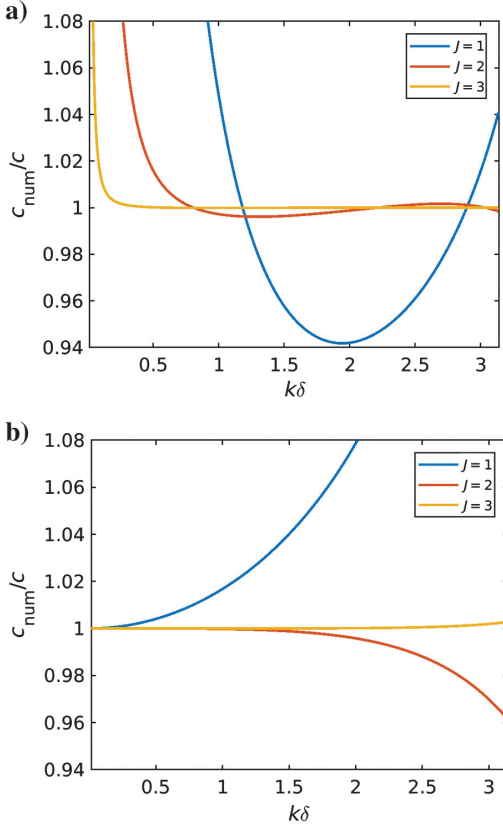


Figure 3. Dispersion error curves  $c_{\text{num}}/c$  in 1D against the wave-number  $k\delta$  for (a) REM and (b) LWM, using the first three polynomial terms. In all cases, the stability number is fixed at  $S = 0.62$ .

$$S_{\max} \leq \frac{c_{\max}}{c_{\min}}. \quad (25)$$

Thus, the condition  $S \geq 1$  can assure the Nyquist resolution in space and time, and it may also allow accurate simulations in heterogeneous simulations.

In summary, the REM and LWM dispersions behave differently along the  $k$  spectrum; hence, we have introduced a novel metric to appropriately compare their accuracy. In both methods, accuracy increases with  $J$ , as long as the stability number remains in the range  $S \leq 1$ . However, this limit can be relaxed up to  $c_{\max}/c_{\min}$  for a material of variable wave speed  $c$ .

## COMPUTATIONAL COST ON HOMOGENEOUS SIMULATIONS

We have already presented an analysis of the stability and the dispersion errors for REM and LWM. However, there are still some features that should be studied to get a fair comparison between

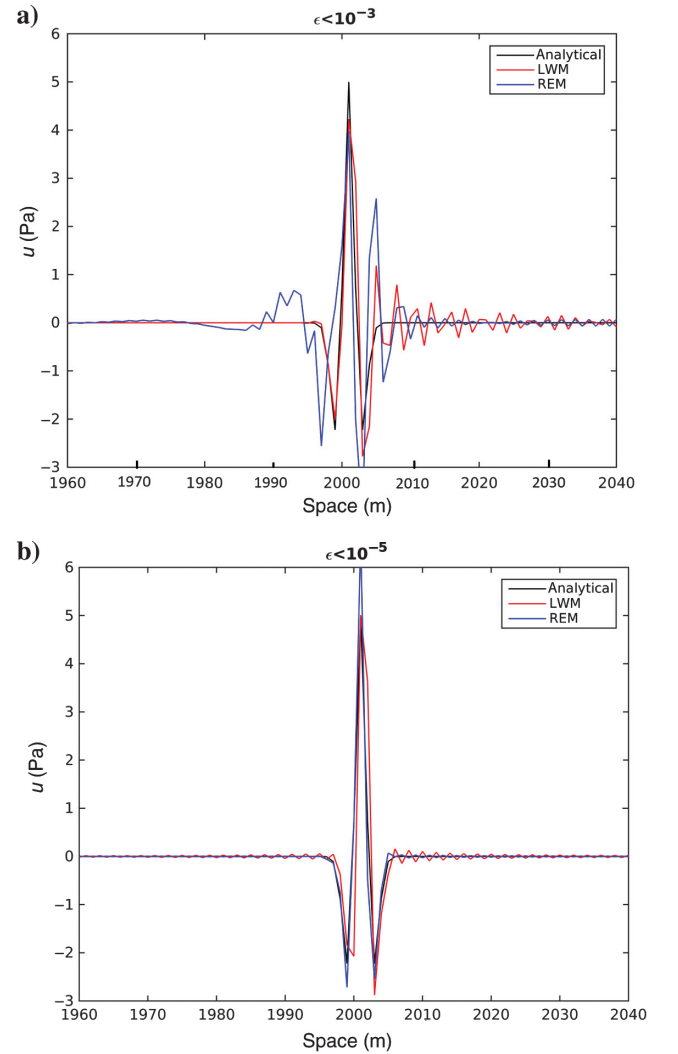


Figure 4. Comparison of numerical and analytical solutions in 1D for LWM and REM. The simulations satisfied (a)  $\epsilon < 10^{-3}$  and (b)  $\epsilon < 10^{-5}$ .

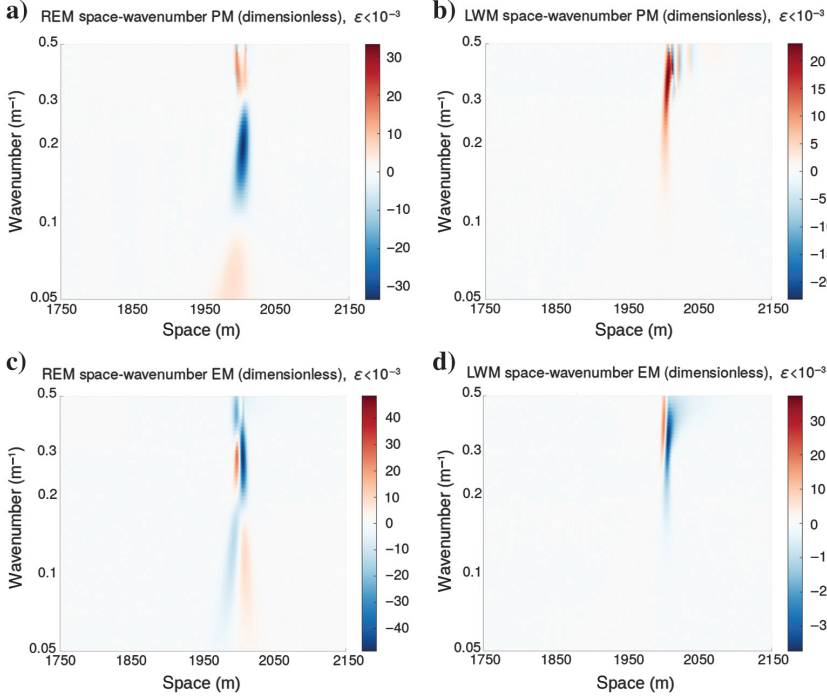


Figure 5. Space-wavenumber phase (upper panels) and envelope (the lower panels) misfits for REM (the left panels) and LWM (the right panels) in percent, for 1D simulations at  $\epsilon < 10^{-3}$ .

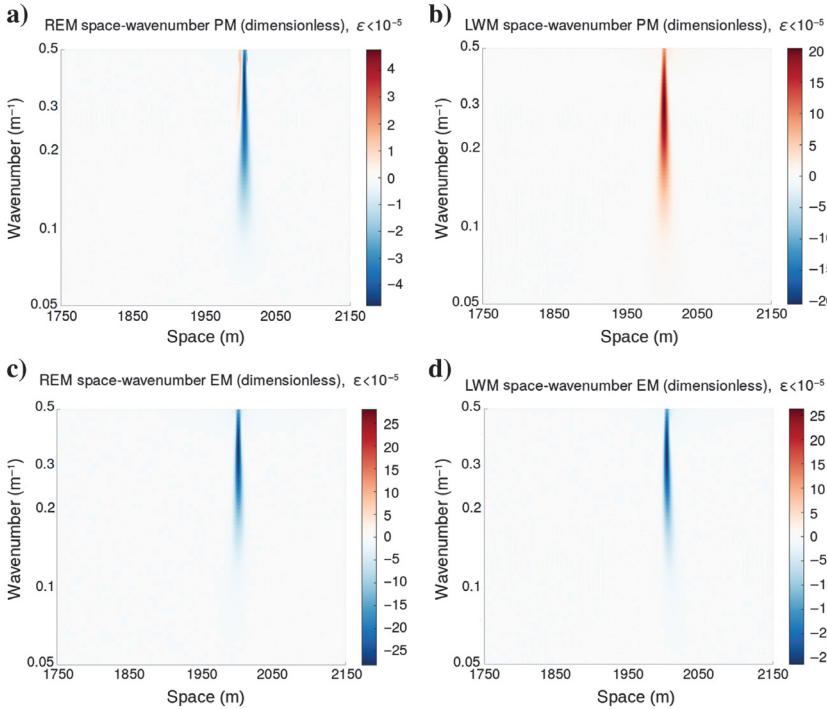


Figure 6. Space-wavenumber phase (the upper panels) and envelope (the lower panels) misfits for REM (the left panels) and LWM (the right panels) in percent, for 1D simulations at  $\epsilon < 10^{-5}$ .

both methods and to give solid conclusions about which method is more computationally efficient. Therefore, in this section, we focus the analysis on the balance between accuracy and computational cost of the algorithms, which is necessary to establish a fair criterion for comparing both implementations. In particular, we will study error and cost as a function of  $S$  and the number of polynomial terms  $J$ . Critical to the scope of this paper, REM- and LWM-PSTD schemes involve, for each time step,  $J$  applications of the operator  $\mathbf{L}^2$ ; hence, the cost of each time step is directly proportional to  $J$  for both methods. On the other hand, the stability number  $S$  fixes the total number of time steps required to reach a prescribed simulation time. The computational cost of a simulation with both methods is roughly proportional to  $J$  and inversely proportional to  $S$ ; hence, we will use the cost proxy  $C = J/S$  in the following. The term  $C$  is not an absolute measure, but it can be used to compare simulations of the same numerical problem using the same spatial grid.

We analyze, for each maximum error threshold  $\epsilon$ , which  $(J, S)$  pair attains the prescribed accuracy with minimum cost. The results are presented in Figure 8a for REM and Figure 8b for LWM. Moreover, we analyze in detail eight accuracy thresholds  $\epsilon < 10^{-(2+i)}$  for  $i = 1, \dots, 8$ . The first conclusion that is observed from the results is that, regardless of the method used, the lower the error threshold the higher the cost  $C$ , as expected. In general, from  $\epsilon < 10^{-3}$  to  $\epsilon < 10^{-10}$ , the lowest values of the computational cost  $C$  are obtained with REM, in almost all of the threshold errors tested, except for  $\epsilon < 10^{-3}$ , where the minimum values of  $C$  are identical in both high-order methods, and  $\epsilon < 10^{-4}$ , where LWM provides a better performance. To illustrate this point, in Table 2, the  $(J, S)$  pairs are presented for both high-order methods, considering the lowest values of the computational cost  $C$  obtained at each corresponding threshold of the weighted dispersion error. We clearly observe that REM almost always provides lower  $C$  values than LWM and hence, in broad terms, REM performs better than LWM with few exceptions. In particular, in the range  $\epsilon < 10^{-3}$  to  $\epsilon < 10^{-6}$ , both methods require the same number of polynomial terms, except at  $\epsilon < 10^{-4}$ , and the main difference between both methods is the larger stability number in REM. In contrast, for more restrictive thresholds of the weighted dispersion error, more specifically from  $\epsilon < 10^{-7}$  to  $\epsilon < 10^{-10}$ , LWM needs one additional polynomial term. In any case, the observation that higher accuracies are attained in REM by means of smaller  $J$  values and higher  $S$  values than LWM is interesting. A simple corollary, for both methods, is that stability numbers in the



range  $S \in [0.58 - 0.86]$  appear to be optimal in terms of cost to attain a fixed error.

In summary, REM seems to reach computationally accurate results for a smaller computational cost in most but not all of the cases studied. To make the results comparable with classic time integration methods, we can take a look at LWM with  $J = 1$ , which is equivalent to the second-order FDs in time. For this case, it is typically acknowledged that we need approximately  $S = 0.1$  to obtain dispersion-free results. This results, for  $\epsilon < 10^{-4}$ , in a cost of  $C \sim 10$ . To put this value in perspective, either REM or LWM at the same cost result in accuracies in the order of  $\epsilon < 10^{-10}$ . Alternatively, either REM or LWM can obtain the same errors with a cost  $C \sim 3.4$  hence, it is approximately three to four times faster than classic PSTD with second-order finite differencing in time.

## DISPERSION ASSESSMENT IN HETEROGENEOUS MEDIA

Realistic seismic applications involve heterogeneous media with significant variations in wave speed. Typical geologic models exhibit wave-propagation velocities generally increasing with depth, albeit large lateral variations are also common. When using fixed (i.e., regular and Cartesian) computational grids in methods such as FDTD and PSTD,  $\Delta t$  and  $\delta$  are generally suboptimal to accommodate the dispersion and stability requirements of the whole model.

Specifically, the region of lowest velocity restricts the spatial sampling  $\delta$  to satisfy the minimum ppw, whereas the region of fastest propagation determines the maximum temporal sampling  $\Delta t$  to enforce stability. Hence, having fixed the space-time grid, the stability number  $S$  and the resolution in terms of ppw vary with wave speed, being high in the fast velocity regions and low in the slow regions.

## Dispersion analysis

We first focus on assessing analytically the effect of wave-speed heterogeneity on the dispersion errors of PSTD using REM and LWM time integration. Hence, the results of this first analysis do not include the interface treatment, only the accuracy for waves inside each region, individually. The objective is thus having an upper limit on the accuracy in real applications, which would add errors coming from the interface treatment to those shown here.

When we use the PSTD method in a Cartesian and regular space-time grid,  $\delta$  and  $\Delta t$  are fixed quantities, regardless of heterogeneities in the grid. Let us assume that  $c$  is a variable quantity in space such that  $c \in [c_{\min}, c_{\max}]$ . If we define the maximum wavenumber in the modeled space as

$$k_{\max} \equiv \max |\mathbf{k}(c = c_{\min})|, \quad (26)$$

then the Shannon-Nyquist theorem imposes a minimum spatial sampling

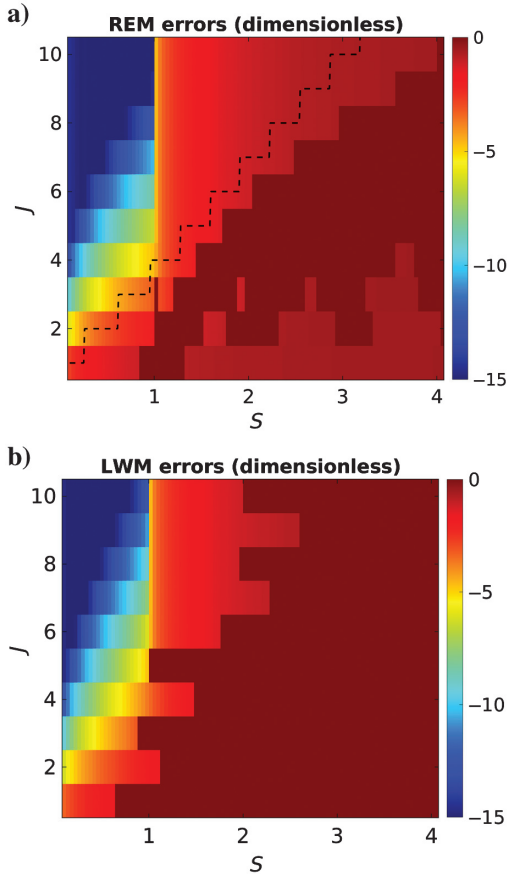


Figure 7. Dispersion errors from equation 22 for 1D in logarithmic scale for (a) REM and (b) LWM.

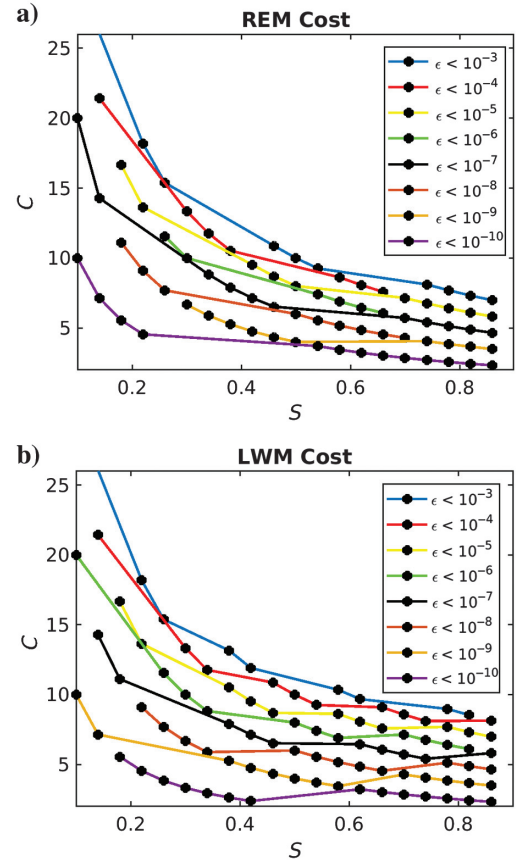


Figure 8. Computational cost  $C$  versus the stability number  $S$  in 1D for (a) REM and (b) LWM for eight different error thresholds. Each dot represents a different  $(J, S)$  pair.

$$\delta = \pi/k_{\max}, \quad (27)$$

and the stability condition in equation 7 results in

$$\Delta t = \frac{\pi S}{k_{\max} c}. \quad (28)$$

Similarly, the wavenumber is limited by the spatial sampling; hence,

$$\mathbf{k} = |\hat{\mathbf{k}}| \frac{\pi c_{\min}}{\delta c}. \quad (29)$$

In addition, for REM, we must use a single  $R$  value that we take as the maximum in the domain (see, e.g., Tal-Ezer et al., 1987, and Pestana and Stoffa, 2010); hence, from equation 6, we obtain

$$R_{\max} = c_{\max} \sqrt{D} k_{\max}. \quad (30)$$

Now, we can rewrite REMs stability, equation 16, for heterogeneous media as

$$\left| \mathcal{J}_0 \left( \pi S \sqrt{D} \frac{c_{\max}}{c} \right) + 2 \sum_{j=1}^J \mathcal{J}_{2j} \left( \pi S \sqrt{D} \frac{c_{\max}}{c} \right) \mathcal{Q}_{2j} \left( \frac{|\hat{\mathbf{k}}| c}{\sqrt{D} c_{\max}} \right) \right| \leq 1, \quad (31)$$

where we have further used the relation  $|\mathbf{k}|/k_{\max} = c_{\min}/c$ . Similarly, for LWM, we obtain the following stability condition:

$$\left| 1 + \sum_{j=1}^J (-1)^j \frac{\left( \pi S |\hat{\mathbf{k}}| \frac{c_{\min}}{c} \right)^{2j}}{(2j)!} \right| \leq 1. \quad (32)$$

We note that, for LWM, the same stability limit can be found as in the homogeneous case because the numerator in equation 32 is the same in every region (i.e.,  $S$  is proportional to  $c$ ). Hence, the value of  $S$  for the maximum velocity layer determines the largest possible time step. Similarly, for REM we find that the stability equation 31 coincides with equation 18 if we take  $c = c_{\max}$ ; hence, the stability limit in the maximum velocity region is in fact the stability limit for all regions. In conclusion, and as expected, the maximum time step

for both methods in heterogeneous media is the same as that in a homogeneous medium with  $c = c_{\max}$ .

Regarding dispersion, similarly, we obtain for REM

$$\frac{c_{\text{num}}}{c} = \frac{1}{S \pi |\hat{\mathbf{k}}| c_{\min}} \arccos \left( \mathcal{J}_0 \left( \pi S \sqrt{D} \frac{c_{\max}}{c} \right) + 2 \sum_{j=1}^J \mathcal{J}_{2j} \left( \pi S \sqrt{D} \frac{c_{\max}}{c} \right) \mathcal{Q}_{2j} \left( \frac{|\hat{\mathbf{k}}| c}{\sqrt{D} c_{\max}} \right) \right), \quad (33)$$

whereas for LWM, we obtain

$$\frac{c_{\text{num}}}{c} = \frac{1}{S \pi |\hat{\mathbf{k}}| c_{\min}} \arccos \left( 1 + \sum_{j=1}^J (-1)^j \frac{\left( \pi S |\hat{\mathbf{k}}| \frac{c_{\min}}{c} \right)^{2j}}{(2j)!} \right). \quad (34)$$

In this case, for LWM, the dispersion in the minimum velocity region, for a given  $S$  value, is the same as it would be in a homogeneous simulation with  $c = c_{\min}$ . Furthermore, the same accuracy is reached in each layer because of the proportionality of  $S$  to  $c$ . In the case of REM, we cannot reach such simple conclusions. As expected, in the homogeneous limit, equations 33 and 34 become equations 21 and 20, respectively.

Let us consider the hypothetical test of a 1D medium with three different velocity layers:  $c_0$ ,  $2c_0$ , and  $4c_0$ , all sharing the same spatial grid, i.e., the same  $\delta$  spacing. Because of the properties of PSTD, we use 2 ppw for the slowest layer; hence, the other layers have 4 and 8 ppw, respectively. In Figure 9, we present the results of the dispersion error in the highest velocity layer using (Figure 9a) REM and (Figure 9b) LWM. When compared with Figure 7, it is clear that the dispersion breakdown has moved from  $S \geq 1$  to  $S \geq 4$ . We remark that this  $S$  value is local to the layer under study and that the low-velocity layer is still limited by  $S \geq 1$ . However, and crucially, we note that LWM is, for a given  $S$  value, more accurate in this layer than in the  $c_0$  layer (which is equivalent to Figure 7) according to its higher ppw. This is due to the  $c/c_{\min}$  factor in equation 34, which, effectively, results in the dispersion at the  $4c_0$  layer for a given  $S$  being the same as the dispersion at the  $c = c_0$  medium at  $S/4$ . The case of REM is less encouraging and, overall, results in worse dispersion for all  $S$  values.

**Table 2. Minimum computational cost  $C$  for each error threshold on metric, equation 22, across multiple  $(J, S)$  pairs using REM and LWM.**

REM	$\epsilon < 10^{-3}$	$\epsilon < 10^{-4}$	$\epsilon < 10^{-5}$	$\epsilon < 10^{-6}$	$\epsilon < 10^{-7}$	$\epsilon < 10^{-8}$	$\epsilon < 10^{-9}$	$\epsilon < 10^{-10}$
$J$	2	3	3	4	4	5	5	6
$S$	0.860	0.860	0.700	0.860	0.660	0.860	0.660	0.860
$C$	2.326	3.488	4.286	4.651	6.061	5.814	7.576	6.997
LWM	$\epsilon < 10^{-3}$	$\epsilon < 10^{-4}$	$\epsilon < 10^{-5}$	$\epsilon < 10^{-6}$	$\epsilon < 10^{-7}$	$\epsilon < 10^{-8}$	$\epsilon < 10^{-9}$	$\epsilon < 10^{-10}$
$J$	2	2	3	4	5	6	6	7
$S$	0.860	0.580	0.660	0.740	0.820	0.860	0.740	0.820
$C$	2.326	3.448	4.545	5.405	6.098	6.977	8.108	8.537

Finally, in Table 3, we present the weighted dispersion errors obtained from equation 22 for REM and LWM for our three-layered media test. For all of the cases analyzed, we only consider three different polynomial terms,  $J = 4$ ,  $J = 6$ , and  $J = 8$ , and the stability number is fixed according to the maximum stability number allowed for each method and polynomial term. First, the results in Table 3 for REM quantify that errors are reduced with increasing  $J$  and are similar for all layers, albeit accuracy is much worse than would be obtained in the homogeneous case. We remark that the  $J$  values used are fully compliant with the  $J$  restrictions proposed

in the literature and summarized in equations 9 and 10, for all three cases. Conversely, for LWM the dispersion error remains very low and constant for the three different layers, thus resulting in better accuracy than in the case of REM. Note that this behavior is observed independent of the number of terms accounted in the temporal integration, and it is complementary to observations in high-order-in-time FDTD results in Moczo et al. (2014) and Rojas et al. (2017).

## 2D numerical test

To support the findings of the previous section, we study REM and LWM numerical solutions for 2D heterogeneous media. The first numerical experiment corresponds to a square grid with three different flat, homogeneous layers of velocity  $c_0$ ,  $2c_0$ , and  $4c_0$ , from top to bottom. These are the same velocities studied analytically in the previous section. In this case, we use a Ricker point source with two samples at the minimum period, located at the top of the lowest velocity layer. We remark that we model sources without high-order time corrections. We set the spatial sampling to  $\delta = 40$  m and the lowest wave speed to  $c_0 = 2000$  m · s<sup>-1</sup>. Our spatial grid has  $N \times N$  points, with  $N = 150$ , and we let the system evolve during 1 s for each given  $(J, S)$  pair. As a reference, in Figure 10a, we show the REM simulation snapshot at 1 s using  $J = 10$  and  $S = 1.2$ . Note that this  $(J, S)$  pair guarantees stability and approximate dispersion errors of  $\epsilon < 10^{-8}$ , according to the results from Figure 9. The corresponding waveform along the vertical 1D grid at the source axis location is plotted in Figure 10b.

For more practical applications, we also study the REM and LWM accuracy when using the low range of polynomial terms,  $J = 1, 2, 3$ . In these tests, we fix  $S = 0.4$  in the fastest layer of speed  $4c_0$ . Note that this particular  $S$  selection guarantees stability for both methods and all  $J$  values tested according to Table 1 (the 2D case).

As a consequence, we revisit Figure 9 and list the REM and LWM dispersion errors, equation 22, in the cases of  $J \leq 3$  in Table 4. Whereas, LWM yields moderate accuracies approximately  $\epsilon < 10^{-4}$  even for  $J = 1$ , REM exhibits larger errors  $\epsilon > 10^{-3}$  for  $J = 1, 2$ , and finally becomes competitive when  $J = 3$  terms are used. To qualitatively assess the dispersion effects on numerical waveforms for this low range of  $J$ , we perform the corresponding 2D simulations, and the results, along the source axis, are depicted in Figure 11a for REM and Figure 11b for LWM. The accuracy level shown by these waveforms is consistent with the analytical errors in Table 4. We clearly observe that, only for  $J = 3$ , REM provides acceptable results, in contrast to the early high precision

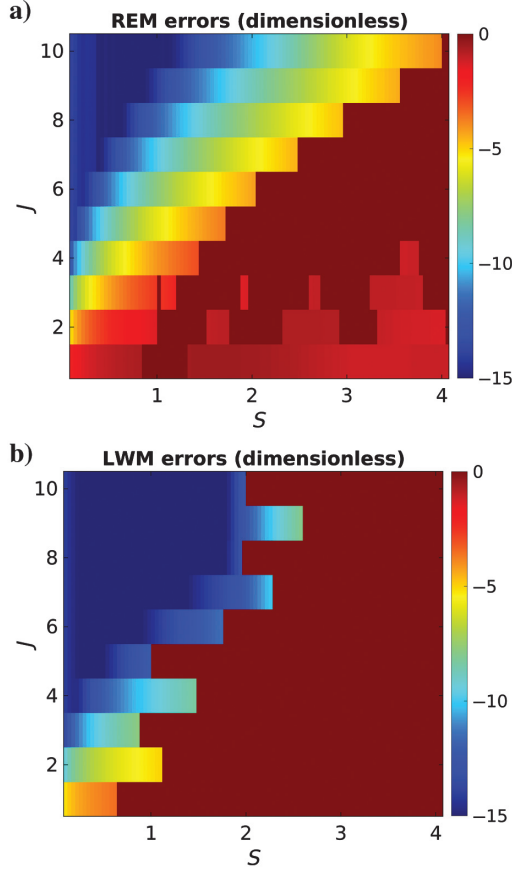


Figure 9. Dispersion errors in 1D for (a) REM and (b) LWM at the  $4c_0$  layer in a grid restricted in space by  $c_0$ .

**Table 3.** Example of errors using equation 22, in logarithmic scale, obtained at each region of a three-layered medium, for three different  $J$  values.

REM	$c_0$	$2c_0$	$4c_0$	LWM	$c_0$	$2c_0$	$4c_0$
$S$	0.350	0.700	1.400	$S$	0.360	0.720	1.440
$\epsilon(J = 4)$	-2.541	-2.751	-3.105	$\epsilon(J = 4)$	-8.123	-8.079	-8.483
$S$	0.490	0.980	1.92	$S$	0.440	0.880	1.750
$\epsilon(J = 6)$	-4.011	-4.112	-4.312	$\epsilon(J = 6)$	-11.261	-11.140	-11.440
$S$	0.590	1.180	2.360	$S$	0.480	0.960	1.920
$\epsilon(J = 8)$	-5.601	-5.793	-6.206	$\epsilon(J = 8)$	-15.022	-14.510	-13.360

noted in LWM results after  $J = 1$ . It is worth noting that, for the problems computed here, our  $J$  choice for REM, according to equations 9 and 10, should be at least  $J = 2$ . Clearly, such a  $J$  choice gives worse results than LWM with  $J = 1$  according to Table 4; thus, LWM outperforms REM in terms of computing cost.

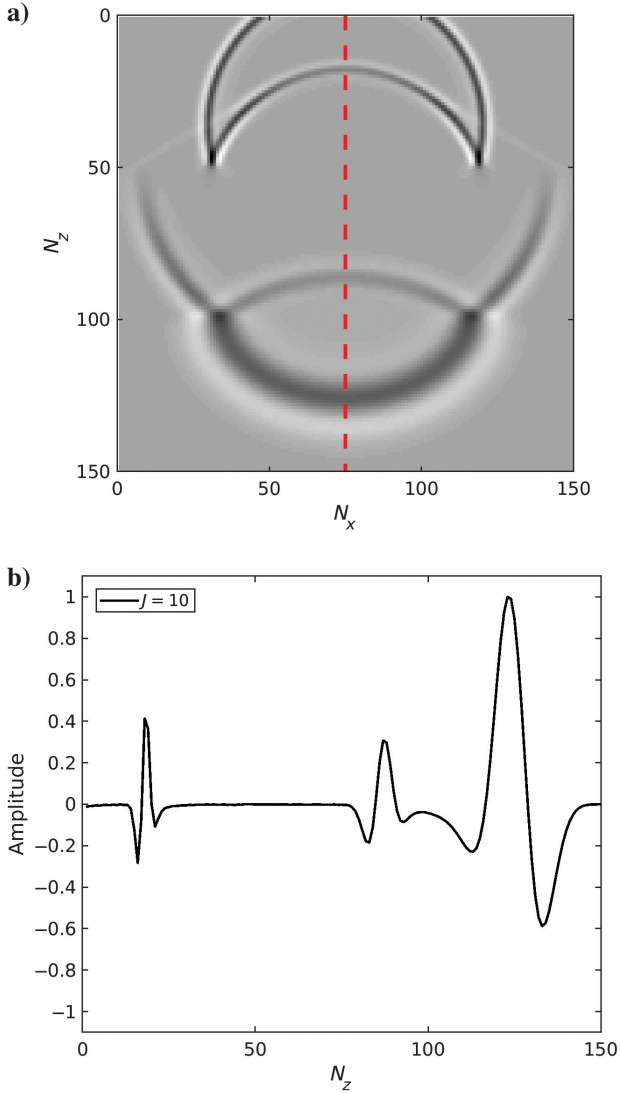


Figure 10. (a) A snapshot of the REM 2D simulation of the three-layered media at 1 s of time. (b) REM waveform along a straight-line cut of the simulation along the  $y$ -axis. These results are obtained fixing  $J = 10$  and  $S = 1.2$  (at the highest velocity layer).

**Table 4. The 2D dispersion errors, equation 22, in logarithmic scale, obtained at each region of a three-layered medium, for the first three  $J$  terms, after fixing  $S = 0.4$ .**

REM	$c_0$	$2c_0$	$4c_0$	LWM	$c_0$	$2c_0$	$4c_0$
$\epsilon(J=1)$	-0.041	-0.342	-0.632	$\epsilon(J=1)$	-3.325	-3.622	-3.923
$\epsilon(J=2)$	-1.261	-1.551	-1.821	$\epsilon(J=2)$	-6.219	-6.501	-6.802
$\epsilon(J=3)$	-3.015	-3.251	-3.621	$\epsilon(J=3)$	-9.259	-9.506	-9.807

## NUMERICAL EXPERIMENT: SEG/EAGE SALT MODEL

As a practical application, we use the SEG/EAGE Salt Model. We choose a 2D slice of the model centered at inline = 6800 m

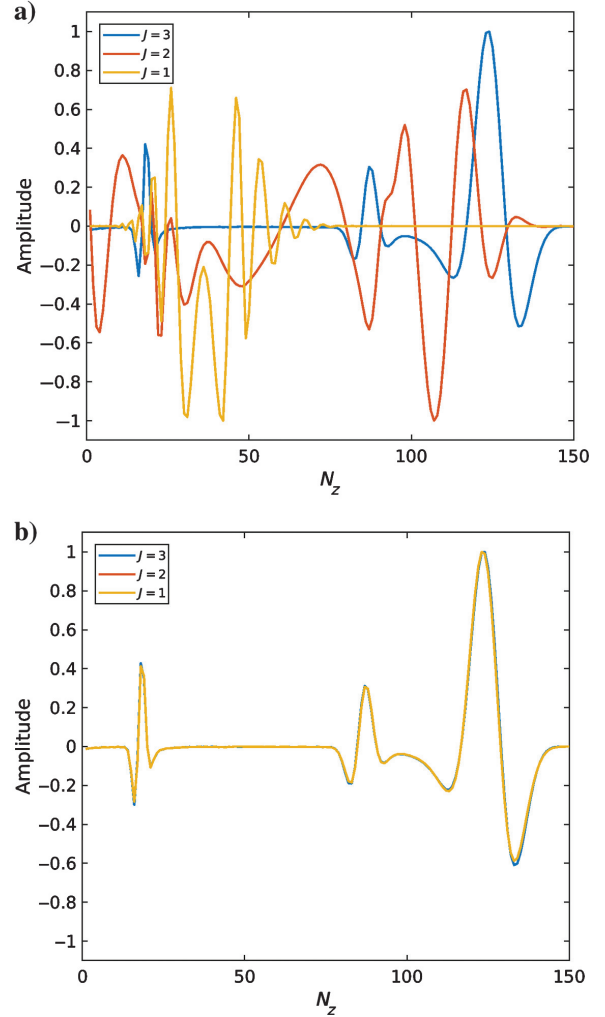


Figure 11. Numerical waveforms from simulations using the first three polynomial terms, after fixing  $S = 0.4$  in the highest velocity medium for (a) REM and (b) LWM.

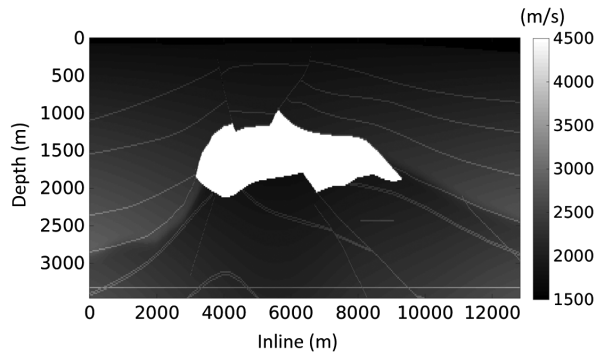


Figure 12. The slice of the SEG/EAGE Salt Model used for the numerical experiment at inline = 6800 m.



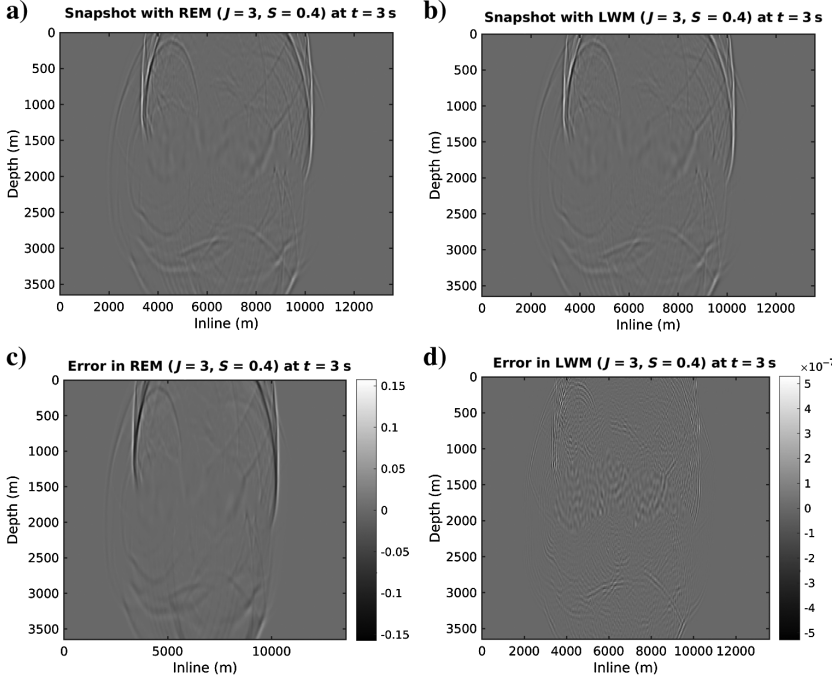


Figure 13. Numerical waveforms obtained with (a) REM and (b) LWM and the corresponding error for (c) REM and (d) LWM fixing  $S = 0.4$  in the highest velocity medium.

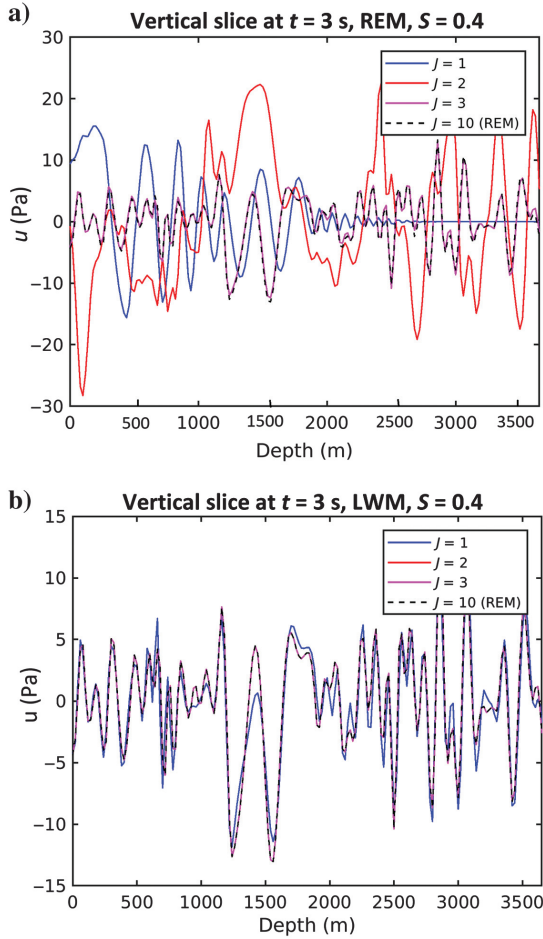


Figure 14. Numerical waveforms from simulations at the vertical direction below the source for different  $J$  values (a) for REM and (b) for LWM.

(see Figure 12) and locate a source at  $(x_s, z_s) = (6800, 0)$  m. The source has maximum frequency content at 37.5 Hz, and the grid spacing enforces 2 ppw. We use REM and LWM time integration with a fixed  $S = 0.4$  and  $J$  ranging from one to three. As a reference, we run a simulation with REM at  $J = 10$  and same  $S$  value.

In Figure 13a–13d, we show the results for  $J = 3$  obtained with both methods at time  $t = 3s$  together with the relative error of the difference between the reference signal and the results at different  $J$ . Although the snapshots of Figure 13a and 13b seem similar, we observe in Figure 13c and 13d maximum differences of 15% with REM, whereas, at the same computational cost, LWM attains errors at least five orders of magnitude smaller. In the same way, we see that REM at  $J = 10$  and LWM at  $J = 3$ , for a fixed  $S = 0.4$ , obtain very similar results, within  $5 \cdot 10^{-7}$  maximum differences, albeit REMs result required more than three times the computational cost.

In Figure 14a and 14b, we see details of the wavefields at the same instant, along the vertical direction coincident with the source location. It is clear that low  $J$  values have a significant impact on the accuracy of REM, whereas LWM is accurate for  $J \geq 1$ . This result further exposes the advantages of LWM with respect to REM for heterogeneous computations. Once again, the recommendation for spectral convergence and accuracy leads to a minimum  $J = 2$  in REM, which is clearly insufficient to attain high-quality results in this test.

### CONCLUSION

We have developed a stability and dispersion analysis of REM and LWM for time integration coupled to PS differentiation in space for the wave equation. The expansion structure of both methods, based on applying the spatial operator for each expansion term in a finite series, makes the methods easy to compare to one another. Our analyses for homogeneous media extensively quantify the dependence of both numerical properties on the total number of

expansion terms  $J$  and the stability CFL number  $S$ . In the case of LWM, the stability limits exhibit a fluctuating increase for growing  $J$  and they are consistently lower than REM bounds when  $J \geq 5$ . In addition, we find the expansion set  $J = 1, 2, \dots, 10$ , enough to show relevant differences between both integration methods. On the useful  $S$  range, i.e., the temporal sampling fulfills the minimum Nyquist resolution, REM dispersion errors are mostly smaller and less sensitive to  $S$  than LWM errors. For a given  $J$  value, dispersion errors for both methods significantly decay as  $S$  reduces. When comparing costs for both methods to attain a given accuracy, REM is typically slightly better than LWM. Nevertheless, and relevant to geophysical problems, in heterogeneous media REMs accuracy diminishes severely. This is related to the value  $R$ , used to scale the Chebyshev expansion terms in the whole computational domain, and given by the highest velocity in the model. As a consequence, in heterogeneous cases, LWM is much more accurate than REM for the same  $J$  value. Particularly important is finding that LWM can attain very accurate results in practical applications, such as the SEG/EAGE Salt Model case, with only one or two polynomial terms. Finally, we remark that both high-order methods are more efficient than PS with second-order FD time integration by a factor higher than three in low-accuracy thresholds, and even more at higher accuracy thresholds.

## ACKNOWLEDGMENTS

We would like to thank E. Koene for helpful comments in the review process. C. Spa has received funding from the Chilean Agency CONICYT under the project FONDECYT 11140212, whereas O. Rojas and J. de la Puente have received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 777778 MATHROCKS. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the ChEESE project, grant agreement No. 823844. We also acknowledge funding from the Spanish Ministry Project Geofísica de Altas Prestaciones TIN2016-80957-P.

## DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

## REFERENCES

- Amundsen, L., and Ø. Pedersen, 2017, Time step n-tupling for wave equations: *Geophysics*, **82**, no. 6, T249–T254, doi: [10.1190/geo2017-0377.1](https://doi.org/10.1190/geo2017-0377.1).
- Blanch, J., and J. Robertsson, 1997, A modified Lax-Wendroff correction for wave propagation in media described by Zener elements: *Geophysical Journal International*, **131**, 381–386, doi: [10.1111/j.1365-246X.1997.tb01229.x](https://doi.org/10.1111/j.1365-246X.1997.tb01229.x).
- Bohlen, T., and F. Wittkamp, 2016, Three-dimensional viscoelastic time-domain finite-difference seismic modelling using the staggered Adams-Bashforth time integrator: *Geophysical Journal International*, **204**, 1781–1788, doi: [10.1093/gji/ggv546](https://doi.org/10.1093/gji/ggv546).
- Chen, J., 2009, Lax-Wendroff and Nyström methods for seismic modelling: *Geophysical Prospecting*, **57**, 931–941, doi: [10.1111/j.1365-2478.2009.00802.x](https://doi.org/10.1111/j.1365-2478.2009.00802.x).
- Chu, C., P. Stoffa, and R. Seif, 2009, 3D elastic wave modeling using modified high-order time stepping schemes with improved stability conditions: 79th Annual International Meeting, SEG, Expanded Abstracts, 2662–2666, doi: [10.1190/1.3255400](https://doi.org/10.1190/1.3255400).
- Cruse, E., 1990, High-order (space and time) finite-difference modeling of the elastic wave equation: 60th Annual International Meeting, SEG, Expanded Abstracts, 987–991, doi: [10.1190/1.1890407](https://doi.org/10.1190/1.1890407).
- Dablain, M., 1986, The application of high-order differencing to the scalar wave equation: *Geophysics*, **51**, 54–66, doi: [10.1190/1.1442040](https://doi.org/10.1190/1.1442040).
- Fomberg, B., 1998, *A practical guide to pseudospectral methods*. Cambridge monographs on applied and computational mathematics: Cambridge University Press.
- Gazdag, J., 1981, Modeling of the acoustic wave equation with transforms methods: *Geophysics*, **46**, 854–859, doi: [10.1190/1.1441223](https://doi.org/10.1190/1.1441223).
- Ghrist, M., B. Fomberg, and T. Driscoll, 2000, Staggered time integrators for wave equations: *SIAM Journal on Numerical Analysis*, **38**, 718–741, doi: [10.1137/S0036142999351777](https://doi.org/10.1137/S0036142999351777).
- Koene, E. F. M., J. O. A. Robertsson, F. Broggini, and F. Andersson, 2017, Eliminating time dispersion from seismic wave modeling: *Geophysical Journal International*, **213**, 169–180, doi: [10.1093/gji/ggx563](https://doi.org/10.1093/gji/ggx563).
- Kosloff, D., and E. Baysal, 1982, Forward modeling by a Fourier method: *Geophysics*, **47**, 1402–1412, doi: [10.1190/1.1441288](https://doi.org/10.1190/1.1441288).
- Kosloff, D., A. Filho, E. Tessmer, and A. Behle, 1989, Numerical solution of the acoustic and elastic wave equations by a new rapid expansion method: *Geophysical Prospecting*, **37**, 383–394, doi: [10.1111/j.1365-2478.1989.tb02212.x](https://doi.org/10.1111/j.1365-2478.1989.tb02212.x).
- Kosloff, D., M. Reshef, and D. Loewenthal, 1984, Elastic wave calculations by the Fourier method: *Bulletin of the Seismological Society of America*, **74**, 875–891.
- Kristeková, M., J. Kristek, P. Moczo, and S. Day, 2006, Misfit criteria for quantitative comparison of seismograms: *Bulletin of the Seismological Society of America*, **96**, 1836–1850, doi: [10.1785/0120060012](https://doi.org/10.1785/0120060012).
- Moczo, P., J. Kristek, and M. Galis, 2014, *The finite-difference modelling of earthquake motions: Waves and ruptures*: Cambridge University Press.
- Pestana, R., and P. Stoffa, 2010, Time evolution of the wave equation using rapid expansion method: *Geophysics*, **75**, no. 4, T121–T131, doi: [10.1190/1.3449091](https://doi.org/10.1190/1.3449091).
- Reshef, M., D. Kosloff, M. Edwards, and C. Hsiung, 1988, Three-dimensional elastic modeling by the Fourier method: *Geophysics*, **53**, 1184–1193, doi: [10.1190/1.1442558](https://doi.org/10.1190/1.1442558).
- Rojas, O., C. Spa, and J. de la Puente, 2017, High-order leapfrog and rapid expansion time integrations on staggered finite difference wave simulations: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201700521](https://doi.org/10.3997/2214-4609.201700521).
- Soubaras, R., and Y. Zhang, 2008, Two-step explicit marching method for reverse time migration: 78th Annual International Meeting, SEG, Expanded Abstracts, 2272–2276, doi: [10.1190/1.3059337](https://doi.org/10.1190/1.3059337).
- Spa, C., J. Escolano, A. Garriga, and T. Mateos, 2009, Methodology for studying the numerical speed of sound in finite differences schemes: *Acta Acustica United with Acustica*, **95**, 690–695, doi: [10.3813/AAA.918197](https://doi.org/10.3813/AAA.918197).
- Stoffa, P., and R. Pestana, 2009, Numerical solution of the acoustic wave equation by the rapid expansion method (REM) — A one step time evolution algorithm: 79th Annual International Meeting, SEG, Expanded Abstracts, 2672–2676, doi: [10.1190/1.3255402](https://doi.org/10.1190/1.3255402).
- Stork, C., 2013, Eliminating nearly all dispersion error from FD modeling and RTM with minimal cost increase: 75th Annual International Conference and Exhibition Incorporating SPE EUROPEC, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20130478](https://doi.org/10.3997/2214-4609.20130478).
- Tal-Ezer, H., 1986, Spectral methods in time for hyperbolic equations: *SIAM Journal on Numerical Analysis*, **23**, 11–26, doi: [10.1137/0723002](https://doi.org/10.1137/0723002).
- Tal-Ezer, H., D. Kosloff, and Z. Koren, 1987, An accurate scheme for seismic forward modelling: *Geophysical Prospecting*, **35**, 479–490, doi: [10.1111/j.1365-2478.1987.tb00830.x](https://doi.org/10.1111/j.1365-2478.1987.tb00830.x).
- Tessmer, E., 2011, Using the rapid expansion method for accurate time-stepping in modeling and reverse-time migration: *Geophysics*, **76**, no. 4, T177–T185, doi: [10.1190/1.3587217](https://doi.org/10.1190/1.3587217).
- Zhang, W., Z. Zhang, and X. Chen, 2012, Three-dimensional elastic wave numerical modelling in the presence of surface topography by a collocated-grid finite-difference method on curvilinear grids: *Geophysical Journal International*, **190**, 358–378, doi: [10.1111/j.1365-246X.2012.05472.x](https://doi.org/10.1111/j.1365-246X.2012.05472.x).
- Zhang, W., Z. Zhang, and X. Chen, 2014, Three-dimensional curved grid finite-difference modelling for non-planar rupture dynamics: *Geophysical Journal International*, **199**, 860–879, doi: [10.1093/gji/ggu308](https://doi.org/10.1093/gji/ggu308).