

# A fourth-order optimal finite difference scheme for the Helmholtz equation with PML

Hatef Dastour\*, Wenyuan Liao

Department of Mathematics & Statistics, University of Calgary, AB, T2N 1N4, Canada



## ARTICLE INFO

### Article history:

Received 5 July 2018

Received in revised form 25 February 2019

Accepted 9 May 2019

Available online 19 August 2019

### Keywords:

Helmholtz equation

PML

Optimal finite difference scheme

Numerical dispersion

## ABSTRACT

In this paper, 17-point and 25-point finite difference (FD) schemes for the Helmholtz equation with perfectly matched layer (PML) in the two-dimensional domain are presented. It is shown that the 17-point FD scheme is inconsistent in the presence of PML; however, the 25-point FD scheme is pointwise consistent. An error analysis for the numerical approximation of the exact wavenumber is also presented. We present the global and refined 25-point finite difference schemes based on minimizing the numerical dispersion. Numerical experiments are given to illustrate the improvement of the accuracy and the reduction of the numerical dispersion.

© 2019 Elsevier Ltd. All rights reserved.

## 1. Introduction

Full-waveform inversion (FWI) is a challenging data-fitting procedure based on full-wavefield modeling to extract quantitative information from seismograms [1]. A key ingredient of FWI is an efficient forward-modeling engine. In particular, in the frequency domain, solving the Helmholtz equation numerically with high wavenumbers is still a challenging task in the field of computational mathematics [2–7]. This is mainly due to the fact that the accuracy of the numerical results usually deteriorates as the wavenumber  $k$  increases. This phenomenon in which the solution of the Helmholtz equation oscillates severely is the pollution effect of high wavenumbers [5,8,9]. In particular, for two and three dimensional Helmholtz equations, the pollution effect of high wavenumbers is almost inevitable [5,8]. Due to the pollution effect of high wavenumbers, the wavenumber of the numerical solution is different from the one of the exact solution, which is known as numerical dispersion [10]. Since the numerical dispersion is closely related to the pollution effect, minimizing the numerical dispersion will reduce the pollution effect [2–4].

Perfectly matched layer (PML) was introduced by Bérenger in 1994 [11] and is used to eliminate artificial reflection near the boundary. PML technique basically introduces an artificial layer with an attenuation parameter around the interior area (the domain of interest). As a result of applying PML, one can be ensured that almost no reflection, in theory, at the interface between the interior medium and the artificial absorbing medium (for more details about PML see [11–13]).

Finite-difference frequency-domain (FDFD) modeling for the generation of synthetic seismograms and cross-hole tomography has been an active field of research since the 1980s [4]. Pratt and Worthington [14] developed the classical 5-point finite difference scheme which requires ten gridpoints per wavelength. However, this can lead to a linear system with a huge and ill-conditioned matrix for large wavenumbers. Direct methods might fail to perform well for solving the linear system associated with these huge and ill-conditioned matrices. Usually, iterative methods with pre-conditioners are considered to solve these linear systems [15–17].

\* Corresponding author.

E-mail address: [hatef.dastour@ucalgary.ca](mailto:hatef.dastour@ucalgary.ca) (H. Dastour).

In 1996, Jo et al. presented the rotated 9-point finite difference method (FDM) for the Helmholtz equation which consists of linearly combining the two discretizations of the second derivative operator on the classical Cartesian coordinate system and the 45° rotated system [2]. In 1998, Shin and Sohn extended the idea of the rotated 9-point scheme to the 25-point formula, and they obtained a group of optimal parameters by the singular-value decomposition method [4]. Although the 25-point formula reduces the number of gridpoints per wavelength to 2, the resulting matrix’s bandwidth is much wider than that of the 9-point scheme. The rotated 9-point FDM was followed by another 9-point FDM for the Helmholtz equation which is consistent with PML [3]. The authors also proposed refined choice strategies for choosing optimal parameters of the scheme based on minimizing the numerical dispersion. This idea later was extended to the optimal 27-point finite difference scheme for the 3D Helmholtz equation with PML [10]. The rotated 9-point scheme was also extended to a generalized optimal 9-point scheme for frequency-domain scalar wave equation [6]. Moreover, in 2017, Cheng et al. [18] presented a new dispersion minimizing the finite difference in which the combination weights are determined by minimizing the numerical dispersion with a flexible selection strategy.

Higher-order finite difference schemes have been developed by many authors to improve numerical accuracy. In particular, Harari and Turkel presented various fourth-order methods for time-harmonic wave propagation [19]. Their method depends on the wave’s propagation angle. Singer and Turkel [20] presented a fourth-order compact finite difference scheme which depends on uniform grids for the 2D Helmholtz equation with constant wavenumbers. Moreover, Wu presented a dispersion minimizing compact FD scheme [5] and proposed a refined choice strategy based on minimizing the numerical dispersion for choosing weight parameters. Sixth-order compact FD schemes for the 2D and 3D Helmholtz equation with constant coefficients were discussed in [21]. There have been some researches regarding sixth-order FD schemes for the 2D and 3D Helmholtz equation with variable coefficients as well and the interested readers are referred to [21,22].

However, many of these higher-order schemes require the source term to be smooth enough to obtain higher-order accuracy, and this is not always the case in many practical problems. We propose a 25-point FDM which is non-compact and does not need this requirement.

The rest of this paper is organized as follows. In Section 2, an extension of the rotated 9-point FDM to a 17-point FDM is investigated. From [3], we know that the rotated 9-point FDM is not pointwise consistent with the Helmholtz equation with PML. In Section 2, it will be shown that our extension of the rotated 9-point FDM to the 17-point FDM is also not pointwise consistent with the Helmholtz equation with PML. This section is followed by Section 3 where we present an optimal 25-point fourth-order FDM. For this method, we use the global and refined choice strategies for choosing optimal parameters [3], and introduce the global and refined 25-point schemes. Moreover, the consistency of the 25-point FDM with the Helmholtz equation with PML and the fourth-order accuracy of the 25-point FDM are proved in Section 3. For the 25-point schemes, we then analyze the error between the numerical wavenumber and the exact wavenumber and present the global and refined 25-point finite difference schemes based on minimizing the numerical dispersion. In Section 4, numerical experiments are given to demonstrate the efficiency of the scheme. We show that the scheme proposed in Section 3 improves the accuracy and reduces the numerical dispersion significantly. Finally, Section 5 contains the conclusions of this paper.

**2. The 17-point finite difference scheme for the Helmholtz equation with PML**

In this section, we investigate an extension of the rotated 9-point finite difference scheme for the Helmholtz equation with PML [2,3] to a fourth-order method using 17 points. Let us start with a brief review of the Helmholtz equation with PML [12,13]. Consider the 2D Helmholtz equation

$$\Delta p + k^2 p = g, \tag{1}$$

where  $\Delta = \partial^2/\partial x^2 + \partial^2/\partial z^2$  is the Laplacian,  $k = 2\pi f/v$  is the wavenumber in which  $f$  and  $v$  represent the frequency and the velocity respectively,  $p$  is the Fourier component of the wavefield pressure, and  $g$  is the Fourier transform of the source function.

Applying the PML technique to truncate the infinite domain into a bounded rectangular domain and taking the same notations as in [3] lead to the equation

$$\frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) + C k^2 p = \tilde{g}, \tag{2}$$

where  $A = \frac{s_x}{s_x}$ ,  $B = \frac{s_z}{s_z}$ , and  $C = s_x s_z$  in which  $s_x = 1 - i \frac{\sigma_x}{\omega}$ ,  $s_z = 1 - i \frac{\sigma_z}{\omega}$ , and  $\omega = 2\pi f$  denotes the angular frequency, and  $\tilde{g} = \begin{cases} 0, & \text{inside PML,} \\ g, & \text{outside PML.} \end{cases}$

Moreover,  $\sigma_x$  and  $\sigma_z$  are usually chosen as a differentiable function only depending on the variable  $x$  and  $z$  respectively. One may consider defining them as follows,

$$\sigma_x = \begin{cases} 2\pi a_{ofM} \left( \frac{l_x}{L_{PML}} \right)^2, & \text{inside PML,} \\ 0, & \text{outside PML,} \end{cases} \quad \sigma_z = \begin{cases} 2\pi a_{ofM} \left( \frac{l_z}{L_{PML}} \right)^2, & \text{inside PML,} \\ 0, & \text{outside PML,} \end{cases} \tag{3}$$

where  $f_M$  is the peak frequency of the source,  $L_{PML}$  is the thickness of PML,  $l_x$  and  $l_z$  are the distance from the point  $(x, z)$  inside PML to the interface between the interior region and PML region. Furthermore,  $a_0$  is a constant, and we choose  $a_0 = 1.79$  according to the paper [23].

We next investigate an extension of the rotated 9-point finite difference scheme [2,3] to a 17-point fourth-order scheme. This scheme is referred as the 17-point scheme in this article.

Consider the network of grid points  $(x_m, z_n)$ , where  $x_m = x_0 + mh$  and  $z_n = z_0 + nh$  for  $m, n = 0, 1, 2, \dots$ . Let  $p_{m,n} = p|_{x=x_m, z=z_n}$  and  $k_{m,n} = k|_{x=x_m, z=z_n}$  represent the pressure of the wavefield and the wavenumber at the location  $(x_m, z_n)$ , respectively. The key idea of the 17-point scheme is to approximate  $\Delta p$  by linearly combining the two discretizations  $\Delta_{h,0^\circ}$  and  $\Delta_{h,45^\circ}$ . Let

$$\Delta_h p|_{x=x_m, z=z_n} = a\Delta_{h,0^\circ} p|_{x=x_m, z=z_n} + (1-a)\Delta_{h,45^\circ} p|_{x=x_m, z=z_n}, \tag{4a}$$

$$\begin{aligned} \Delta_{h,0^\circ} p|_{x=x_m, z=z_n} &= \frac{-p_{m,n-2} + 16p_{m,n-1} - 30p_{m,n} + 16p_{m,n+1} - p_{m,n+2}}{12h^2} \\ &+ \frac{-p_{m-2,n} + 16p_{m-1,n} - 30p_{m,n} + 16p_{m+1,n} - p_{m+2,n}}{12h^2}, \end{aligned} \tag{4b}$$

$$\begin{aligned} \Delta_{h,45^\circ} p|_{x=x_m, z=z_n} &= \frac{-p_{m-2,n-2} + 16p_{m-1,n-1} - 30p_{m,n} + 16p_{m+1,n+1} - p_{m+2,n+2}}{24h^2} \\ &+ \frac{-p_{m-2,n+2} + 16p_{m-1,n+2} - 30p_{m,n} + 16p_{m+1,n-1} - p_{m+2,n-2}}{24h^2}, \end{aligned} \tag{4c}$$

where  $a \in (0, 1]$  is a parameter to be determined.

We next establish a fourth-order approximation of  $k^2 p$ . Let

$$\begin{aligned} I(k^2 p)|_{x=x_m, z=z_n} &= c(k^2 p)|_{x=x_m, z=z_n} + dI_{h,0^\circ}(k^2 p)|_{x=x_m, z=z_n} \\ &+ eI_{h,45^\circ}(k^2 p)|_{x=x_m, z=z_n}. \end{aligned} \tag{5}$$

where  $c, d$  and  $e$  are parameters satisfying  $c + d + e = 1$ , and

$$\begin{aligned} I_{h,0^\circ}(k^2 p)|_{x=x_m, z=z_n} &= -\frac{1}{12}k_{m-2,n}^2 p_{m-2,n} + \frac{1}{3}k_{m-1,n}^2 p_{m-1,n} + \frac{1}{3}k_{m+1,n}^2 p_{m+1,n} \\ &- \frac{1}{12}k_{m+2,n}^2 p_{m+2,n} - \frac{1}{12}k_{m,n-2}^2 p_{m,n-2} + \frac{1}{3}k_{m,n-1}^2 p_{m,n-1} + \frac{1}{3}k_{m,n+1}^2 p_{m,n+1} \\ &- \frac{1}{12}k_{m,n+2}^2 p_{m,n+2}, \end{aligned} \tag{6a}$$

$$\begin{aligned} I_{h,45^\circ}(k^2 p)|_{x=x_m, z=z_n} &= -\frac{1}{12}k_{m-2,n-2}^2 p_{m-2,n-2} + \frac{1}{3}k_{m-1,n-1}^2 p_{m-1,n-1} \\ &+ \frac{1}{3}k_{m+1,n+1}^2 p_{m+1,n+1} - \frac{1}{12}k_{m+2,n+2}^2 p_{m+2,n+2} - \frac{1}{12}k_{m-2,n+2}^2 p_{m-2,n+2} \\ &+ \frac{1}{3}k_{m-1,n+1}^2 p_{m-1,n+1} + \frac{1}{3}k_{m+1,n-1}^2 p_{m+1,n-1} - \frac{1}{12}k_{m+2,n-2}^2 p_{m+2,n-2}. \end{aligned} \tag{6b}$$

Note that coefficients in Eqs. (6a) and (6b) are considered in a way that  $I_{h,0^\circ}$  and  $I_{h,45^\circ}$  approximate  $k^2 p$  with fourth-order accuracy.

As a result, the following the 17-point FD approximation for the Helmholtz (1) can be obtained,

$$\Delta_h p|_{x=x_m, z=z_n} + I(k^2 p)|_{x=x_m, z=z_n} = 0. \tag{7}$$

We need to discretize  $A(x, z)$ ,  $B(x, z)$  and  $C(x, z)$  in order to develop a 17-point scheme for solving the Helmholtz-PML equation (2). We have

$$\begin{cases} A_{m+\frac{j}{2}, n+\frac{l}{2}} = A(x_m + \frac{j}{2}h, z_n + \frac{l}{2}h), \\ B_{m+\frac{j}{2}, n+\frac{l}{2}} = B(x_m + \frac{j}{2}h, z_n + \frac{l}{2}h), \quad j, l \in \{-3, -1, 0, 1, 3\}. \\ C_{m,n} = C(x_m, z_n), \end{cases} \tag{8}$$

In addition, we define

$$\mathcal{L}_h p|_{x=x_m, z=z_n} = a\mathcal{L}_{h,0^\circ} p|_{x=x_m, z=z_n} + (1-a)\mathcal{L}_{h,45^\circ} p|_{x=x_m, z=z_n}, \tag{9}$$

where  $a \in (0, 1]$ , and

$$\begin{aligned} \mathcal{L}_{h,0^\circ} p|_{x=x_m, z=z_n} &= \frac{1}{h^2} \left[ \frac{1}{24}A_{m-\frac{3}{2}, n} \left( -\frac{11}{12}p_{m-2,n} + \frac{17}{24}p_{m-1,n} + \frac{3}{8}p_{m,n} - \frac{5}{24}p_{m+1,n} \right. \right. \\ &\left. \left. + \frac{1}{24}p_{m+2,n} \right) - \frac{9}{8}A_{m-\frac{1}{2}, n} \left( \frac{1}{24}p_{m-2,n} - \frac{9}{8}p_{m-1,n} + \frac{9}{8}p_{m,n} - \frac{1}{24}p_{m+1,n} \right) \right] \end{aligned}$$

$$\begin{aligned}
 & + \frac{9}{8}A_{m+\frac{1}{2},n} \left( \frac{1}{24}p_{m-1,n} - \frac{9}{8}p_{m,n} + \frac{9}{8}p_{m+1,n} - \frac{1}{24}p_{m+2,n} \right) \\
 & - \frac{1}{24}A_{m+\frac{3}{2},n} \left( -\frac{1}{24}p_{m-2,n} + \frac{5}{24}p_{m-1,n} - \frac{3}{8}p_{m,n} - \frac{17}{24}p_{m+1,n} + \frac{11}{12}p_{m+2,n} \right) \Big] \\
 & + \frac{1}{h^2} \left[ \frac{1}{24}B_{m,n-\frac{3}{2}} \left( -\frac{11}{12}p_{m,n-2} + \frac{17}{24}p_{m,n-1} + \frac{3}{8}p_{m,n} - \frac{5}{24}p_{m,n+1} + \frac{1}{24}p_{m,n+2} \right) \right. \\
 & - \frac{9}{8}B_{m,n-\frac{1}{2}} \left( \frac{1}{24}p_{m,n-2} - \frac{9}{8}p_{m,n-1} + \frac{9}{8}p_{m,n} - \frac{1}{24}p_{m,n+1} \right) \\
 & + \frac{9}{8}B_{m,n+\frac{1}{2}} \left( \frac{1}{24}p_{m,n-1} - \frac{9}{8}p_{m,n} + \frac{9}{8}p_{m,n+1} - \frac{1}{24}p_{m,n+2} \right) \\
 & \left. - \frac{1}{24}B_{m,n+\frac{3}{2}} \left( -\frac{1}{24}p_{m,n-2} + \frac{5}{24}p_{m,n-1} - \frac{3}{8}p_{m,n} - \frac{17}{24}p_{m,n+1} + \frac{11}{12}p_{m,n+2} \right) \right]. \tag{10}
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{L}_{h,45^\circ} p|_{x=x_m, z=z_n} & = \frac{1}{2h^2} \left[ \frac{1}{24}A_{m-\frac{3}{2},n+\frac{3}{2}} \left( -\frac{11}{12}p_{m-2,n+2} + \frac{17}{24}p_{m-1,n+1} + \frac{3}{8}p_{m,n} \right. \right. \\
 & - \frac{5}{24}p_{m+1,n-1} + \frac{1}{24}p_{m+2,n-2} \Big) - \frac{9}{8}A_{m-\frac{1}{2},n+\frac{1}{2}} \left( \frac{1}{24}p_{m-2,n+2} - \frac{9}{8}p_{m-1,n+1} + \frac{9}{8}p_{m,n} \right. \\
 & - \frac{1}{24}p_{m+1,n-1} \Big) + \frac{9}{8}A_{m+\frac{1}{2},n-\frac{1}{2}} \left( \frac{1}{24}p_{m-1,n+1} - \frac{9}{8}p_{m,n} + \frac{9}{8}p_{m+1,n-1} - \frac{1}{24}p_{m+2,n-2} \right) \\
 & - \frac{1}{24}A_{m+\frac{3}{2},n-\frac{3}{2}} \left( -\frac{1}{24}p_{m-2,n+2} + \frac{5}{24}p_{m-1,n+1} - \frac{3}{8}p_{m,n} - \frac{17}{24}p_{m+1,n-1} + \frac{11}{12}p_{m+2,n-2} \right) \Big] \\
 & + \frac{1}{2h^2} \left[ \frac{1}{24}B_{m-\frac{3}{2},n-\frac{3}{2}} \left( -\frac{11}{12}p_{m,n-2} + \frac{17}{24}p_{m,n-1} + \frac{3}{8}p_{m,n} - \frac{5}{24}p_{m,n+1} + \frac{1}{24}p_{m,n+2} \right) \right. \\
 & - \frac{9}{8}B_{m-\frac{1}{2},n-\frac{1}{2}} \left( \frac{1}{24}p_{m,n-2} - \frac{9}{8}p_{m,n-1} + \frac{9}{8}p_{m,n} - \frac{1}{24}p_{m,n+1} \right) \\
 & + \frac{9}{8}B_{m+\frac{1}{2},n+\frac{1}{2}} \left( \frac{1}{24}p_{m,n-1} - \frac{9}{8}p_{m,n} + \frac{9}{8}p_{m,n+1} - \frac{1}{24}p_{m,n+2} \right) \\
 & \left. - \frac{1}{24}B_{m+\frac{3}{2},n+\frac{3}{2}} \left( -\frac{1}{24}p_{m,n-2} + \frac{5}{24}p_{m,n-1} - \frac{3}{8}p_{m,n} - \frac{17}{24}p_{m,n+1} + \frac{11}{12}p_{m,n+2} \right) \right]. \tag{11}
 \end{aligned}$$

Please refer to Appendix A to see how Eqs. (10) and (11) are derived.

Moreover,  $k^2 Cp|_{x=x_m, z=z_n}$ , in Eq. (2), can be approximated using Eq. (5)

$$\begin{aligned}
 I(k^2 Cp)|_{x=x_m, z=z_n} & = c(k^2 Cp)|_{x=x_m, z=z_n} + dI_{h,0^\circ}(k^2 Cp)|_{x=x_m, z=z_n} \\
 & + eI_{h,45^\circ}(k^2 Cp)|_{x=x_m, z=z_n}. \tag{12}
 \end{aligned}$$

Thereby, the following the 17-point FD approximation for the Helmholtz-PML equation (2) can be obtained,

$$\mathcal{L}_h p|_{x=x_m, z=z_n} + I(k^2 Cp)|_{x=x_m, z=z_n} = 0. \tag{13}$$

We recall the concept of consistency [3,24] to analyze the relationship between the finite difference approximation (13) and the Helmholtz-PML equation (2).

**Definition 2.1.** Let  $(x_m, z_n) = (x_0 + m\Delta x, z_0 + n\Delta z)$  for  $m, n = 0, 1, 2, \dots$ , and suppose that the partial differential equation under consideration is  $(\Delta + k^2)p = g$ , and the corresponding finite difference approximation is  $\mathcal{L}P_{m,n} = G_{m,n}$  where  $G_{m,n} = g(x_n, y_n)$ . The finite difference scheme  $\mathcal{L}P_{m,n} = G_{m,n}$  is pointwise consistent with the partial differential equation  $(\Delta + k^2)p = g$  at  $(x, z)$  if for any smooth function  $\phi(x, z)$ ,

$$\left\| \left( (\Delta + k^2)\phi - g \right) |_{x=x_m, z=z_n} - [\mathcal{L}\phi(x_m, z_n) - G_{m,n}] \right\| \rightarrow 0 \tag{14}$$

as  $\Delta x, \Delta z \rightarrow 0$ .

**Proposition 2.1.** The 17-point FD approximation (13) is not pointwise consistent with the Helmholtz-PML equation (2).

**Proof.** Assume that  $x_m \leq x < x_{m+1}$  and  $z_n \leq z < z_{n+1}$ . It follows from Eqs. (10)–(12) and the Taylor theorem that

$$\mathcal{L}_{h,0^\circ} p|_{x=x_m, z=z_n} = \frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) + \mu_1 h^4 + \mathcal{O}(h^6), \tag{15a}$$

$$\begin{aligned} \mathcal{L}_{h,45^\circ} p|_{x=x_m, z=z_n} &= \frac{1}{2} \left[ \left( \frac{\partial}{\partial x} - \frac{\partial}{\partial z} \right) \left( A \left( \frac{\partial}{\partial x} - \frac{\partial}{\partial z} \right) p \right) \right. \\ &\quad \left. + \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \right) \left( B \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \right) p \right) \right] + \mu_2 h^4 + \mathcal{O}(h^6), \end{aligned} \tag{15b}$$

$$I(k^2 Cp)|_{x=x_m, z=z_n} = k^2 Cp + \mu_3 h^4 + \mathcal{O}(h^6), \tag{15c}$$

where  $\mu_1, \mu_2$  and  $\mu_3$  are available in (B.1)–(B.3).

Combining Eqs. (15a)–(15c) yields that the left-hand side of the 17-point scheme (13) is equivalent to

$$\begin{aligned} a \left[ \frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) \right] + \frac{1-a}{2} \left[ \left( \frac{\partial}{\partial x} - \frac{\partial}{\partial z} \right) \left( A \left( \frac{\partial}{\partial x} - \frac{\partial}{\partial z} \right) p \right) \right. \\ \left. + \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \right) \left( B \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \right) p \right) \right] + k^2 Cu + \zeta h^4 + \mathcal{O}(h^6), \end{aligned} \tag{16}$$

where  $\zeta = a\mu_1 + (1-a)\mu_2 + \mu_3$ .

From the PML's formulation, we know that there exists some area satisfying  $e_z \equiv 1$  and  $e_x \neq 1$ . Hence, in such area there hold  $A = \frac{1}{e_x}, B = e_x$  and  $C = e_x$ . It follows from (13) that (16) can be written as

$$\begin{aligned} a \left[ \frac{\partial}{\partial x} \left( \frac{1}{e_x} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( e_x \frac{\partial p}{\partial z} \right) \right] + \frac{1-a}{2} \left[ \frac{\partial}{\partial x} \left( \left( \frac{1}{e_x} + e_x \right) \frac{\partial p}{\partial x} \right) \right. \\ \left. + \frac{\partial}{\partial z} \left( \left( \frac{1}{e_x} + e_x \right) \frac{\partial p}{\partial z} \right) + \left( e_x - \frac{1}{e_x} \right) \left( \frac{\partial^2 p}{\partial x \partial z} + \frac{\partial^2 p}{\partial z \partial x} \right) + \frac{\partial}{\partial x} \left( \left( \frac{1}{e_x} - e_x \right) \frac{\partial p}{\partial z} \right) \right] \\ + k^2 e_x u + \zeta h^4 + \mathcal{O}(h^6). \end{aligned} \tag{17}$$

As there exist the terms  $\left( \frac{\partial^2 p}{\partial x \partial z} + \frac{\partial^2 p}{\partial z \partial x} \right)$  and  $\frac{\partial p}{\partial z}$ , we have the conclusion of this proposition for the area which satisfy both  $e_z \equiv 1$  and  $e_x \neq 1$ . We can get similar results almost everywhere in PML. Thus, we come to the conclusion of this proposition.  $\square$

From [3], we knew that the rotated 9-point finite difference scheme is not pointwise consistent with the Helmholtz-PML equation. Proposition 2.1 indicates that our extension of the rotated 9-point finite difference to the 17-point FDM is also pointwise inconsistent. As the convergence of the finite difference scheme requires that the finite difference scheme should be consistent with the Helmholtz-PML equation, the 17-point FDM approximation for the Helmholtz-PML equation is not a recommended choice.

### 3. A consistent 25-point fourth-order finite difference scheme for the Helmholtz equation with PML

In this section, an optimal 25-point FDM which is fourth-order and consistent with the Helmholtz-PML equation is proposed. The error between the numerical wavenumber and the exact wavenumber is analyzed. We also present the global and refined 25-point finite difference schemes using the global and the refined rules [3]. Moreover, the 25-point FDM is generalized to the case that different step sizes are used for different variables.

We follow the approach of constructing finite difference scheme in [3,13] to construct a consistent fourth-order 25-point difference scheme for the Helmholtz equation with PML. This method can be extended to the case of the Helmholtz-PML equation in the three-dimensional domain. Let

$$\begin{aligned} \mathcal{L}_{h,x} p|_{(m, n+j)} &= \frac{1}{h^2} \left[ \frac{1}{24} A_{m-\frac{3}{2}, n+j} \left( -\frac{11}{12} p_{m-2, n+j} + \frac{17}{24} p_{m-1, n+j} + \frac{3}{8} p_{m, n+j} \right. \right. \\ &\quad - \frac{5}{24} p_{m+1, n+j} + \frac{1}{24} p_{m+2, n+j} \Big) - \frac{9}{8} A_{m-\frac{1}{2}, n+j} \left( \frac{1}{24} p_{m-2, n+j} - \frac{9}{8} p_{m-1, n+j} + \frac{9}{8} p_{m, n+j} \right. \\ &\quad \left. \left. - \frac{1}{24} p_{m+1, n+j} \right) + \frac{9}{8} A_{m+\frac{1}{2}, n+j} \left( \frac{1}{24} p_{m-1, n+j} - \frac{9}{8} p_{m0, n+j} + \frac{9}{8} p_{m+1, n+j} - \frac{1}{24} p_{m+2, n+j} \right) \right. \\ &\quad \left. - \frac{1}{24} A_{m+\frac{3}{2}, n+j} \left( -\frac{1}{24} p_{m-2, n+j} + \frac{5}{24} p_{m-1, n+j} - \frac{3}{8} p_{m, n+j} - \frac{17}{24} p_{m+1, n+j} \right. \right. \\ &\quad \left. \left. + \frac{11}{12} p_{m+2, n+j} \right) \right], \quad j \in \mathbb{Z}_3, \end{aligned} \tag{18a}$$

$$\begin{aligned} \mathcal{L}_{h,z} p|_{(m+j, n)} &= \frac{1}{h^2} \left[ \frac{1}{24} B_{m+j, n-\frac{3}{2}} \left( -\frac{11}{12} p_{m+j, n-2} + \frac{17}{24} p_{m+j, n-1} + \frac{3}{8} p_{m+j, n} \right. \right. \\ &\quad \left. \left. - \frac{5}{24} p_{m+j, n+1} + \frac{1}{24} p_{m+j, n+2} \right) - \frac{9}{8} B_{m+j, n-\frac{1}{2}} \left( \frac{1}{24} p_{m+j, n-2} - \frac{9}{8} p_{m+j, n-1} + \frac{9}{8} p_{m+j, n} \right) \right. \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{24}p_{m+j,n+1} \Big) + \frac{9}{8}B_{m+j,n+\frac{1}{2}} \left( \frac{1}{24}p_{m+j,n-1} - \frac{9}{8}p_{m+j,n0} + \frac{9}{8}p_{m+j,n+1} - \frac{1}{24}p_{m+j,n+2} \right) \\
 & -\frac{1}{24}B_{m+j,n+\frac{3}{2}} \left( -\frac{1}{24}p_{m+j,n-2} + \frac{5}{24}p_{m+j,n-1} - \frac{3}{8}p_{m+j,n} - \frac{17}{24}p_{m+j,n+1} \right. \\
 & \left. + \frac{11}{12}p_{m+j,n+2} \right) \Big], j \in \mathbb{Z}_3. \tag{18b}
 \end{aligned}$$

Define

$$\begin{aligned}
 \tilde{\mathcal{L}}_{h,x}p_{m,n} := & b\mathcal{L}_{h,x}p_{m,n} + (1-b) \left( -\frac{1}{6}\mathcal{L}_{h,x}p_{m,n-2} + \frac{2}{3}\mathcal{L}_{h,x}p_{m,n-1} + \frac{2}{3}\mathcal{L}_{h,x}p_{m,n+1} \right. \\
 & \left. -\frac{1}{6}\mathcal{L}_{h,x}p_{m,n+2} \right), \tag{19a}
 \end{aligned}$$

$$\begin{aligned}
 \tilde{\mathcal{L}}_{h,z}p_{m,n} := & b\mathcal{L}_{h,z}p_{m,n} + (1-b) \left( -\frac{1}{6}\mathcal{L}_{h,z}p_{m-2,n} + \frac{2}{3}\mathcal{L}_{h,z}p_{m-1,n} + \frac{2}{3}\mathcal{L}_{h,z}p_{m+1,n} \right. \\
 & \left. -\frac{1}{6}\mathcal{L}_{h,z}p_{m+2,n} \right). \tag{19b}
 \end{aligned}$$

where  $b \in (0, 1]$ .

Let  $\tilde{\mathcal{L}}_h = \tilde{\mathcal{L}}_{h,x} + \tilde{\mathcal{L}}_{h,z}$  and approximate the first two terms of the left-hand side of (2) as

$$\frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) \approx \tilde{\mathcal{L}}_h p_{m,n}. \tag{20}$$

We obtain the following fourth-order 25-point finite difference approximation for the Helmholtz-PML equation (2),

$$\tilde{\mathcal{L}}_h p|_{x=x_m, z=z_n} + I(k^2 Cp)|_{x=x_m, z=z_n} = 0. \tag{21}$$

The next proposition presents the convergence analysis for the 25-point finite difference scheme (21).

**Proposition 3.1.** *If  $b \in (0, 1]$  and  $c + d + e = 1$ , then the 25-point finite difference approximation (21) is pointwise consistent with the Helmholtz-PML equation (2) and is a fourth-order scheme.*

**Proof.** Assume that  $x_m \leq x < x_{m+1}$  and  $z_n \leq z < z_{n+1}$ . It follows from the Taylor expansion that

$$\tilde{\mathcal{L}}_{h,x} \Big|_{x=x_m, z=z_n} = \frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) + \xi_1 h^4 + \mathcal{O}(h^6), \tag{22a}$$

$$\tilde{\mathcal{L}}_{h,z} \Big|_{x=x_m, z=z_n} = \frac{\partial}{\partial z} \left( A \frac{\partial p}{\partial z} \right) + \xi_2 h^4 + \mathcal{O}(h^6) \tag{22b}$$

where  $\xi_1$  and  $\xi_2$  are available in (B.4)–(B.5).

Let  $\zeta = \xi_1 + \xi_2 + \mu_3$ . It follows from (22) and (15c) that the left-hand side of the 25-point finite difference approximation (21) is equivalent to:

$$\begin{aligned}
 \tilde{\mathcal{L}}_h p|_{x=x_m, z=z_n} + I(k^2 Cp)|_{x=x_m, z=z_n} = & \frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) \\
 & + Ck^2 p + \zeta h^4 + \mathcal{O}(h^6). \tag{23}
 \end{aligned}$$

The results of this proposition can be concluded from (23) and (2). Eq. (23) also shows that the 25-point finite difference approximation (21) is a fourth-order scheme for arbitrary constants  $b, c, d$  and  $e$  under the conditions  $b \in (0, 1]$  and  $c + d + e = 1$ . Since  $\zeta$  depends on  $k, A$  and  $B$ , as shown in (B.3)–(B.5), the convergence order and accuracy may be affected by the values of  $k$  and its derivatives, especially when  $k$  is large. □

### 3.1. Global and refined 25p finite difference schemes

Convergence order alone is not enough to measure the property of a finite difference scheme since the solution of the Helmholtz equation is oscillating seriously for large wavenumbers and the accuracy of the numerical solution deteriorates with increasing wavenumber  $k$ . This phenomenon is known as the pollution effect (please see [3,8,9] for more details). As a result, one may consider minimizing the numerical dispersion by minimizing the error between the numerical wavenumber and the exact wavenumber. The optimal difference scheme for the Helmholtz-PML equation is a finite difference scheme that the parameters are chosen such that the scheme has minimal numerical dispersion in the interior area, and it has optimal convergence order [3,25].

The global and the refined choice parameter choice strategies [3] are used here to optimize the 25-point scheme (21). Let  $P(x, z) = \exp(-ik(x \cos \theta + z \sin \theta))$ , where  $\theta$  is the propagation angle from the z-axis and the wavenumber  $k$  here is assumed to be a positive constant.

In the interior area,  $A = B = C = 1$ , thus replacing  $p_{m+i,n+j}$  with  $P_{m+i,n+j}$  ( $i, j \in \mathbb{Z}_3$ ) in the formula (21) gives

$$\begin{aligned} & \hat{T}_1 P_{m-2,n-2} + \hat{T}_2 P_{m-1,n-2} + \hat{T}_3 P_{m,n-2} + \hat{T}_2 P_{m+1,n-2} + \hat{T}_1 P_{m+2,n-2} \\ & \hat{T}_2 P_{m-2,n-1} + \hat{T}_4 P_{m-1,n-1} + \hat{T}_5 P_{m,n-1} + \hat{T}_4 P_{m+1,n-1} + \hat{T}_2 P_{m+2,n-1} \\ & \hat{T}_3 P_{m-2,n} + \hat{T}_5 P_{m-1,n} + \hat{T}_6 P_{m,n} + \hat{T}_5 P_{m+1,n} + \hat{T}_3 P_{m+2,n} \\ & \hat{T}_2 P_{m-2,n+1} + \hat{T}_4 P_{m-1,n+1} + \hat{T}_5 P_{m,n+1} + \hat{T}_2 P_{m+1,n+1} + \hat{T}_2 P_{m+2,n+1} \\ & \hat{T}_1 P_{m-2,n+2} + \hat{T}_2 P_{m-1,n+2} + \hat{T}_3 P_{m,n+2} + \hat{T}_2 P_{m+1,n+2} + \hat{T}_1 P_{m+2,n+2} = 0, \end{aligned} \tag{24}$$

where

$$\begin{aligned} \hat{T}_1 &= \frac{1-b}{36h^2} - \frac{e}{12} k^2, & \hat{T}_2 &= \frac{5(b-1)}{18h^2}, & \hat{T}_3 &= \frac{5-6b}{12h^2} - \frac{d}{12} k^2, \\ \hat{T}_4 &= \frac{16(1-b)}{9h^2} + \frac{e}{3} k^2, & \hat{T}_5 &= \frac{9b-5}{3h^2} + \frac{d}{3} k^2, & \hat{T}_6 &= -\frac{5b}{h^2} + (1-d-e) k^2. \end{aligned} \tag{25}$$

Let  $\lambda = \frac{2\pi v}{\omega}$  and  $G = \frac{\lambda}{h}$  denote the wavelength and the number of gridpoints per wavelength, respectively. Moreover, denote

$$\begin{cases} P = \cos(k_x h) = \cos(kh \cos \theta) = \cos\left(\frac{2\pi}{G} \cos \theta\right), \\ Q = \cos(k_z h) = \cos(kh \sin \theta) = \cos\left(\frac{2\pi}{G} \sin \theta\right). \end{cases} \tag{26}$$

It follows from substituting  $P_{m,n} = \exp(-ik(x \cos \theta + z \sin \theta))$  into Eq. (24), and simplifying that

$$\begin{aligned} & (4(2P^2 - 1)(2Q^2 - 1)) \hat{T}_1 + (4P(2Q^2 - 1) + 4Q(2P^2 - 1)) \hat{T}_2 \\ & + (4P^2 + 4Q^2 - 4) \hat{T}_3 + (4PQ) \hat{T}_4 + (2P + 2Q) \hat{T}_5 + \hat{T}_6 = 0. \end{aligned} \tag{27}$$

Furthermore, let  $k_N$  represent the numerical wavenumber. It follows from replacing the variable  $k$  in the parameters  $\hat{T}_1, \hat{T}_2, \dots, \hat{T}_6$  with  $k_N$  in Eq. (27) that

$$\frac{k_N}{k} = \frac{1}{h} \sqrt{\frac{N}{D}}, \tag{28}$$

where

$$\begin{aligned} N &= -4(Q - 1)(P - 1)(PQ - 4Q - 4P + 7)b + 13(P^2 + Q^2) \\ &+ 4PQ(PQ - 5Q - 5P + 16) - 20(P + Q), \end{aligned} \tag{29a}$$

$$\begin{aligned} D &= 3(P^2 + Q^2 - 2(P + Q) + 2)d + 6(2PQ(PQ - 1) \\ &- (P^2 + Q^2) + 2)e - 9. \end{aligned} \tag{29b}$$

The next proposition presents the error between the numerical wavenumber  $k_N$  and the exact wavenumber  $k$  for the finite difference scheme (21).

**Proposition 3.2.** For the 25-point finite difference scheme (21), there holds

$$\begin{aligned} (k_N)^2 &= k^2 - \frac{1}{720} [(-15 \cos(4\theta) - 45)d + (60 \cos(4\theta) - 180)e \\ &+ 15b \cos(4\theta) - 12 \cos(4\theta) - 15b + 20] h^4 k^6 + \mathcal{O}(k^{10} h^8), \quad kh \rightarrow 0. \end{aligned} \tag{30}$$

**Proof.** Let  $f_1(\tau)$  and  $f_2(\tau)$  respectively denote the numerator and the denominator of the function under the squared root in Eq. (28), where  $\tau = kh$ . We have,

$$\begin{aligned} f_1(\tau) &= -4(Q(\tau) - 1)(P(\tau) - 1)(P(\tau)Q(\tau) - 4Q(\tau) - 4P(\tau) + 7)b \\ &+ 13(P(\tau)^2 + Q(\tau)^2) + 4P(\tau)Q(\tau)(P(\tau)Q(\tau) - 5Q(\tau) - 5P(\tau) + 16) \\ &- 20(P(\tau) + Q(\tau)), \end{aligned} \tag{31a}$$

$$\begin{aligned} f_2(\tau) &= 3(P(\tau)^2 + Q(\tau)^2 - 2(P(\tau) + Q(\tau)) + 2)d \\ &+ 6(2P(\tau)Q(\tau)(P(\tau)Q(\tau) - 1) - (P(\tau)^2 + Q(\tau)^2) + 2)e - 9. \end{aligned} \tag{31b}$$

Applying Taylor expansions to  $f_1(\tau)$  and  $\frac{1}{f_2(\tau)}$  at the point  $\tau = 0$ , we have,

$$f_1(\tau) = -9\tau^2 + \left( \frac{3b \cos(4\theta)}{16} - \frac{3 \cos(4\theta)}{20} - \frac{3b}{16} + \frac{1}{4} \right) \tau^6 + \mathcal{O}(\tau^8), \tag{32a}$$

$$\frac{1}{f_2(\tau)} = -\frac{1}{9} + \left( \frac{e \cos(4\theta)}{108} - \frac{e}{36} - \frac{d \cos(4\theta)}{432} - \frac{d}{144} \right) \tau^4 + \mathcal{O}(\tau^6). \tag{32b}$$

In addition, we have from (28)

$$(k_N h)^2 = \frac{f_1(\tau)}{f_2(\tau)}. \tag{33}$$

The results of this proposition can be concluded from (32) and (33). □

The above proposition indicates that  $k_N$  approximates  $k$  with fourth-order accuracy. Moreover, the term associated with  $k^6 h^4$  presents the pollution effect, which depends on the wavenumber  $k$ , the parameters of the finite difference formula and the wave’s propagation angle  $\theta$  from the  $z$ -axis.

Since  $h = \frac{2\pi}{Gk}$ , the relationship of the numerical wavenumber  $k_N$  and the exact wavenumber  $k$  can be presented as follows,

$$\frac{k_N}{k} = \frac{G}{2\pi} \sqrt{\frac{N}{D}}. \tag{34}$$

Moreover, the normalized numerical phase velocity and the normalized numerical group velocity can be found respectively as [2–4,26],

$$\frac{V_{ph}^N}{v} = \frac{G}{2\pi} \sqrt{\frac{N}{D}}, \tag{35a}$$

$$\frac{V_{gr}^N}{v} = \frac{v}{V_{ph}^N} \left[ \frac{\left( \frac{1}{h} \frac{\partial N}{\partial k} \right) D - N \left( \frac{1}{h} \frac{\partial D}{\partial k} \right)}{D^2} \right]. \tag{35b}$$

Let

$$J(b, d, e; G, \theta) = \frac{G}{2\pi} \sqrt{\frac{N}{D}} - 1 \tag{36}$$

for  $b \in (0, 1]$ , and  $d, e \in \mathbb{R}$ , and  $(G, \theta) \in I_G \times I_\theta$ , where  $I_G$  and  $I_\theta$  are two intervals. In general, one can choose  $I_\theta = [0, \frac{\pi}{2}]$  and  $I_G = [G_{\min}, G_{\max}] \subseteq [2, 400]$  (see [4]). We remark that the interval  $[0, \frac{\pi}{2}]$  can be replaced by  $[0, \frac{\pi}{4}]$  because of the symmetry, and  $G_{\min} \geq 2$  based on the Nyquist sampling limit (see [4]).

The optimal parameters  $b, d$  and  $e$  are chosen in a way that the numerical dispersion is minimized. In other words, minimizing the error between  $k_N$  and  $k$  is equivalent to minimizing the norm  $\|J(b, d, e; G, \theta)\|_{\infty, I_G \times I_\theta}$ . We use the global and the refined choice strategies (rules 3.4 and 3.8 from [3]) to minimize  $J(b, d, e; G, \theta)$ .

To implement the global and the refined choice strategies, we can solve

$$(b, d, e) = \arg \min \{ \|J(b, d, e; G, \theta)\|_{I_G \times I_\theta} : b \in (0, 1], d, e \in \mathbb{R} \}, \tag{37}$$

by using the least-squares method. Thereby, it follows from  $J(b, d, e; G, \theta) = 0$  that

$$\begin{aligned} 4G^2(Q-1)(P-1)(PQ-4Q-4P+7)b + 12\pi^2(P^2-2P+Q^2-2Q+2)d \\ + 24\pi^2(2P^2Q^2-P^2-2PQ-Q^2+2)e = [4PQ(PQ-5Q-5P+16) \\ + P(13P-20) + Q(13Q-20) - 14]G^2 + 36\pi^2 \end{aligned} \tag{38}$$

Note that  $P$  and  $Q$  are functions of  $G$  and  $\theta$ . We choose

$$\begin{cases} \theta = \theta_m = \frac{(m-1)\pi}{4(l-1)} \in I_\theta = [0, \frac{\pi}{2}], & m = 1, 2, \dots, l, \\ \frac{1}{G} = \frac{1}{G_n} = \frac{1}{G_{\max}} + (n-1) \frac{G_{\min} - G_{\max}}{r-1} \in \left[ \frac{1}{G_{\max}}, \frac{1}{G_{\min}} \right], & n = 1, 2, \dots, r. \end{cases}$$

Then, Eq. (38) leads to the following linear system

$$\begin{bmatrix} S_{1,1}^1 & S_{1,1}^2 & S_{1,1}^3 \\ \vdots & \vdots & \vdots \\ S_{1,r}^1 & S_{1,r}^2 & S_{1,r}^3 \\ \vdots & \vdots & \vdots \\ S_{m,n}^1 & S_{m,n}^2 & S_{m,n}^3 \\ \vdots & \vdots & \vdots \\ S_{l,r}^1 & S_{l,r}^2 & S_{l,r}^3 \end{bmatrix} \begin{bmatrix} b \\ d \\ e \end{bmatrix} = \begin{bmatrix} S_{1,1}^4 \\ \vdots \\ S_{1,r}^4 \\ \vdots \\ S_{m,n}^4 \\ \vdots \\ S_{l,r}^4 \end{bmatrix}, \tag{39}$$



**Table 1**  
Refined optimal parameters for the 25-point scheme.

$I_G$	[4,5]	[5,6]	[6,8]	[8,10]	[10,400]
$b$	0.932846	0.944137	0.949495	0.954379	0.956171
$d$	0.172909	0.161464	0.154458	0.149138	0.146164
$e$	-0.007484	-0.006366	-0.005545	-0.005081	-0.004726

where

$$S_{m,n}^1 = 4 G_n^2 (Q_{m,n} - 1) (P_{m,n} - 1) (P_{m,n} Q_{m,n} - 4 Q_{m,n} - 4 P_{m,n} + 7), \tag{40a}$$

$$S_{m,n}^2 = 12 \pi^2 (P_{m,n}^2 - 2 P_{m,n} + Q_{m,n}^2 - 2 Q_{m,n} + 2), \tag{40b}$$

$$S_{m,n}^3 = 24 \pi^2 (2 P_{m,n}^2 Q_{m,n} - P_{m,n}^2 - 2 P_{m,n} Q_{m,n} - Q_{m,n}^2 + 2), \tag{40c}$$

$$S_{m,n}^4 = 36 \pi^2 + [4 P_{m,n} Q_{m,n} (P_{m,n} Q_{m,n} - 5 Q_{m,n} - 5 P_{m,n} + 16) + P_{m,n} (13 P_{m,n} - 20) + Q_{m,n} (13 Q_{m,n} - 20) - 14] G_n^2, \tag{40d}$$

in which

$$P_{m,n} = \cos\left(\frac{2\pi}{G_n} \cos(\theta_m)\right), \quad Q_{m,n} = \cos\left(\frac{2\pi}{G_n} \sin(\theta_m)\right). \tag{41}$$

The coefficient matrix in linear system (39) has  $l \times r$  rows and 3 columns, thus it is an over-determined system. For example, we can solve the linear system (39) with  $l = 100$ ,  $r = 1000$  and  $I_G = [2, 400]$  and obtain a set of optimal parameters for the 25-point scheme (21) as follows,

$$b = 0.791472, \quad d = 0.292964, \quad \text{and} \quad e = -0.020518. \tag{42}$$

In this article, the finite difference scheme (21) with parameters (42) is referred as the global 25-point scheme for the Helmholtz-PML equation (or simply the global 25p).

We can also implement the refined choice strategy to reduce the numerical dispersion and improve the accuracy of the difference scheme. Several groups of refined parameters are used for plotting normalized phase and group velocity curves are available in Table 1. However, in general,  $I_G$  is estimated by using a priori information before choosing parameters. For example, for a given step size  $h$ ,

$$I_G = \left[ \frac{v_{\min}}{hf_{\max}}, \frac{v_{\max}}{hf_{\min}} \right],$$

where  $f \in [f_{\min}, f_{\max}]$  and  $v \in [v_{\min}, v_{\max}]$  respectively are the frequency and the velocity.

Fig. 1 shows the normalized phase and group velocity curves for refined and global 25p for different propagation angles from the z-axis which include  $0^\circ$ ,  $15^\circ$ ,  $30^\circ$  and  $45^\circ$ . The normalized phase and group velocity curves for the global 25p and the refined 25p are generated using parameters (42) and parameters from Table 1, respectively. In comparison with other optimal methods presented in [2,3,6], the refined 25p provides a satisfactory level of dispersion.

### 3.2. A generalization

In many practical applications, we usually need to use different step sizes  $\Delta x$  and  $\Delta z$  for variables  $x$  and  $z$ , respectively. Here, we present the generalized 25-point scheme (21). Let

$$\begin{aligned} \mathcal{L}_{\Delta x, x} p_{m,n+j} = & \frac{1}{\Delta x^2} \left[ \frac{1}{24} A_{m-\frac{3}{2}, n+j} \left( -\frac{11}{12} p_{m-2, n+j} + \frac{17}{24} p_{m-1, n+j} + \frac{3}{8} p_{m, n+j} \right. \right. \\ & - \frac{5}{24} p_{m+1, n+j} + \left. \frac{1}{24} p_{m+2, n+j} \right) - \frac{9}{8} A_{m-\frac{1}{2}, n+j} \left( \frac{1}{24} p_{m-2, n+j} - \frac{9}{8} p_{m-1, n+j} + \frac{9}{8} p_{m, n+j} \right. \\ & - \left. \frac{1}{24} p_{m+1, n+j} \right) + \frac{9}{8} A_{m+\frac{1}{2}, n+j} \left( \frac{1}{24} p_{m-1, n+j} - \frac{9}{8} p_{m0, n+j} + \frac{9}{8} p_{m+1, n+j} - \frac{1}{24} p_{m+2, n+j} \right) \\ & - \left. \frac{1}{24} A_{m+\frac{3}{2}, n+j} \left( -\frac{1}{24} p_{m-2, n+j} + \frac{5}{24} p_{m-1, n+j} - \frac{3}{8} p_{m, n+j} - \frac{17}{24} p_{m+1, n+j} \right. \right. \\ & \left. \left. + \frac{11}{12} p_{m+2, n+j} \right) \right], \quad j \in \mathbb{Z}_3, \tag{43} \end{aligned}$$

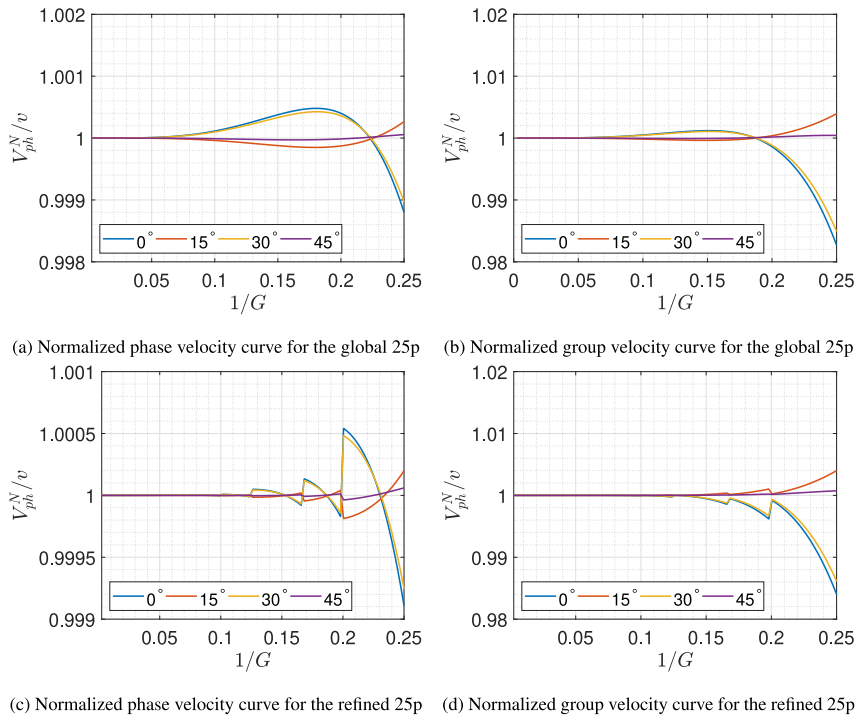


Fig. 1. Normalized phase and group velocity curves for global and refined 25p schemes.

and define,

$$\begin{aligned} \tilde{\mathcal{L}}_{\Delta x, x} p_{m, n} := & b \mathcal{L}_{\Delta x, x} p_{m, n} + (1 - b) \left( -\frac{1}{6} \mathcal{L}_{\Delta x, x} p_{m, n-2} + \frac{2}{3} \mathcal{L}_{\Delta x, x} p_{m, n-1} + \frac{2}{3} \mathcal{L}_{\Delta x, x} p_{m, n+1} \right. \\ & \left. - \frac{1}{6} \mathcal{L}_{\Delta x, x} p_{m, n+2} \right), \end{aligned} \tag{44}$$

where  $b \in (0, 1]$ .

Moreover,  $\tilde{\mathcal{L}}_{\Delta z, z} p_{m, n}$  can be defined in a similar way. Letting  $\tilde{\mathcal{L}}_{\Delta x, \Delta z} = \tilde{\mathcal{L}}_{\Delta x, x} + \tilde{\mathcal{L}}_{\Delta z, z}$ , Eq. (2) can be approximated as

$$\tilde{\mathcal{L}}_{\Delta x, \Delta z} p|_{x=x_m, z=z_n} + I(k^2 Cp)|_{x=x_m, z=z_n} = 0. \tag{45}$$

**Remark 3.1.** If  $c + d + e = 1$ , then the 25-point FD approximation (45) is consistent with Helmholtz-PML equation (2).

Let  $\Delta x = h$ ,  $\Delta z = \gamma h$  and  $\eta = 1 + \frac{1}{\gamma^2}$  where  $\gamma$  is a positive constant. It follows from performing classical dispersion analysis to the finite difference method (45) that

$$k_N = \frac{1}{h} \sqrt{\frac{\tilde{N}}{\tilde{D}}}, \tag{46}$$

where

$$\begin{aligned} \tilde{N} = & -2(\tilde{Q} - 1)(P - 1) \left( (P\tilde{Q} - \tilde{Q} - 7P + 7)\eta + 6P - 6\tilde{Q} \right) b \\ & + 3(P - \tilde{Q})(4P\tilde{Q} - 5\tilde{Q} - 5P + 12) + (\tilde{Q} - 1)(\tilde{Q} - 7)(2P^2 - 4P - 1)\eta, \end{aligned} \tag{47}$$

$$\begin{aligned} \tilde{D} = & 3(P^2 + \tilde{Q}^2 - 2(P + \tilde{Q}) + 2)d + 6(2(P^2\tilde{Q}^2 - P\tilde{Q}) \\ & - (P^2 + \tilde{Q}^2) + 2)e - 9. \end{aligned} \tag{48}$$

in which  $\tilde{Q} = \cos(k_z \Delta z) = \cos(\gamma kh \sin \theta)$ .

As  $h = \frac{2\pi}{Gk}$ , we conclude that

$$k_N = \frac{G}{2\pi} \sqrt{\frac{\tilde{N}}{\tilde{D}}} - 1. \tag{49}$$

The optimal parameters  $b, d$  and  $e$  can be chosen by minimizing the numerical dispersion. To do this, we define the functional

$$J_\eta(b, d, e; G, \theta) = \frac{G}{2\pi} \sqrt{\frac{\tilde{N}}{\tilde{D}}} - 1. \tag{50}$$

The global and refined choice strategies also can be incorporated here; however, the main difference is that  $J_\eta(b, d, e; G, \theta)$  is used instead of  $J(b, d, e; G, \theta)$ .

Another benefit of using the 25p FDM over the 17p FDM is that the 25p FDM can be easily generalized. Note that,

$$\begin{aligned} \Delta_{h,45^\circ} p|_{x=x_m, z=z_n} &= \left(\frac{\gamma^2 + 1}{2\gamma^2}\right) \frac{\partial^2}{\partial x^2} p + \left(\frac{\gamma^2 + 1}{2}\right) \frac{\partial^2}{\partial z^2} p \\ &\quad - \left(\frac{\gamma^2 - 1}{\gamma}\right) \frac{\partial}{\partial z} \frac{\partial}{\partial x} p + \mathcal{O}(h^4), \end{aligned} \tag{51}$$

This means that  $\Delta_{h,45^\circ} p|_{x=x_m, z=z_n}$  can approximate  $\Delta p$  with fourth-order accuracy if  $\gamma = 1$  (or  $\Delta x = \Delta z$ ).

#### 4. Numerical experiments

In this section, we present three numerical examples. Section 4.1 is meant for illustrating the accuracy of the schemes introduced in this article and comparing with some other optimal finite difference methods that have been used for the Helmholtz equation with PML. Sections 4.2 and 4.3 are considered for more practical usages of the new schemes, the refined 25p, and the refined 17p schemes. Section 4.2 will show that there could be almost no difference between the refined 25p and the refined 17p schemes for some problems; however, Section 4.3 will show that we could get better results by using the refined 25p scheme in more complex examples.

##### 4.1. Problem 1

In this experiment, we solve the Helmholtz equation

$$\Delta p + k^2 p = g(x, z), \quad \text{in } \Omega := (0, 1) \times (0, 1), \tag{52}$$

with

$$g(x, z) = k_0^2 e^{-2(x+z)} (2e^{x+z} + 1) e^{ik_0(x \cos(\theta) + z \sin(\theta))}, \tag{53}$$

$$(k_x, k_z) = k_0 (e^{-x-z} + 1) (\cos(\theta), \sin(\theta)). \tag{54}$$

Dirichlet boundary conditions are imposed on this problem, and its analytical solution is expressed as follows,

$$p(x, z) := e^{ik_0(x \cos(\theta) + z \sin(\theta))}. \tag{55}$$

We compare the 25p and 17p finite difference methods with a number of popular optimal finite difference schemes that are used for the Helmholtz equation with variable wavenumbers. The new schemes are compared with the 9p FDM [3], rotated 9p FDM [2,3], and the point-weighting 9p [18].

Moreover, the parameters of all refined schemes are estimated using the refined choice strategy [3] in which  $I_C = [G_{\min}, G_{\max}] = \left[\frac{2\pi}{hk_{\max}}, \frac{2\pi}{hk_{\min}}\right]$  is estimated by using a priori information. The parameters of global schemes are estimated using the global choice strategy [3] with  $I_C = [2, 400]$ .

Furthermore, the optimal rotated 9p represents the rotated 9-point FDM [2,3] with parameters  $a = 0.5461, d = 0.3752$  and  $e = -4 \times 10^{-5}$ . This group of optimal parameters is provided by Jo, Shin and Suh [2,3] as optimal parameters. The point-weighting 9p is the point-weighting scheme [18] that its parameters are estimated using the flexible strategy for selection of weights (rule 3.3 at page 2354 of [18]).

All the experiments in this example are performed with MATLAB 9.5.0.1033004 (R2018b) Update 2 (64-bit) on a Dell laptop equipped with Windows 10 Home Edition (64-bit), Intel(R) Core(TM) i5-4210U CPU, and 8.00 GB Physical Memory (RAM). We used an unsymmetric-pattern multifrontal (UMF) method for sparse LU factorization [27,28] for solving linear systems generated by each FDM.

Additionally, the error between the numerical solution and the exact solution is measured in C-norm [3], which is defined for any complex vector  $\mathbf{z} = [z_1, z_2, \dots, z_M]$ , as

$$\|\mathbf{z}\|_C = \max_{1 \leq j \leq M} |z_j|. \tag{56}$$

where  $|z_j|$  is the complex modulus of  $z_j$ .

**Table 2**The error in the C-norm for  $k_0 = 100$ .

$N$	161	321	641
Refined 25p	1.7619e-02	1.9780e-03	1.4992e-04
Global 25p	7.6060e-02	6.7285e-03	3.3061e-04
Refined 17p	2.0475e-02	2.0235e-03	1.2071e-04
Global 17p	3.8338e-02	1.8634e-03	2.4496e-04
Refined 9p	7.9471e-01	2.2035e-01	5.6881e-02
Global 9p	2.3546e+00	4.3470e-01	5.7789e-02
Refined rotated 9p	7.8142e-01	2.2044e-01	5.6887e-02
Optimal rotated 9p	6.4635e-01	1.8743e-01	5.4399e-02
Point-weighting 9p	7.8142e-01	2.2044e-01	5.6887e-02

**Table 3**The error in the C-norm for  $k_0 = 200$ .

$N$	321	641	1281
Refined 25p	7.0579e-02	2.6068e-03	8.1386e-05
Global 25p	3.3057e-02	1.6795e-02	8.3128e-05
Refined 17p	6.2034e-02	2.6828e-03	6.5833e-05
Global 17p	5.7523e-02	4.1146e-02	1.5328e-04
Refined 9p	4.6659e+00	1.4241e+00	5.5252e-01
Global 9p	2.5599e+00	2.2858e+00	1.5675e-01
Refined rotated 9p	9.2534e+00	1.5010e+00	5.4948e-01
Optimal rotated 9p	3.2353e+00	6.8168e-01	6.9660e-01
Point-weighting 9p	9.2534e+00	1.5010e+00	5.4948e-01

**Table 4**The error in the C-norm for  $k_0 = 300$ .

$N$	321	641	1281
Refined 25p	6.2525e-01	1.6934e-02	2.5024e-03
Global 25p	1.5566e-01	6.8584e-03	1.5261e-03
Refined 17p	3.1920e-01	2.0192e-02	2.4412e-03
Global 17p	6.3903e-01	2.9267e-02	1.2975e-03
Refined 9p	8.7409e+01	1.0444e+00	4.2841e-01
Global 9p	1.4680e+01	8.1048e+00	3.9829e-01
Refined rotated 9p	4.3196e+00	3.2259e+00	4.3469e-01
Optimal rotated 9p	7.1256e+00	1.8943e+00	9.2993e+00
Point-weighting 9p	4.3196e+00	3.2259e+00	4.3469e-01

Tables 2–4 show the error in the C-norm for different schemes with different gridpoints  $N$  per line with  $k_0 = 100$ , 200 and 300, respectively.

Table 2 highlights that the refined 25p and the refined 17p provide the best accuracy among all methods when  $k_0 = 100$ . This table also shows that the four new schemes are fourth-order schemes. This can be seen since the C-norm gets roughly  $2^4$  better by decreasing the step-size by half each time. In addition, it can be seen that the refined 25p only needs a quarter of number of gridpoints to obtain almost three times better than the accuracy of the 9p schemes.

Similarly, it can be seen from Tables 3 and 4 that the refined 25p and the refined 17p have provided, overall, a better level of accuracy than the other schemes when  $k_0 = 200$  and  $k_0 = 300$ . It can be also seen that C-norm does not always decrease by a factor of  $2^4$  when  $N$  is increased by a factor of two. When  $k_0$  is very large and  $h$  is very small,  $h^4$  can be comparable to the round off error. Moreover, the round off error is a function of the optimal parameters ( $b$ ,  $d$  and  $e$ ). Therefore, the round off error can get sometimes too small for some values of  $h$ . Then, the round off error cannot decrease by a factor of  $2^4$  when the size of increment  $h$  is decreased to  $h/2$ . This is more noticeable for larger values of  $k_0$ , where the difference between  $G_{\min}$  and  $G_{\max}$  is large. In summary, the new method is fourth-order accurate, as long as  $h$  is not too small. For example, consider the global 25p scheme in Table 4, when  $N$  is increased by a factor of 1.25, from 641 to 801, the error decreases to  $2.5135e-03$ . It can be seen that the error gets roughly  $(1.25)^4$  smaller, which confirms the method is fourth-order. However, when  $h$  gets very small, the random round-off error cannot be ignored, together with the effect from large value of  $k$ , the overall accuracy will be affected. As can be seen, then  $N$  is increased to 1281, the error is not decreasing as expected. Therefore, too small  $h$  is not recommended when the presented method is used to discretize the Helmholtz equation with PML.

Furthermore, Table 5 shows the error in the C-norm for  $k_0 = 400$ ,  $N = 901$  and  $\theta = 0, \frac{\pi}{16}, \dots, \frac{\pi}{4}$ . It can be seen from the table that the refined 25p and the refined 17p provide, overall, a better level of accuracy than the global 25p and the global 17p.

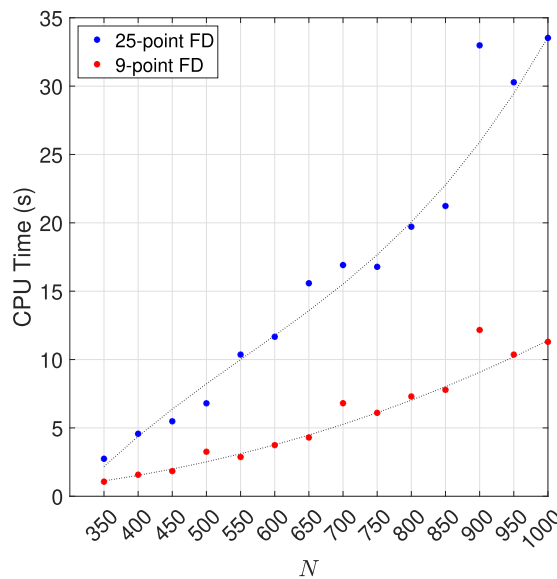
Table 6 and Fig. 2 compare the CPU time of a 25-point FDM and a 9-point FDM. The matrix in the linear system associated with each FDM is a sparse matrix with complex values. We used the UMF method for sparse LU factorization

**Table 5**  
The error in the C-norm for  $k_0 = 400$  and  $N = 901$ .

$\theta$	0	$\frac{\pi}{16}$	$\frac{\pi}{8}$	$\frac{3\pi}{16}$	$\frac{\pi}{4}$
Refined 25p	2.5990e-02	2.8542e-02	2.4396e-02	1.6240e-02	2.0541e-02
Global 25p	5.7248e-01	2.9213e-01	2.3036e-01	1.1509e-01	1.0332e-01
Refined 17p	2.4487e-02	2.6372e-02	2.3109e-02	1.5984e-02	2.0808e-02
Global 17p	6.3795e-02	4.2729e-02	3.3544e-02	1.9199e-02	2.3303e-02

**Table 6**  
CPU time of a 25-point FDM and a 9-point FDM.

$N$	350	400	450	500	550	600	650
25-point FD	2.74	4.57	5.48	6.80	10.36	11.67	15.58
9-point FD	1.07	1.58	1.84	3.26	2.87	3.74	4.30
$N$	700	750	800	850	900	950	1000
25-point FD	16.91	16.78	19.71	21.23	32.98	30.28	33.53
9-point FD	6.81	6.10	7.30	7.78	12.16	10.36	11.30



**Fig. 2.** CPU time of a 25-point FDM and a 9-point FDM.

to solve the linear systems. It can be seen that solving a 25-point FDM is almost 2.92 computationally costlier than solving a 9-point FDM. However, to reach a certain accuracy, the CPU time of the refined 25p is much less than that of the refined 9p. For example, when  $k_0 = 100$ , the refined 25p can reach the accuracy of  $1.7619e-02$  using  $N = 161$  gridpoints in 0.44 s; on the contrary, to reach a comparable accuracy ( $1.7678e-02$ ), the refined 9p needs much finer grids ( $N = 1151$ ) and the CPU time of 16.87 s.

The results of this example can be summarized as follows. The refined 25p is computationally costlier than the 9p schemes for Helmholtz equation with PML; however, it has been shown that the CPU time of the refined 25p is much less than that of the refined 9p to reach a certain accuracy. Moreover, in many practical cases that involve large domains, steps sizes are usually large and higher order methods tend to provide better efficiency.

#### 4.2. Problem 2: a fault model

In this example, we consider a fault P-wave velocity model, shown in Fig. 3, to test the refined 25p and refined 17p finite difference methods. Horizontal and vertical samplings are  $n_x = n_z = 200$  with sampling intervals  $\Delta x = \Delta z = 10$  m, and the time sampling is  $\Delta t = 8$  ms. A point source  $\delta(x-x_s, z-z_s)R(\omega, f_M)$  is located at the point  $(x_s, z_s) = (1000 \text{ m}, 0 \text{ m})$ , where  $R(\omega, f_M)$  is the Ricker wavelet, defined in Eq. (57), with the peak frequency  $f_M = 15$  Hz.

$$R(t, f_M) = (1 - 2\pi^2 f_M^2 t^2) / \exp(\pi^2 f_M^2 t^2). \tag{57}$$

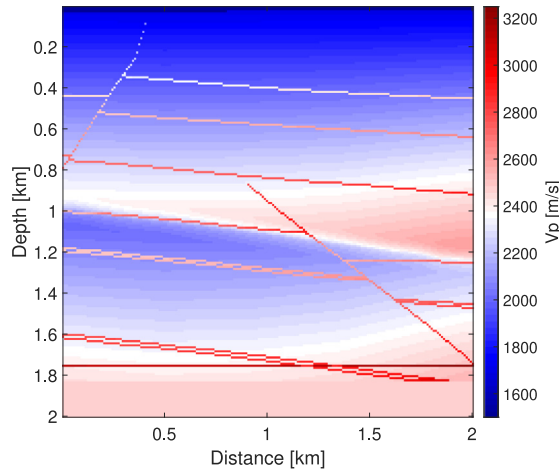


Fig. 3. A fault model.

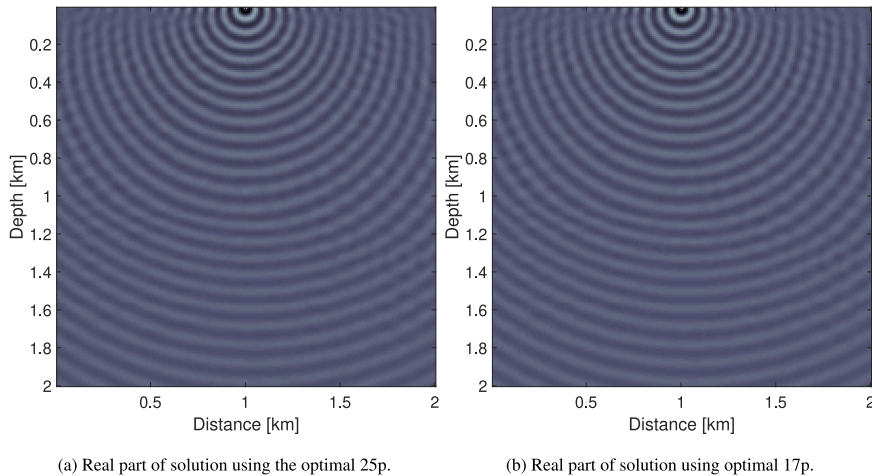


Fig. 4. The monofrequency wavefields (real part) for  $f = 37.6984$  Hz.

For the two optimal schemes,  $I_G = [G_{\min}, G_{\max}] = \left[ \frac{v_{\min}}{f h}, \frac{v_{\max}}{f h} \right]$  is estimated by using a priori information. In addition, the damping profiles here are defined using  $\sigma_x$  and  $\sigma_z$  from (3) with  $L_{PML} = 500$  m and  $a_0 = 1.7900$ .

Moreover, the monofrequency wavefields (real parts) for  $f = 37.6984$  Hz obtained by the refined 25p (Fig. 4(a)) and the refined 17p (Fig. 4(b)) are presented in Fig. 4.

Furthermore, Fig. 5 includes two snapshots for the time being  $t = 400$  ms generated using respectively the refined 25p (Fig. 5(a)) and the refined 17p (Fig. 5(b)) schemes. We have also included the shot records for a shot at a x-position of 1000 m in Fig. 6.

What can be seen from Figs. 4–6 can be summarized as follows. First off, the PML’s absorption is efficient, and almost no boundary reflections exist. Additionally, the upward incident waves, the downward incident waves and transmissive waves are all clear.

Moreover, from Proposition 2.1, we know that the 17p FDM approximation is not pointwise consistent with the Helmholtz-PML equation; however, there is no significant difference between the two schemes, the refined 25p and the refined 17p. That is why we next investigate a more complex example, the Marmousi model.

### 4.3. Problem 3: Marmousi model

Fig. 7 shows the Marmousi P-wave velocity model. In this example, horizontal and vertical samplings are  $n_x = 150$  and  $n_z = 450$ , respectively. The sampling intervals are  $\Delta x = \Delta z = 20$  m, and the time sampling is  $\Delta t = 8$  ms. A Ricker wavelet with a peak frequency of 15 Hz is placed at  $(x_s, z_s) = (4500 \text{ m}, 0 \text{ m})$  as a source. For this example, the damping profiles defined using  $\sigma_x$  and  $\sigma_z$  from (3) with  $L_{PML} = 1000$  m and  $a_0 = 1.7900$ .

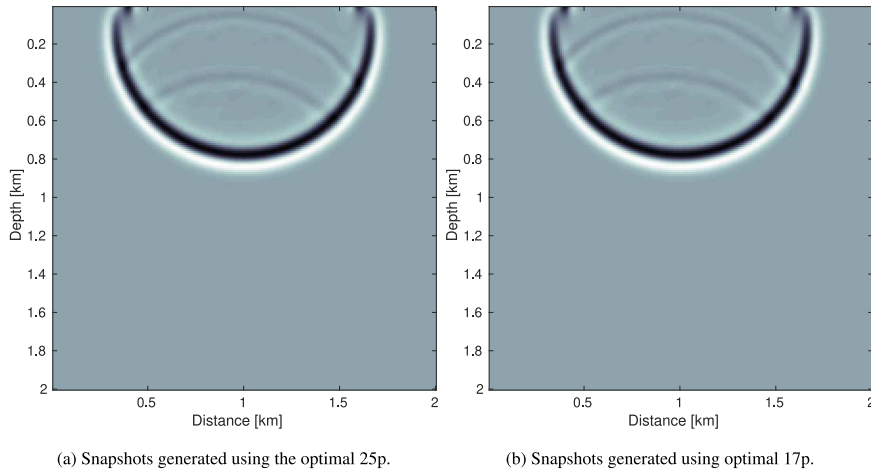


Fig. 5. Snapshots generated at the time 400 ms.

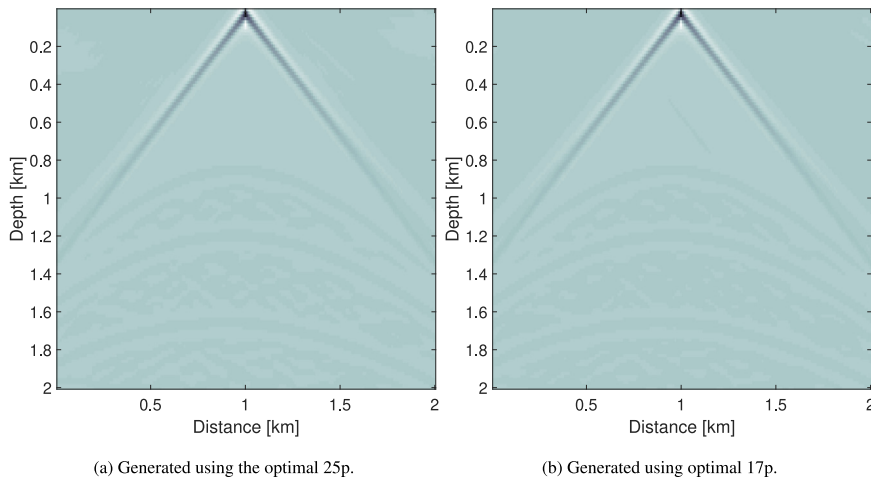


Fig. 6. Shot records for a shot at a X position of 1000 m.

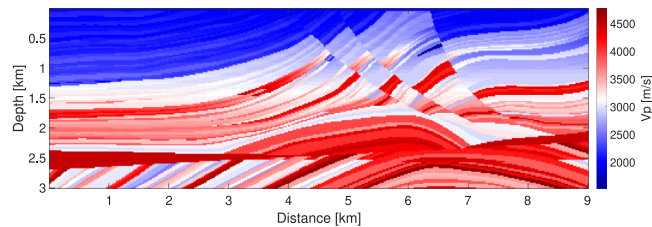


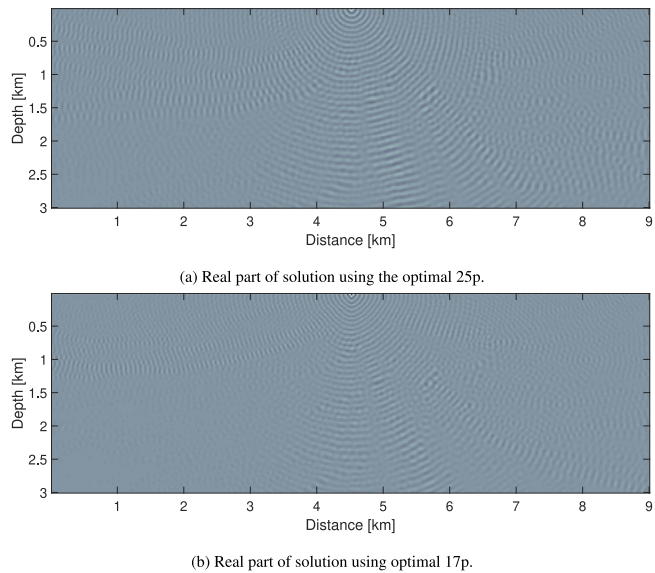
Fig. 7. Marmousi model.

Moreover, the monofrequency wavefields (real part) for  $f = 20.2793$  Hz obtained by the refined 25p (Fig. 8(a)) and the refined 17p (Fig. 8(b)) are presented in Fig. 8. As it can be seen, the refined 25p works better here than the refined 17p. Consider center top of Figs. 8(a) and 8(b). As it can be seen, the wavefields in the top of Fig. 8(b) no longer have a circular shape. This is due to the numerical dispersion error caused by the refined 17p finite difference method.

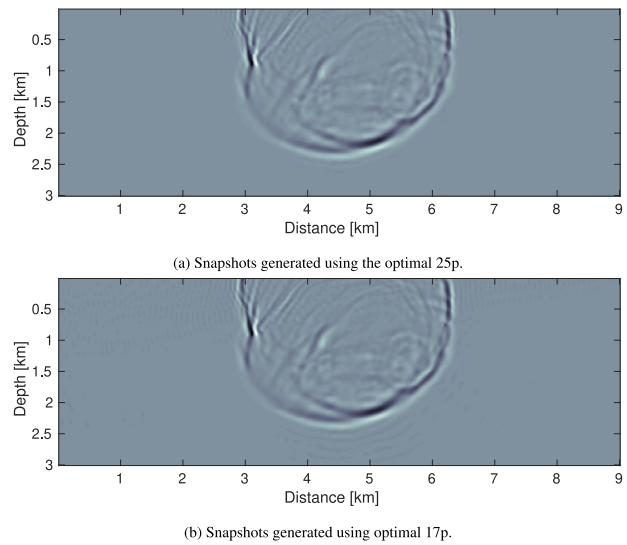
In addition, Fig. 9 shows the snapshots for the time  $t = 792$  ms obtained using the refined 25p (Fig. 9(a)) and the refined 17p (Fig. 9(b)) schemes, respectively. The snapshot generated by the optimal 25p is much clearer than that of generated by the refined 17p. In addition, the shot records for a shot at a x-position of 4.5 km can be found in Fig. 10.

In this example, we tested the refined 25p and the refined 17p with the Marmousi model, and the refined 25p showed more reliable results than the refined 17p.





**Fig. 8.** The monofrequency wavefields (real part) for  $f = 20.2793$  Hz.



**Fig. 9.** Snapshots generated at the time 792 ms.

### 5. Conclusions

In this section, we summarize and comment on the finite difference methods. In Section 2, our extension of the rotated 9-point FDM to the 17-point FDM is also pointwise inconsistent with the Helmholtz-PML equation (2). In addition, a consistent 25-point FDM for the Helmholtz-PML equation was presented in Section 3. For this method, we presented an error analysis for the numerical approximation of the exact wavenumber, and applied the global and the refined choice strategies for choosing optimal parameters based on minimizing the numerical dispersion. Finally, numerical experiments were presented in Section 4 to practice the theoretical results.

Moreover, the refined 25p and the refined 17p are computationally costlier than any other optimal 9p methods; however, given the fact that an almost same accuracy can be obtained using a much fewer number of gridpoints, the overall computational costs are still comparable. Furthermore, in many practical cases that involve large domains, steps sizes are usually large and higher order methods tend to provide better efficiency. Sections 4.2 and 4.3 were considered for more practical usages of the using new schemes. In particular, Section 4.3 has shown that we could get better results by using the refined 25p in more complex examples.



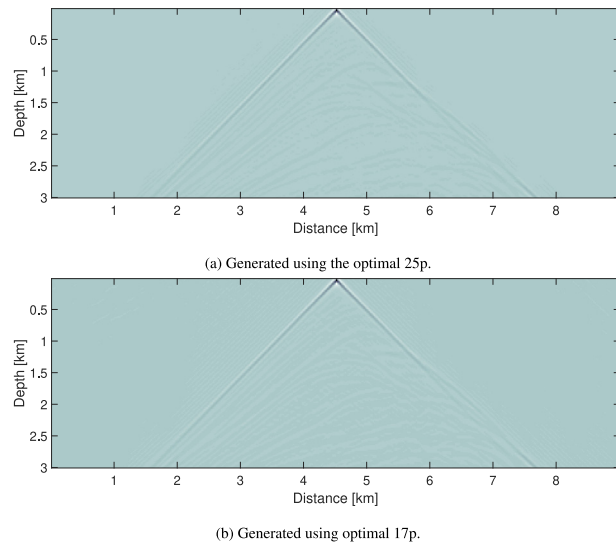


Fig. 10. Shot records for a shot at a X position of 4.5 km.

**Acknowledgments**

The work is supported by the Natural Sciences & Engineering Research Council of Canada (NSERC) through the individual Discovery Grant program. The authors would like to thank anonymous reviewers whose comments have greatly improved this manuscript. The first author is also thankful for the Alberta Innovates Graduate Student Scholarship that he received during his PhD studies.

**Appendix A. Deriving FD approximation for the Helmholtz-PML equation**

It follows from Eq. (2) that

$$\begin{aligned} \frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) \Big|_{x=x_m, z=z_n} &= a_1 \left( A \frac{\partial p}{\partial x} \right) \Big|_{x=x_m - \frac{3h}{2}, z=z_n} + a_2 \left( A \frac{\partial p}{\partial x} \right) \Big|_{x=x_m - \frac{h}{2}, z=z_n} \\ &+ a_3 \left( A \frac{\partial p}{\partial x} \right) \Big|_{x=x_m + \frac{h}{2}, z=z_n} + a_4 \left( A \frac{\partial p}{\partial x} \right) \Big|_{x=x_m + \frac{3h}{2}, z=z_n}, \end{aligned} \tag{A.1}$$

$$\begin{aligned} \frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) \Big|_{x=x_m, z=z_n} &= b_1 \left( B \frac{\partial p}{\partial z} \right) \Big|_{x=x_m, z=z_n - \frac{3h}{2}} + b_2 \left( B \frac{\partial p}{\partial z} \right) \Big|_{x=x_m, z=z_n - \frac{h}{2}} \\ &+ b_3 \left( B \frac{\partial p}{\partial z} \right) \Big|_{x=x_m, z=z_n + \frac{h}{2}} + b_4 \left( B \frac{\partial p}{\partial z} \right) \Big|_{x=x_m, z=z_n + \frac{3h}{2}}. \end{aligned} \tag{A.2}$$

We need to find  $a_i$  and  $b_i$  ( $i = 1, \dots, 4$ ) such that (A.1) and (A.2) can approximate  $\frac{\partial}{\partial x} \left( A \frac{\partial p}{\partial x} \right) \Big|_{x=x_m, z=z_n}$  and  $\frac{\partial}{\partial z} \left( B \frac{\partial p}{\partial z} \right) \Big|_{x=x_m, z=z_n}$  with 4th-order accuracy. Applying the Taylor theorem on right-hand sides of (A.1) and (A.2), and then solving the generated linear systems for  $a_i$  and  $b_i$  with  $i = 1, \dots, 4$ , we have

$$a_1 = b_1 = \frac{1}{24h}, \quad a_2 = b_2 = -\frac{9}{8h}, \quad a_3 = b_3 = \frac{9}{8h}, \quad \text{and} \quad a_4 = b_4 = -\frac{1}{24h}.$$

Now, we need to approximate  $\frac{\partial p}{\partial x}$  and  $\frac{\partial p}{\partial z}$  at points  $(x_m - \frac{3h}{2}, z_n), \dots, (x_m, z_n + \frac{3h}{2})$  with fourth-order accuracy. For example, we can find  $d_i$  for  $i = 1, \dots, 5$  as follows

$$\begin{aligned} \frac{\partial p}{\partial x} \Big|_{x=x_m - \frac{3h}{2}, z=z_n} &= d_1 p|_{x=x_m - 2h, z=z_n} + d_2 p|_{x=x_m - h, z=z_n} + d_3 p|_{x=x_m, z=z_n} \\ &+ d_4 p|_{x=x_m + h, z=z_n} + d_5 p|_{x=x_m + 2h, z=z_n}, \end{aligned} \tag{A.3}$$

and

$$d_1 = -\frac{11}{12h}, \quad d_2 = \frac{17}{24h}, \quad d_3 = \frac{3}{8h}, \quad d_4 = -\frac{5}{24h} \quad \text{and} \quad d_5 = \frac{1}{24h}. \tag{A.4}$$

**Appendix B. Long expressions**

$$\begin{aligned} \mu_1 = & -\frac{1}{90} \left( A \frac{\partial^6 p}{\partial x^6} + B \frac{\partial^6 p}{\partial z^6} \right) - \frac{1}{30} \left( \frac{\partial A}{\partial x} \frac{\partial^5 p}{\partial x^5} + \frac{\partial B}{\partial z} \frac{\partial^5 p}{\partial z^5} \right) \\ & - \frac{3}{64} \left( \frac{\partial^2 A}{\partial x^2} \frac{\partial^4 p}{\partial x^4} + \frac{\partial^2 B}{\partial z^2} \frac{\partial^4 p}{\partial z^4} \right) - \frac{3}{64} \left( \frac{\partial^3 A}{\partial x^3} \frac{\partial^3 p}{\partial x^3} + \frac{\partial^3 B}{\partial z^3} \frac{\partial^3 p}{\partial z^3} \right) \\ & - \frac{3}{128} \left( \frac{\partial^4 A}{\partial x^4} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^4 B}{\partial z^4} \frac{\partial^2 p}{\partial z^2} \right) - \frac{3}{640} \left( \frac{\partial^5 A}{\partial x^5} \frac{\partial p}{\partial x} + \frac{\partial^5 B}{\partial z^5} \frac{\partial p}{\partial z} \right), \end{aligned} \tag{B.1}$$

$$\begin{aligned} \mu_2 = & -\frac{1}{180} (A + B) \left( \frac{\partial^6 p}{\partial x^6} + 15 \frac{\partial^2}{\partial z^2} \frac{\partial^4 p}{\partial x^4} + 15 \frac{\partial^4}{\partial z^4} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^6 p}{\partial z^6} \right) + \frac{1}{90} (A - B) \left( 3 \frac{\partial}{\partial z} \frac{\partial^5 p}{\partial x^5} \right. \\ & + 10 \frac{\partial^3}{\partial z^3} \frac{\partial^3 p}{\partial x^3} + 3 \frac{\partial^5}{\partial z^5} \frac{\partial p}{\partial x} \left. \right) - \frac{1}{60} \left( \frac{\partial^5 p}{\partial x^5} + 10 \frac{\partial^2}{\partial z^2} \frac{\partial^3 p}{\partial x^3} + 5 \frac{\partial^4}{\partial z^4} \frac{\partial p}{\partial x} \right) \left( \frac{\partial A}{\partial x} + \frac{\partial B}{\partial x} \right. \\ & \left. - \frac{\partial A}{\partial z} + \frac{\partial B}{\partial z} \right) + \frac{1}{60} \left( 5 \frac{\partial}{\partial z} \frac{\partial^4 p}{\partial x^4} + 10 \frac{\partial^3}{\partial z^3} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^5 p}{\partial z^5} \right) \left( \frac{\partial A}{\partial x} - \frac{\partial B}{\partial x} - \frac{\partial A}{\partial z} - \frac{\partial B}{\partial z} \right) \\ & - \frac{3}{128} \left( \frac{\partial^4 p}{\partial x^4} + 6 \frac{\partial^2}{\partial z^2} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^4 p}{\partial z^4} \right) \left( \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 A}{\partial z^2} + \frac{\partial^2 B}{\partial z^2} \right) + \frac{3}{32} \left( \frac{\partial}{\partial z} \frac{\partial^3 p}{\partial x^3} \right. \\ & \left. + \frac{\partial^3}{\partial z^3} \frac{\partial p}{\partial x} \right) \left( \frac{\partial^2 A}{\partial x^2} - \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 A}{\partial z^2} - \frac{\partial^2 B}{\partial z^2} \right) - \frac{3}{128} \left( \frac{\partial^3 p}{\partial x^3} + 3 \frac{\partial^2}{\partial z^2} \frac{\partial p}{\partial x} \right) \left( \frac{\partial^3 B}{\partial x^3} - \frac{\partial^3 A}{\partial z^3} \right. \\ & \left. + \frac{\partial^3 B}{\partial z^3} + \frac{\partial^3 A}{\partial x^3} \right) + \frac{3}{128} \left( 3 \frac{\partial}{\partial z} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^3 p}{\partial z^3} \right) \left( \frac{\partial^3 A}{\partial x^3} - \frac{\partial^3 A}{\partial z^3} - \frac{\partial^3 B}{\partial z^3} - \frac{\partial^3 B}{\partial x^3} \right) \\ & - \frac{9}{128} \left( 3 \frac{\partial}{\partial z} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^3 p}{\partial z^3} \right) \left( \frac{\partial}{\partial z} \frac{\partial^2 A}{\partial x^2} + \frac{\partial}{\partial z} \frac{\partial^2 B}{\partial x^2} - \frac{\partial^2}{\partial z^2} \frac{\partial A}{\partial x} + \frac{\partial^2}{\partial z^2} \frac{\partial B}{\partial x} \right) \\ & + \frac{9}{128} \left( \frac{\partial^3 p}{\partial x^3} + 3 \frac{\partial^2}{\partial z^2} \frac{\partial p}{\partial x} \right) \left( \frac{\partial}{\partial z} \frac{\partial^2 A}{\partial x^2} - \frac{\partial}{\partial z} \frac{\partial^2 B}{\partial x^2} - \frac{\partial^2}{\partial z^2} \frac{\partial A}{\partial x} - \frac{\partial^2}{\partial z^2} \frac{\partial B}{\partial x} \right) \\ & - \frac{3}{256} \left( \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} \right) \left( \frac{\partial^4 A}{\partial x^4} + \frac{\partial^4 B}{\partial x^4} + \frac{\partial^4 A}{\partial z^4} + \frac{\partial^4 B}{\partial z^4} \right) \\ & + \frac{3}{128} \left( \frac{\partial}{\partial z} \frac{\partial p}{\partial x} \right) \left( \left( \frac{\partial^4 A}{\partial x^4} - \frac{\partial^4 B}{\partial x^4} + \frac{\partial^4 A}{\partial z^4} - \frac{\partial^4 B}{\partial z^4} \right) - 4 \left( \frac{\partial}{\partial z} \frac{\partial^3 A}{\partial x^3} + \frac{\partial}{\partial z} \frac{\partial^3 B}{\partial x^3} + \frac{\partial^3}{\partial z^3} \frac{\partial A}{\partial x} \right. \right. \\ & \left. \left. + \frac{\partial^3}{\partial z^3} \frac{\partial B}{\partial x} \right) \right) + \frac{3}{64} \left( \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} \right) \left( \frac{\partial}{\partial z} \frac{\partial^3 A}{\partial x^3} - \frac{\partial}{\partial z} \frac{\partial^3 B}{\partial x^3} + \frac{\partial^3}{\partial z^3} \frac{\partial A}{\partial x} - \frac{\partial^3}{\partial z^3} \frac{\partial B}{\partial x} \right) \\ & - \frac{3}{1280} \left( \frac{\partial p}{\partial x} \left( \frac{\partial^5 A}{\partial x^5} + \frac{\partial^5 B}{\partial x^5} - \frac{\partial^5 A}{\partial z^5} + \frac{\partial^5 B}{\partial z^5} \right) - \frac{\partial p}{\partial z} \left( \frac{\partial^5 A}{\partial x^5} - \frac{\partial^5 B}{\partial x^5} - \frac{\partial^5 A}{\partial z^5} - \frac{\partial^5 B}{\partial z^5} \right) \right) \\ & - \frac{3}{256} \frac{\partial p}{\partial z} \left( \frac{\partial}{\partial z} \frac{\partial^4 A}{\partial x^4} + \frac{\partial}{\partial z} \frac{\partial^4 B}{\partial x^4} - \frac{\partial^4}{\partial z^4} \frac{\partial A}{\partial x} + \frac{\partial^4}{\partial z^4} \frac{\partial B}{\partial x} \right) + \frac{3}{256} \frac{\partial p}{\partial x} \left( \frac{\partial}{\partial z} \frac{\partial^4 A}{\partial x^4} - \frac{\partial}{\partial z} \frac{\partial^4 B}{\partial x^4} \right. \\ & \left. - \frac{\partial^4}{\partial z^4} \frac{\partial A}{\partial x} - \frac{\partial^4}{\partial z^4} \frac{\partial B}{\partial x} \right) + \left( \frac{\partial p}{\partial x} + \frac{\partial p}{\partial z} \right) \left( \frac{\partial^2}{\partial z^2} \frac{\partial^3 B}{\partial x^3} + \frac{\partial^3}{\partial z^3} \frac{\partial^2 B}{\partial x^2} \right) \\ & - \left( 2 \frac{\partial^4 p}{\partial x^4} - 8 \frac{\partial}{\partial z} \frac{\partial^3 p}{\partial x^3} + 12 \frac{\partial^2}{\partial z^2} \frac{\partial^2 p}{\partial x^2} - 8 \frac{\partial^3}{\partial z^3} \frac{\partial p}{\partial x} + 2 \frac{\partial^4 p}{\partial z^4} \right) \frac{\partial}{\partial z} \frac{\partial A}{\partial x} \\ & + \left( 2 \frac{\partial^4 p}{\partial x^4} + 8 \frac{\partial}{\partial z} \frac{\partial^3 p}{\partial x^3} + 12 \frac{\partial^2}{\partial z^2} \frac{\partial^2 p}{\partial x^2} + 8 \frac{\partial^3}{\partial z^3} \frac{\partial p}{\partial x} + 2 \frac{\partial^4 p}{\partial z^4} \right) \left( \frac{\partial}{\partial z} \frac{\partial B}{\partial x} + \frac{\partial^2}{\partial z^2} \frac{\partial^2 A}{\partial x^2} \right) \\ & + \left( 3 \frac{\partial^2 p}{\partial x^2} + 6 \frac{\partial}{\partial z} \frac{\partial p}{\partial x} + 3 \frac{\partial^2 p}{\partial z^2} \right) \frac{\partial^2}{\partial z^2} \frac{\partial^2 B}{\partial x^2} + \left( \frac{\partial p}{\partial x} - \frac{\partial p}{\partial z} \right) \left( \frac{\partial^2}{\partial z^2} \frac{\partial^3 A}{\partial x^3} - \frac{\partial^3}{\partial z^3} \frac{\partial^2 A}{\partial x^2} \right), \end{aligned} \tag{B.2}$$

$$\mu_3 = -\frac{1}{12} (d + 2e) \left( \frac{\partial^4}{\partial x^4} (k^2 Cp) + \frac{\partial^4}{\partial z^4} (k^2 Cp) \right) - e \frac{\partial^4}{\partial z^2 \partial x^2} (k^2 Cp). \quad (\text{B.3})$$

$$\begin{aligned} \xi_1 = & \left( \frac{2b-2}{3} \right) \left[ \frac{\partial^5}{\partial z^3 \partial x^2} p \frac{\partial}{\partial z} A + \frac{\partial^2}{\partial z \partial x} A \frac{\partial^3}{\partial z^3} \frac{\partial}{\partial x} p + \frac{\partial^2}{\partial z \partial x} p \frac{\partial^3}{\partial z^3} \frac{\partial}{\partial x} A + \frac{\partial^3}{\partial z^3} A \frac{\partial^3}{\partial z \partial x^2} p \right] \\ & - \frac{1}{30} \left[ \frac{1}{3} A \frac{\partial^6}{\partial x^6} p + \frac{\partial}{\partial x} A \frac{\partial^5}{\partial x^5} p + \frac{9}{64} \frac{\partial}{\partial x} p \frac{\partial^5}{\partial x^5} A + \frac{45}{64} \frac{\partial^4}{\partial x^4} A \frac{\partial^2}{\partial x^2} p \right] \\ & - \frac{3}{64} \left[ \frac{\partial^3}{\partial x^3} A \frac{\partial^3}{\partial x^3} p + \frac{\partial^4}{\partial x^4} p \frac{\partial^2}{\partial x^2} A \right] + (b-1) \left( \frac{\partial^3}{\partial z^2 \partial x} p \frac{\partial^3}{\partial z^2 \partial x} A + \frac{\partial^4}{\partial z^2 \partial x^2} p \frac{\partial^2}{\partial z^2} A \right) \\ & + \left( \frac{b-1}{6} \right) \left[ A \frac{\partial^6}{\partial z^4 \partial x^2} p + \frac{\partial}{\partial x} A \frac{\partial^5}{\partial z^4 \partial x} p + \frac{\partial^5}{\partial z^4 \partial x} A \frac{\partial}{\partial x} p + \frac{\partial^4}{\partial z^4} A \frac{\partial^2}{\partial x^2} p \right], \quad (\text{B.4}) \end{aligned}$$

$$\begin{aligned} \xi_2 = & \left( \frac{2b-2}{3} \right) \left[ \frac{\partial^5}{\partial z^2 \partial x^3} p \frac{\partial}{\partial x} B + \frac{\partial^4}{\partial z \partial x^3} p \frac{\partial^2}{\partial z \partial x} B + \frac{\partial^4}{\partial z \partial x^3} B \frac{\partial^2}{\partial z \partial x} p + \frac{\partial^3}{\partial z^2 \partial x} p \frac{\partial^3}{\partial x^3} B \right] \\ & - \frac{1}{30} \left[ \frac{9}{64} \frac{\partial}{\partial z} p \frac{\partial^5}{\partial z^5} B + \frac{45}{64} \frac{\partial^4}{\partial z^4} B \frac{\partial^2}{\partial z^2} p + \frac{\partial^5}{\partial z^5} p \frac{\partial}{\partial z} B \right] \\ & - \frac{3}{64} \left[ \frac{\partial^3}{\partial z^3} p \frac{\partial^3}{\partial z^3} B + \frac{\partial^4}{\partial z^4} p \frac{\partial^2}{\partial z^2} B \right] + (b-1) \left[ \frac{\partial^3}{\partial z \partial x^2} p \frac{\partial^3}{\partial z \partial x^2} B + \frac{\partial^4}{\partial z^2 \partial x^2} p \frac{\partial^2}{\partial x^2} B \right] \\ & + \left( \frac{b-1}{6} \right) \left[ B \frac{\partial^6}{\partial z^2 \partial x^4} p + \frac{\partial}{\partial z} B \frac{\partial^5}{\partial z \partial x^4} p + \frac{\partial^5}{\partial z \partial x^4} B \frac{\partial}{\partial z} p + \frac{\partial^4}{\partial x^4} B \frac{\partial^2}{\partial z^2} p \right]. \quad (\text{B.5}) \end{aligned}$$

## References

- [1] J. Virieux, S. Operto, An overview of full-waveform inversion in exploration geophysics, *Geophysics* 74 (6) (2009) WCC1–WCC26.
- [2] C.-H. Jo, C. Shin, J.H. Suh, An optimal 9-point, finite-difference, frequency-space, 2-D scalar wave extrapolator, *Geophysics* 61 (2) (1996) 529–537.
- [3] Z. Chen, D. Cheng, W. Feng, T. Wu, An optimal 9-point finite difference scheme for the Helmholtz equation with PML, *Int. J. Numer. Anal. Model.* 10 (2) (2013).
- [4] C. Shin, H. Sohn, A frequency-space 2-D scalar wave extrapolator using extended 25-point finite-difference operator, *Geophysics* 63 (1) (1998) 289–296.
- [5] T. Wu, A dispersion minimizing compact finite difference scheme for the 2D Helmholtz equation, *J. Comput. Appl. Math.* 311 (2017) 497–512.
- [6] J.-B. Chen, A generalized optimal 9-point scheme for frequency-domain scalar wave equation, *J. Appl. Geophys.* 92 (2013) 1–7.
- [7] M. Nabavi, M.K. Siddiqui, J. Dargahi, A new 9-point sixth-order accurate compact finite-difference method for the Helmholtz equation, *J. Sound Vib.* 307 (3–5) (2007) 972–982.
- [8] F. Ihlenburg, I. Babuška, Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation, *Internat. J. Numer. Methods Engrg.* 38 (22) (1995) 3745–3774.
- [9] F. Ihlenburg, I. Babuška, Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM, *Comput. Math. Appl.* 30 (9) (1995) 9–37.
- [10] Z. Chen, D. Cheng, T. Wu, A dispersion minimizing finite difference scheme and preconditioned solver for the 3D Helmholtz equation, *J. Comput. Phys.* 231 (24) (2012) 8152–8175.
- [11] J.-P. Berenger, A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.* 114 (2) (1994) 185–200.
- [12] E. Turkel, A. Yefet, Absorbing PML boundary layers for wave-like equations, *Appl. Numer. Math.* 27 (4) (1998) 533–557.
- [13] I. Singer, E. Turkel, A perfectly matched layer for the Helmholtz equation in a semi-infinite strip, *J. Comput. Phys.* 201 (2) (2004) 439–465.
- [14] R.G. Pratt, M. Worthington, Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method, *Geophys. Prospect.* 38 (3) (1990) 287–310.
- [15] P.M. De Zeeuw, Matrix-dependent prolongations and restrictions in a blackbox multigrid solver, *J. Comput. Appl. Math.* 33 (1) (1990) 1–27.
- [16] Y.A. Erlangga, C.W. Oosterlee, C. Vuik, A novel multigrid based preconditioner for heterogeneous Helmholtz problems, *SIAM J. Sci. Comput.* 27 (4) (2006) 1471–1492.
- [17] M.B. van Gijzen, Y.A. Erlangga, C. Vuik, Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian, *SIAM J. Sci. Comput.* 29 (5) (2007) 1942–1958.
- [18] D. Cheng, X. Tan, T. Zeng, A dispersion minimizing finite difference scheme for the Helmholtz equation based on point-weighting, *Comput. Math. Appl.* 73 (11) (2017) 2345–2359.
- [19] I. Harari, E. Turkel, Accurate finite difference methods for time-harmonic wave propagation, *J. Comput. Phys.* 119 (2) (1995) 252–270.
- [20] I. Singer, E. Turkel, High-order finite difference methods for the Helmholtz equation, *Comput. Methods Appl. Mech. Engrg.* 163 (1–4) (1998) 343–358.
- [21] E. Turkel, D. Gordon, R. Gordon, S. Tsynkov, Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number, *J. Comput. Phys.* 232 (1) (2013) 272–287.
- [22] T. Wu, R. Xu, An optimal compact sixth-order finite difference scheme for the Helmholtz equation, *Comput. Math. Appl.* 75 (7) (2018) 2520–2537.
- [23] Y. Zeng, J. He, Q. Liu, The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media, *Geophysics* 66 (4) (2001) 1258–1266.
- [24] J.W. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*, Vol. 22, Springer Science & Business Media, 2013.
- [25] H. Ren, H. Wang, T. Gong, Seismic modeling of scalar seismic wave propagation with finite-difference scheme in frequency-space domain, *Geophys. Prospect.* *Petroleum* 48 (2009) 20–26.
- [26] L.N. Trefethen, Group velocity in finite difference schemes, *SIAM Rev.* 24 (2) (1982) 113–136.
- [27] T.A. Davis, I.S. Duff, An unsymmetric-pattern multifrontal method for sparse LU factorization, *SIAM J. Matrix Anal. Appl.* 18 (1) (1997) 140–158.
- [28] T.A. Davis, Algorithm 832: UMPACK V4. 3—an unsymmetric-pattern multifrontal method, *ACM Trans. Math. Softw.* 30 (2) (2004) 196–199.