

# 基于预条件共轭残差法截断牛顿法的全波形反演

曾钵祁, 何清龙\*

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2023年10月21日; 录用日期: 2023年11月22日; 发布日期: 2023年11月29日

## 摘要

全波形反演(FWI)是一个大规模的非线性不适定问题, 其二阶梯度信息有着重要的作用, 但牛顿型方法需要计算量和存储量巨大。本文基于共轭残差方法和信赖域全局化策略提出了一种高效的截断牛顿全波形反演方法, 该全波形反演方法能够充分利用目标泛函的二阶梯度信息, 从而提高反演精度, 为了加速共轭残差法的收敛速度和计算效率, 本文给出了预处理的共轭残差方法并给出了其相关性质。基于二维2004 BP模型和Sigsbee模型, 验证了预处理共轭残差截断牛顿反演方法的有效性。数值结果表明预处理共轭残差截断牛顿法能充分利用二阶梯度信息, 从而加速算法收敛速度和提高成像精度。

## 关键词

全波形反演, 共轭残差法, 截断牛顿法

# Full Waveform Inversion Based on Preconditioned Conjugate Residual Truncated Newton Method

Boqi Zeng, Qinglong He\*

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

\* 通讯作者。

Received: Oct. 21<sup>st</sup>, 2023; accepted: Nov. 22<sup>nd</sup>, 2023; published: Nov. 29<sup>th</sup>, 2023

## Abstract

Full waveform inversion (FWI) is a large-scale nonlinear ill-posed problem, and its second-order gradient information plays an important role. However, the implementation of Newton-type method is expensive. In this paper, an efficient truncated Newton full waveform inversion method is proposed based on the conjugate residual method. The full waveform inversion method can make full use of the second-order gradient information of the target functional and improve the inversion accuracy. Aiming at the problem that the truncated Newton method depends on the initial value selection, this paper combines the trust region globalization strategy into the truncated Newton method. In order to accelerate the convergence rate of the conjugate residual method, the preprocessing operator is introduced in this paper. Based on the two-dimensional 2004 BP model and the Sigsbee model, the effectiveness of the preconditioned conjugate residual truncated Newton inversion method is verified. The numerical results show that the preconditioned conjugate residual truncated Newton method can make full use of the second-order gradient information, thus accelerating the convergence speed of the algorithm and improving the imaging accuracy.

## Keywords

Full Waveform Inversion, Conjugate Residual Method, Truncated Newton Method

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

全波形反演(Full Waveform Inversion, 简称为FWI)通过拟合波场模拟数据和实际观测数据, 以获取刻画地震波传播相关的物性参数(如波速、密度、衰减和各向异性等)。该方法充分利用了波场的全部信息内容, 显著的提高地震成像的分辨率 [1-6]。

FWI问题是一个大规模不适定的强非线性反问题。该反问题常转化为一个非线性最小二乘问

题, 因此常常使用基于局部收敛的算法进行求解。最速下降方法是最小化目标函数的一种最简单的优化方法 [7], 其下降方向为目标函数的负梯度方向, 因此收敛速度往往较慢。牛顿法是基于目标函数的二阶梯度(Hessian 矩阵)的一种快速收敛的局部收敛型优化方法, 相较于一阶梯度型方法其具有较快的收敛速度 [8], 尤其是在最优值附近时, 收敛速度是二次的。然而, 针对FWI大规模反演问题, 牛顿方法的Hessian矩阵巨大的存储需求和密集的大规模计算成本限制了其在FWI问题中的应用。拟牛顿方法通过存储之前迭代的参数模型和梯度变化信息来近似Hessian矩阵或Hessian矩阵的逆, 在改进了最速下降法型方法收敛速度较慢的同时减少了Hessian矩阵的大规模计算难题, 例如L-BFGS方法 [9]。然而, 随着油气勘探地质条件难度的增加, 基于Hessian信息的牛顿型方法在解决该类复杂FWI反问题时的优势正逐渐凸显。目标函数的二阶梯度(Hessian矩阵)能补偿有限的震源照明提高深度结构的反演精度, 其次该类方法还能充分利用波场的多次散射信息, 可以有效地提高复杂地质结构的重建分辨率 [10]。

截断牛顿型方法是一类能充分利用目标函数Hessian信息, 同时避免Hessian矩阵的巨大存储和计算量需求的高效优化方法。截断牛顿型方法的主要思想为利用迭代法近似求解牛顿方法中的大规模牛顿方程, 因此有效地避免了Hessian矩阵的直接计算, 有效地提高了算法的计算效率。从全局化策略方面可以将截断牛顿法分为两类, 一类为基于线搜索的截断牛顿型反演方法, 一类为基于信赖域全局化策略的截断牛顿型方法。相较于线搜索的截断牛顿型方法, 基于信赖域的截断牛顿方法在处理高度非凸和非线性问题上具有较好的计算表现, 同时基于信赖域全局化策略的截断牛顿方法还具有隐式正则性的作用, 因此更适用于解决FWI反演问题 [11]。

共轭残差法(CR)是一种求解对称线性方程组的迭代型方法, 其由Hestenenel和Stiefel提出 [12]。CR的主要思想是极小化残差, 其在求解大规模对称线性方程组问题时具有良好的表现。Fong和Saunders确定了当矩阵为对称正定时, 二次模型的函数值随CR迭代呈单调递减且迭代近似解的欧式范数也随迭代步成单调递减, CR的单调性使得其非常适合于求解信赖域方法中的信赖域子问题 [13]。

本文基于共轭残差线性方程组求解方法和信赖域全局化策略建立了截断牛顿全波形反演方法。为了进一步加快方法的收敛速度, 本文引入预处理算子并给出了预处理的截断共轭残差法, 提出了预处理信赖域截断牛顿算法, 简记为PCRTRUST-NEWTON方法。考虑到共轭残差算法的重构效率高度依赖于截断条件的选取, 给出了一种有效的截断条件。同时, 基于频率域二维声波方程, 给出数值实验以验证PCRTRUST-NEWTON方法的有效性和高效性; 本文对PCRTRUST-NEWTON方法与有限内存的拟牛顿法(简记为L-BFGS)进行对比实验, 以显示PCRTRUST-NEWTON方法在计算效率和重构精度方面的优越性。

## 2. 理论与方法

### 2.1. 信赖域全局化策略的截断牛顿法

理论上, FWI可以表示为最小二乘优化问题

$$\min_m f(m) = \frac{1}{2} \sum_{s=1}^{N_s} \|Ru_s(m) - d_s\|_2^2, \quad (1)$$

其中,  $N_s$ 为源的个数,  $\|\cdot\|_2$ 表示 $L_2$ 范数,  $R$ 表示每个源 $q_s$ 接收位置的波场,  $m$ 是反演参数,  $d_s$ 为观测数据。  $u_s(m)$ 是合成波场且满足微分方程

$$A(m)u_s = q_s, \quad (2)$$

其中,  $A(m)$ 为微分算子,  $m$ 为介质参数,  $u_s$ 表示对应震源为 $q_s$ 的响应波场。

牛顿法的基本思想是在点 $m_k$ 处对方程(1)进行二次逼近

$$\phi(m_k + \Delta m_k) = f(m_k) + \Delta m_k^T g_k + \frac{1}{2} \Delta m_k^T H_k \Delta m_k, \quad (3)$$

其中,  $g_k = \nabla f(m_k)$ 为目标函数在点 $m_k$ 处的梯度,  $H_k = \nabla^2 f(m_k)$ 表示目标函数的Hessian矩阵。将二次近似的最小值作为函数 $f(m)$ 的最小值, 可以得到

$$H_k \Delta m_k = -g_k, \quad (4)$$

该方程(4)称为牛顿方程, 满足该方程的 $\Delta m_k$ 为牛顿方向, 因此牛顿法的迭代公式为

$$m_{k+1} = m_k + \Delta m_k. \quad (5)$$

FWI是一个大规模非线性问题, 它的Hessian矩阵储存和使用矩阵分解法直接求解都较为困难。因此采用迭代法近似求解牛顿方程较合理。这种利用迭代法近似牛顿方程的方法称为截断牛顿法。

在牛顿方法中, 使用二阶近似来近似目标函数 $f$ , 但只有在 $m$ 的一个合理邻域中才能很好地近似 $f$ 。因此, 引入了信赖域的策略, 增加二次模型对原目标函数的约束, 该二次模型为

$$\min_{\Delta m \in \mathbb{R}^n} \phi(m_k + \Delta m_k) = f(m_k) + \Delta m_k^T g_k + \frac{1}{2} \Delta m_k^T H_k \Delta m_k, \quad (6a)$$

$$s.t. \|\Delta m_k\|_2 \leq \Delta_k, \Delta_k > 0, \quad (6b)$$

其中,  $\Delta_k$ 是信赖域半径, 在这个区域内 $\phi(m_k + \Delta m_k)$ 可以很好的近似函数 $f(m)$ 。方程(6a)和(6b)即为著名的信赖域子问题。

信赖区域半径的选择是一个关键问题, 它与算法的收敛性有关。引入如下的比率指标来度量近似的程度

$$\rho_k = \frac{f(m_k) - f(m_k + \Delta m_k)}{\phi(m_k) - \phi(m_k + \Delta m_k)}, \quad (7)$$

比值参数是目标函数值的实际下降量与预估下降量(即二次近似模型的下降量)的比值。如果 $\rho_k$ 相对较大, 则表示二次模型可以很好地逼近目标函数, 则可以接受近似更新量 $\Delta m_k$ , 并适当地扩大信赖域半径 $\Delta_k$ 。相反地, 应拒绝近似更新量 $\Delta m_k$ , 并减小信赖域半径 $\Delta_k$ , 并重新计算近似更新

量 $\Delta m_k$ 。算法1所示为信赖域全局化策略的截断牛顿法的算法描述。

**算法 1.** 信赖域全局化策略的截断牛顿法

- 1: 初始化: $\eta_0 < \eta_1 < \eta_2 \leq 1, 0 < \sigma_1 < \sigma_2 < 1 < \sigma_3, f(m_0), k := 0$
- 2: **while**  $k < max\_iter$  **do**
- 3:      $g_k := \nabla f(m_k)$
- 4:     **if**  $\|g_k\|_2 < \epsilon_{ga} + \epsilon_{gr} \|g_0\|_2$  **then**
- 5:          $m := m_k$ , break
- 6:     **end if**
- 7:     计算Hessian矩阵 $H_k$
- 8:     根据算法2计算 $\Delta m_k$
- 9:     根据公式(7)计算下降率 $\rho_k$
- 10:     更新 $m_k$ :

$$m_{k+1} = \begin{cases} m_k + \Delta m_k, & \rho_k > \eta_0 \\ m_k, & \rho_k \leq \eta_0 \end{cases}$$

- 11:     更新信赖域半径 $\Delta_{k+1}$

$$\Delta_{k+1} = \begin{cases} \min((\max(\alpha_k^* \sigma_1)) \|\Delta m_k\| \sigma_2 \Delta_k), & \rho_k < \eta_0 \\ \max(\sigma_{1\Delta_k} \min(\alpha_k^* \|\Delta m_k\| \sigma_2 \Delta_k)), & \rho_k \in [\eta_0, \eta_1] \\ \max \sigma_{1\Delta_k} \min(\alpha_k^* \|\Delta m_k\| \sigma_3 \Delta_k), & \rho_k \in (\eta_1, \eta_2) \\ \max(\Delta_k \min(\alpha_k^* \|\Delta m_k\| \sigma_3 \Delta_k)), & \rho_k \geq \eta_2 \end{cases}$$

$$\alpha_k^* = \frac{-\nabla f(m_k)^T \Delta m_k}{2(f((m_k + \Delta m_k) - f(m_k) - \nabla f(m_k)^T \Delta m_k))}$$

- 12:     **if**  $k == max\_iter$  or  $f(m_{k+1}) < \epsilon_f f(m_0)$  **then**
- 13:          $m := m_{k+1}$ , break
- 14:     **end if**
- 15:      $k := k + 1$
- 16: **end while**
- 17: 返回 $m$

## 2.2. 预条件截断共轭残差法

共轭残差法(Conjugate residual method, 简记为CR)的单调性质使其在信赖域全局化策略的截断牛顿法中极为有吸引力。本小节主要介绍用CR 来求解信赖域子问题(6a)。

首先忽略信赖域的约束, 求解信赖域子问题(6a)相当于求解如下问题

$$\min_{m \in \mathbb{R}^n} \varphi(m) = g^T m + \frac{1}{2} m^T H m, \tag{8}$$

其中,  $g = g_k, H = H_k, m = \Delta m_k$ 。设该目标函数的解为 $\Delta m_k$ , 则利用一阶优化方法取最小值的必要条件可知

$$H \Delta m_k = -g.$$

CR 方法将求解上述线性方程组转化为求解如下优化问题

$$\min_{m \in \mathbb{R}^n} \psi(m) = \|Hm + g\|_2.$$

CR的迭代格式如下所示

$$\alpha_k = \frac{r_{k-1}^T H r_{k-1}}{p_{k-1}^T H^2 p_{k-1}}, \quad (9)$$

$$m_k = m_{k-1} + \alpha_k p_{k-1}, \quad (10)$$

$$r_k = r_{k-1} - \alpha_k H p_{k-1}, \quad (11)$$

$$p_k = r_k + \beta_k p_{k-1}, \quad (12)$$

$$\beta_k = \frac{r_k^T H r_k}{r_{k-1}^T H r_{k-1}}, \quad (13)$$

具体推导过程可参考 [14, 15]。

为了提高CR算法的计算效率, 基于pseudo-Hessian对Hessian矩阵对角线元素的良好近似, 构建了对角阵预处理算子 [16]

$$M_k = \text{diag}\left(\sum_{s=0}^{N_s} \left(\frac{\partial A(m_k)}{\partial m} u_s\right)^* \left(\frac{\partial A(m_k)}{\partial m} u_s\right)\right). \quad (14)$$

忽略下标 $k$ , 因为对于每一个信赖域子问题而言,  $M_k$ 是相同的。因此记为 $M$ ,  $M$ 是一个对称正定矩阵。令 $M = C^2 = M^T$ ,  $\tilde{H} \triangleq C^{-1} H C^{-1}$ ,  $\tilde{g} = C^{-1} g$ ,  $m_k = C^{-1} \tilde{m}_k$ 。则预处理的CR算法推导如下

$$\tilde{r}_k = -\tilde{g} - \tilde{H} \tilde{m}_k = -C^{-1} g - C^{-1} H C^{-1} \tilde{m}_k = C^{-1} (-g - H m_k) = C^{-1} r_k$$

记 $M z_k = r_k$ ,  $u_k = H z_k$ ,  $q_k = C \tilde{q}_k$ ,  $p_k = C \tilde{p}_k$  于是

$$\tilde{\alpha}_k = \tilde{\zeta}_{k-1} / \|\tilde{q}_{k-1}\|^2 = \tilde{\zeta}_{k-1} / \|C^{-1} q_k\|^2 = \tilde{\zeta}_{k-1} / \langle M^{-1} q_k, q_k \rangle$$

$$\tilde{m}_k = \tilde{m}_{k-1} + \tilde{\alpha}_k \tilde{p}_{k-1} \Rightarrow C m_k = C m_{k-1} + \tilde{\alpha}_k C^{-1} p_k \Rightarrow m_k = m_{k-1} + \tilde{\alpha}_k M^{-1} p_{k-1}$$

$$\tilde{r}_k = \tilde{r}_{k-1} - \tilde{\alpha}_k \tilde{q}_{k-1} \Rightarrow C^{-1} r_k = C^{-1} r_{k-1} - \tilde{\alpha}_k C^{-1} q_{k-1} \Rightarrow r_k = r_{k-1} - \tilde{\alpha}_k q_{k-1}$$

$$\tilde{\zeta}_k = \tilde{r}_k^T \tilde{u}_k \Rightarrow \tilde{\zeta}_k = (C^{-1} r_k)^T (C^{-1} u_k) = r_k^T M^{-1} u_k = (M^{-1} r_k)^T u_k \Rightarrow \tilde{\zeta}_k = z_k^T u_k$$

$$\tilde{u}_k = \tilde{H} \tilde{r}_k \Rightarrow \tilde{u}_k = C^{-1} H C^{-1} \tilde{r}_k \Rightarrow \tilde{u}_k = C^{-1} H z_k \Rightarrow \tilde{u}_k = C^{-1} H z_k = C^{-1} u_k \Rightarrow u_k = H z_k$$

$$\tilde{\beta}_k = \frac{z_k^T u_k}{z_{k-1}^T u_{k-1}}$$

$$\tilde{p}_k = \tilde{r}_k + \tilde{\beta}_k \tilde{p}_{k-1} \Rightarrow C^{-1} p_k = C^{-1} r_k + \tilde{\beta}_k C^{-1} p_{k-1} \Rightarrow p_k = r_k + \tilde{\beta}_k p_{k-1}$$

$$\tilde{q}_k = \tilde{u}_k + \tilde{\beta}_k \tilde{q}_{k-1} \Rightarrow C^{-1} q_k = C^{-1} u_k + \tilde{\beta}_k C^{-1} q_{k-1} \Rightarrow q_k = u_k + \tilde{\beta}_k q_{k-1}$$

定理1总结了我们经常使用的CR的性质。

**定理1.** 若 $H$ 为正定矩阵, 且对于所有的 $i = 0, \dots, k$ 都有 $r_i \neq 0$ , 若 $\langle p_i, p_j \rangle_{M^{-1}} = p_i^T M^{-1} p_j$ , 则下列性质成立:

$$q_i = HM^{-1}p_i, \quad (15)$$

$$\rho_i = r_i^T M^{-1} r_i = \langle r_i, r_i \rangle_{M^{-1}} = \|r_i\|_{M^{-1}}, \quad (16)$$

$$\zeta_i = r_i^T M^{-1} q_i = \langle r_i, q_i \rangle = r_i^T M^{-1} HM^{-1} p_i = z_i^T z_i = r_i^T M^{-1} HM^{-1} r_i, \quad (17)$$

$$\langle HM^{-1} p_i, HM^{-1} p_j \rangle_{M^{-1}} = \langle q_i, q_j \rangle_{M^{-1}} = 0 \quad i \neq j \quad i, j = 0, \dots, k \quad (18)$$

$$z_i^T H z_i = r_i^T M^{-1} HM^{-1} r_j = 0 \quad i \neq j \quad i, j = 0, \dots, k \quad (19)$$

$$p_i^T M^{-1} HM^{-1} r_j = (M^{-1} p_i)^T H (M^{-1} r_j) = 0 \quad 0 \leq i < j \leq k \quad (20)$$

$$\|m_i\|_{M^{-1}} < \|m_j\|_{M^{-1}} \quad i < j \quad i, j = 0, \dots, k \quad (21)$$

$$\alpha_i > 0 \quad i = 0, \dots, k \quad (22)$$

$$\langle p_i, r_j \rangle_{\mu^{-1}} > 0 \quad 0 \leq i, j \leq k \quad (23)$$

证明. 性质(18)-(23)可以由线性系统 $\tilde{H}\tilde{m} = -\tilde{g}$ 的CR和利用 $\tilde{r}_k = C^{-1}r_k$ ,  $\tilde{q}_k = C^{-1}q_k$ ,  $\tilde{p}_k = C^{-1}p_k$ ,  $M^{-1} = C^{-1}C^{-1}$ 得到. 因此我们只需证(15)-(17)。

由归纳假设法, 可知, 对于 $k = 0$ ,  $q_0 = u_0 = Hz_0 = HM^{-1}r_0 = HM^{-1}p_0$ , 性质(15)成立. 假设对于 $k$ 性质(15)成立, 则对于 $k + 1$ 有

$$\begin{aligned} q_{k+1} &= u_{k+1} + \beta_{k+1}q_k = Hz_{k+1} + \beta_{k+1}HM^{-1}p_k \\ &= HM^{-1}r_{k+1} + \beta_{k+1}HM^{-1}p_k \\ &= HM^{-1}(r_{k+1} + \beta_{k+1}p_k) \\ &= HM^{-1}p_{k+1} \end{aligned}$$

性质(15)成立。

由性质(20)可得

$$\begin{aligned} \zeta_k &= r_k^T M^{-1} q_k = r_k^T M^{-1} HM^{-1} p_k \\ &= r_k^T M^{-1} HM^{-1} (r_k + \beta_k p_{k-1}) \\ &= r_k^T M^{-1} HM^{-1} r_k + \beta_k r_k^T M^{-1} HM^{-1} p_{k-1} \\ &= r_k^T M^{-1} HM^{-1} r_k \end{aligned}$$

性质(17)成立。

由性质(20)和性质(17)可得

$$\begin{aligned}
 \rho_{k+1} &= r_{k+1}^T M^{-1} r_{k+1} = r_{k+1}^T M^{-1} (r_k - \alpha_{k+1} q_k) \\
 &= r_{k+1}^T M^{-1} r_k - \alpha_{k+1} r_{k+1}^T M^{-1} q_k \\
 &= r_{k+1}^T M^{-1} r_k - \alpha_{k+1} r_{k+1}^T M^{-1} H M^{-1} p_k \\
 &= r_{k+1}^T M^{-1} r_k = (r_k - \alpha_{k+1} q_k)^T M^{-1} r_k \\
 &= r_k^T M^{-1} r_k - \alpha_{k+1} q_k^T M^{-1} r_k \\
 &= \rho_k - \alpha_{k+1} \zeta_k
 \end{aligned}$$

性质(16)成立。 □

算法2所示为预条件截断信赖域CR方法的算法描述, 截断CR 相关细节请参考文献 [15]。

### 算法 2. 预条件截断CR方法

- 1: 输入:  $H, g, \Delta_k > 0, \tau_a > 0, \tau_r > 0$
- 2: 初始化:  $k = 0, m_0 = 0, r_0 = -g, Mz_0 = r_0, u_0 = Hz_0, \zeta_0 = z_0^T u_0, q_0 = u_0, p_0 = r_0, \rho_0 = r_0^T z_0, \nu_0 = \sqrt{\rho_0}, \delta_0 = \zeta_0, \mu_0 = \rho_0$
- 3: **while**  $\nu_k > \tau_a + \tau_r \|g_0\|_{M^{-1}}$  **do**
- 4:      $k \leftarrow k + 1$
- 5:     利用算法3计算 $\alpha_k$ 和 $p_{k-1}$
- 6:      $m_k = m_{k-1} + \alpha_k p_{k-1}$
- 7:      $r_k = r_{k-1} - \alpha_k q_{k-1}$
- 8:      $Mz_k = r_k$
- 9:      $\rho_k = \rho_{k-1} - \alpha_k \zeta_{k-1}$   $\triangleright \rho_k = r_k^T M^{-1} r_k$
- 10:      $\nu_k = \sqrt{\rho_k}$
- 11:      $u_k = Hz_k$
- 12:      $\zeta_k = z_k^T u_k$   $\triangleright \zeta_k = r_k^T M^{-1} q_k = z_k^T q_k = r_k^T M^{-1} H M^{-1} r_k$
- 13:      $\beta_k = \zeta_k / \zeta_{k-1} = z_k^T u_k / z_{k-1}^T u_{k-1}$
- 14:      $p_k = r_k + \beta_k p_{k-1}$
- 15:      $q_k = u_k + \beta_k q_{k-1}$   $\triangleright q_k = H M^{-1} p_k$
- 16:      $\mu_k = \rho_k + \beta_k (\mu_{k-1} - \alpha_k \delta_{k-1})$   $\triangleright \mu_k = p_k^T r_k$
- 17:      $\delta_k = \zeta_k + \beta_k^2 \delta_{k-1}$   $\triangleright \delta_k = p_k^T H p_k$
- 18: **end while**
- 19: 返回 $m_k$

### 2.3. 截断条件

截断条件的选择是影响非牛顿精确算法计算效率和精度的关键因素之一。对于内部CR迭代的截断条件 $\tau_a + \tau_r \|g_0\|_{M^{-1}}$ , 选取有效的截断条件

$$\tau_a = 0 \quad \tau_r = \min(1/j, \sqrt{\|g_j\|}), \quad (24)$$

这里的 $j$ 为外层信赖域牛顿法迭代数。



**算法 3.** 算法2的步长计算

- 1: 输入:  $\Delta_k > 0, m_{k-1}, r_{k-1}, \zeta_{k-1}, q_{k-1}, p_{k-1}, \rho_{k-1}, \nu_{k-1}, \delta_{k-1}, \mu_{k-1}$
- 2: **if**  $\delta_{k-1} = 0$  **and**  $\mu_{k-1} = 0$  **then**
- 3:     更新  $\alpha_k$  和  $p_{k-1}$ 

$$\begin{cases} p_{k-1} \leftarrow r_{k-1}, \alpha_k \leftarrow \min(\alpha_r, \alpha_a) & \text{if } \zeta_{k-1} > 0 \\ p_{k-1} \leftarrow r_{k-1}, \alpha_k \leftarrow \alpha_r & \text{if } \zeta_{k-1} < 0 \end{cases}$$
- 4: **else if**  $\delta_{k-1} = 0$  **and**  $\mu_{k-1} \neq 0$  **then**
- 5:     计算  $\alpha_{r^*}$  和  $\alpha_{p^*}$ 

$$\begin{cases} \alpha_{r^*} = \min(\alpha_r, \alpha_a) & \text{if } \zeta_{k-1} > 0 \\ \alpha_{r^*} = \alpha_r & \text{if } \zeta_{k-1} < 0 \\ \alpha_{p^*} = \alpha_{p^+} & \text{if } \mu_{k-1} > 0 \\ \alpha_{p^*} = \alpha_{p^-} & \text{if } \mu_{k-1} < 0 \end{cases}$$
- 6:     计算  $\xi_k = -\alpha_{p^*} \mu_{k-1} + \alpha_{r^*} \rho_{k-1} - \frac{1}{2} \alpha_{r^*}^2 \zeta_{k-1}$
- 7:     更新  $\alpha_k$  和  $p_{k-1}$ 

$$\begin{cases} p_{k-1} \leftarrow r_{k-1}, \alpha_k \leftarrow \alpha_{r^*} & \text{if } \xi_k > 0 \\ p_{k-1} \leftarrow r_{k-1}, \alpha_k \leftarrow \alpha_{p^*} & \text{if } \xi_k < 0 \end{cases}$$
- 8: **else if**  $\delta_{k-1} > 0$  **and**  $\zeta_{k-1} > 0$  **then**
- 9:      $\alpha_k = \zeta_{k-1} / \|q_k\|_{M^{-1}}^2 = \zeta_{k-1} / \langle M^{-1}q_k, q_k \rangle$
- 10:     **if**  $\alpha_k \geq \alpha_{p^+}$  **then**
- 11:         更新  $\alpha_k \leftarrow \alpha_{p^+}$
- 12:     **end if**
- 13: **else**
- 14:     计算  $\alpha_{r^*}$  和  $\alpha_{p^*}$ ,  $\alpha_{r^*}$  的计算同上
$$\begin{cases} \alpha_{p^*} = \alpha_{p^+} & \text{if } \delta_{k-1} < 0, \mu_{k-1} > 0 \\ \alpha_{p^*} = \alpha_{p^-} & \text{if } \delta_{k-1} < 0, \mu_{k-1} < 0 \\ \alpha_{p^*} = \min(\alpha_{p^+}, \alpha_b) & \text{if } \delta_{k-1} > 0, \zeta_{k-1} > 0 \\ \alpha_{p^*} = \min(\alpha_{p^-}, \alpha_b) & \text{if } \delta_{k-1} > 0, \zeta_{k-1} < 0 \end{cases}$$
- 15:     计算  $\xi_k = -\alpha_{p^*} \mu_{k-1} + \alpha_{r^*} \rho_{k-1} + \frac{1}{2} (\alpha_{p^*}^2 \delta_{k-1} - \alpha_{r^*}^2 \zeta_{k-1})$
- 16:     更新  $\alpha_k$  和  $p_{k-1}$ , 更新策略同上
- 17: **end if**
- 18: 返回  $\alpha_k, p_{k-1}$

### 3. 数值实验

本节基于2D频率域声波方程, 从数值计算的角度分析PCRTRUST-NEWTON方法的可行性和计算效率。方程(2)为声波方程, 具体形式为

$$\Delta u(m) + \frac{\omega^2}{m^2} u(m) = q_s(x, x_s), \quad (25)$$

其中,  $q_s$ 是在 $x_s$ 位置的源,  $\omega$ 是角频率,  $m$ 在这里为声速。

声波方程(25)采用混合网格差分形式进行离散 [17]。在离散过程中, 边界条件为完全匹配层吸收边界条件(PML), 降低边界反射对波场的影响 [18]。在本文的所有实验中, 所有边界条件均为PML吸收边界条件。离散线性方程组采用基于LU分解的直接法进行求解 [19]。

信赖域初始半径 $\Delta_0$ 设置为

$$\Delta_0 = \frac{\|\nabla f(m_0)\|_2}{10}$$

信任区域半径的相关参数设置为

$$\gamma_0 = 1.0e - 4 \quad \gamma_1 = 0.25 \quad \gamma_2 = 0.75 \quad \sigma_1 = 0.25 \quad \sigma_2 = 0.5 \quad \sigma_3 = 4$$

方法停止参数设置为

$$\epsilon_f = 1.0e - 10 \quad \epsilon_{ga} = \epsilon_{gr} = 1.0e - 10$$

本文基于2004 BP模型和Sigsbee模型进行数值实验以验证算法的收敛性和可行性, 并将其与L-BFGS方法进行对比以验证该算法在计算效率和重构精度方面的优势。

#### 3.1. 2004 BP模型

2004 BP模型是地震反演的基准模型之一 [20], 该模型的速度值变化在1 km/s到5 km/s之间。该模型沉积物与丘体之间存在高速度比, 且由于大部分能量由表层高速盐丘体所反射, 因此只有很少的能量可以通过这个高速盐体到达深部的沉积层, 这增加了高精度反演该模型的难度。该模型的大小为 $421 \times 212$  (包含10个PML网格), 空间采样步长为 $\Delta x = \Delta z = 20$  m。初始模型是对真实模型利用光滑化模糊而得。图1所示为模型的真实模型(a)和初始模型(b)。100个震源和400个接收点均匀分布在模型上表面, 震源间距为80 m, 接收点间距为20 m, 震源为单位脉冲震源。

在反演过程中, 选取10个离散频率2.5, 3.0, 4.0, 5.5, 6.5, 7.5, 8.5, 9.5, 10.0, 12.0 Hz。从低频到高频逐次进行反演, 每个频率最大迭代步数分别为14步(PCRTRUST-NEWTON)和164步(L-BFGS)。图2所示是频率为2.5 Hz时两种方法的收敛曲线, 其它频率也有类似的曲线。从图2可以看出, PCRTRUST-NEWTON方法相较于L-BFGS方法有着较高的收敛速度和收敛效率。图3(a)所示为PCRTRUST-NEWTON方法每个频率迭代14步后的重构结果、图3(b)所示为L-BFGS方法每个频率迭代164步后的重构结果。图4是不同速度模型在位置 $x = 1.56$  km (a)和 $x = 5.55$  km (b)处的纵向剖面。从图3(a)和图4可以看出, L-BFGS方法在浅层重构存在大量的异常波动, 这在

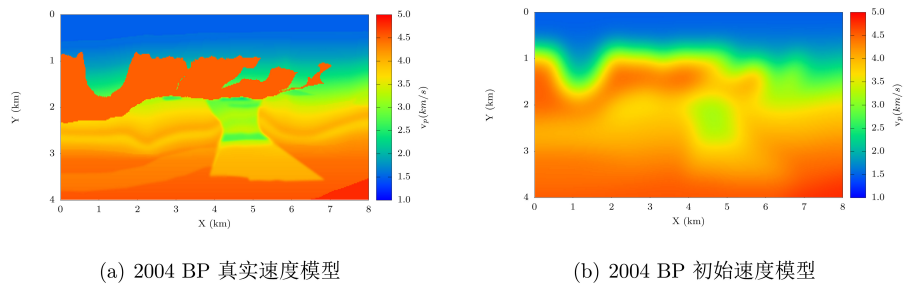


Figure 1. The 2004 BP speed model

图 1. 2004 BP 速度模型

位置  $x = 1.56$  km 处的纵向剖面可以更为明显的看出。而PCRTRUST-NEWTON方法不仅在浅层高速盐丘体重构表现良好, 与真实模型几乎完全匹配, 在中深部的沉积层重构效果也比L-BFGS方法的重构效果显著。这是由于PCRTRUST-NEWTON方法基于Hessian矩阵构造, 利用了多散射场, 提高了重构的精度。同时, Hessian矩阵的逆可以弥补震源照明不足的缺点, 提升了成像分辨率。

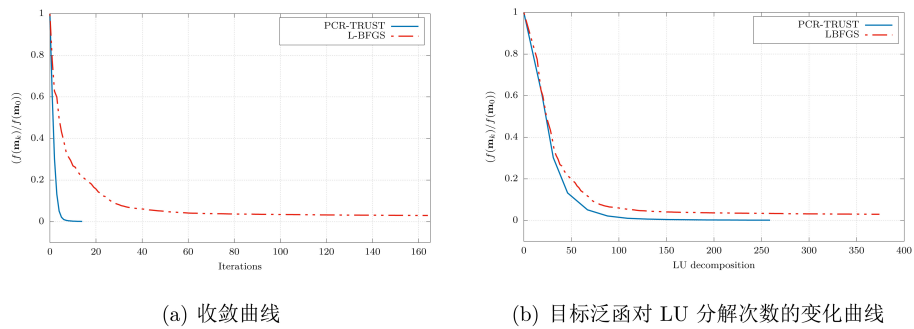


Figure 2. The objective function change curve of 2004 BP model

图 2. 2004 BP模型的目标函数变化曲线

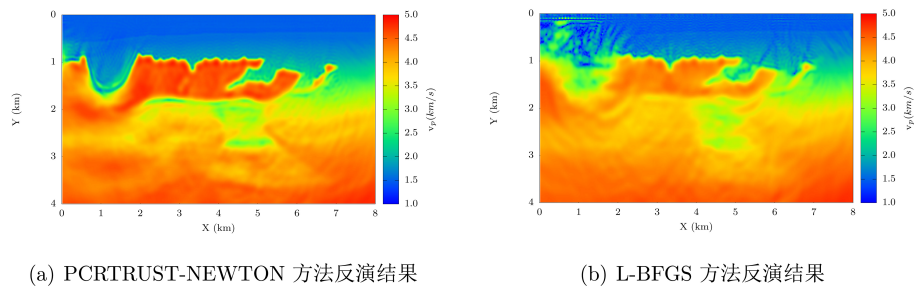


Figure 3. The inversion results of 2004 BP model

图 3. 2004 BP模型的反演结果

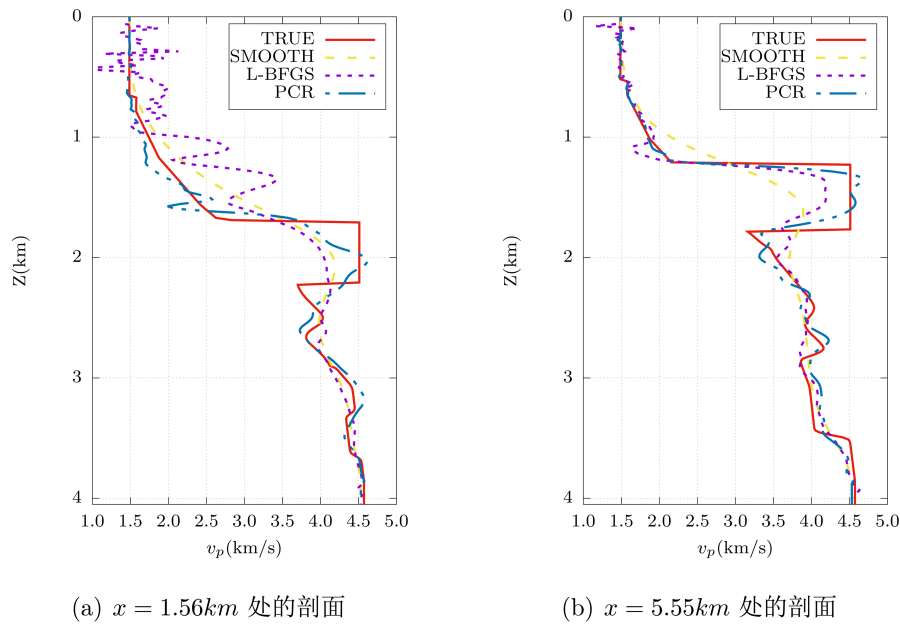


Figure 4. Longitudinal profile of 2004 BP model parameters

图 4. 2004 BP模型参数纵向剖面

### 3.2. Sigsbee模型

Sigsbee模型拥有复杂的盐丘形状, 且顶部构造分界面存在强烈的速度间断 [21]。由于高速盐丘与沉积物的高速度比以及沉积物的渐变结构, 因此会有很强的多次散射波场, 使得精确反演该模型比较困难。选取模型大小为 $411 \times 164$ , 空间采样步长为 $\Delta x = \Delta z = 15$  m。97个震源和390个接收点均匀分布在模型上表面, 震源间距为60 m, 接收点间距为15 m, 震源类型为单位脉冲震源。图 5所示为真实模型(a)和初始模型(b), 初始模型也采用光滑化真实模型得到。

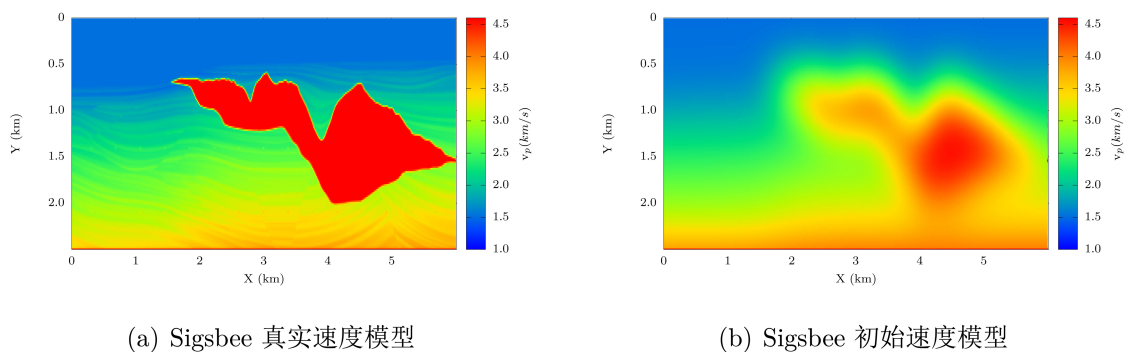


Figure 5. The Sigsbee speed model

图 5. Sigsbee 速度模型

反演过程中, 选取10个频率2.5, 3.0, 4.0, 5.5, 7.0, 9.0, 10.5, 12.0, 15.0, 17.0 Hz。采用从低频到高频的反演策略, PCRTRUST-NEWTON方法的每个频率最大迭代步数为14步, L-BFGS方法的每个频率最大迭代步数为164步。图 6所示是频率为2.5 Hz时两种方法的目标函数

变化曲线。可以看出, PCRTRUST-NEWYON方法收敛速度和计算效率均优于L-BFGS方法。图7(a)所示为PCRTRUST-NEWTON方法每个频率迭代14步后的重构结果、图7(b)所示为L-BFGS方法每个频率迭代164步后的重构结果。图8是不同速度模型在位置 $x = 2.73$  km (a)和 $x = 4.48$  km (b)处的纵向剖面。从图7和图8可以看出, L-BFGS方法反演在模型浅层存在较为明显的伪影, 而由于PCRTRUST-NEWTON方法可以很好的处理多重散射, 所以PCRTUST-NEWTON方法的浅层反演结果伪影较少。同时从图7可以看出, 在中深层结构中L-BFGS方法反演的结果更为模糊和存在较多的伪影, 从图8来看, 在两个位置的纵向剖面图中, L-BFGS方法反演结果都较为明显的偏离真实模型, 而PCRTRUST-NEWTON方法反演结果则更贴合与真实模型。这是由于PCRTRUST-NEWTON方法利用了二阶梯度信息, 所以该方法的中深层结构反演效果也优于L-BFGS方法。

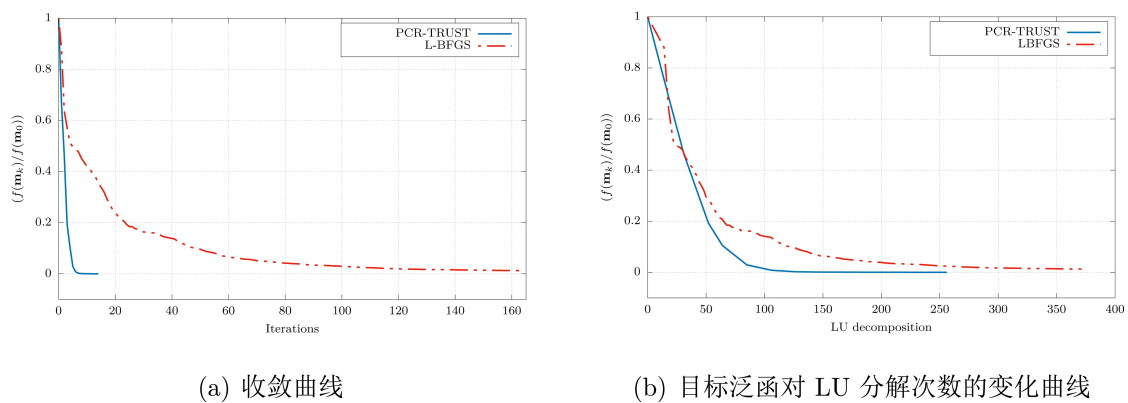


Figure 6. The objective function change curve of Sigsbee model

图 6. Sigsbee 模型的目标函数变化曲线

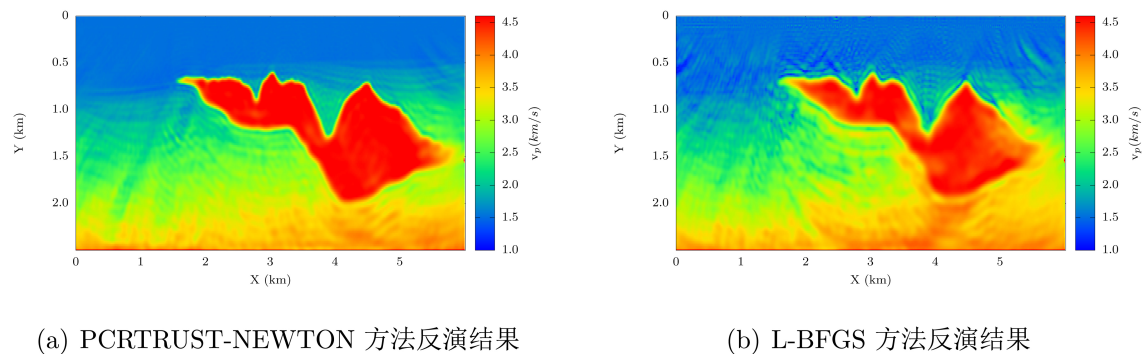
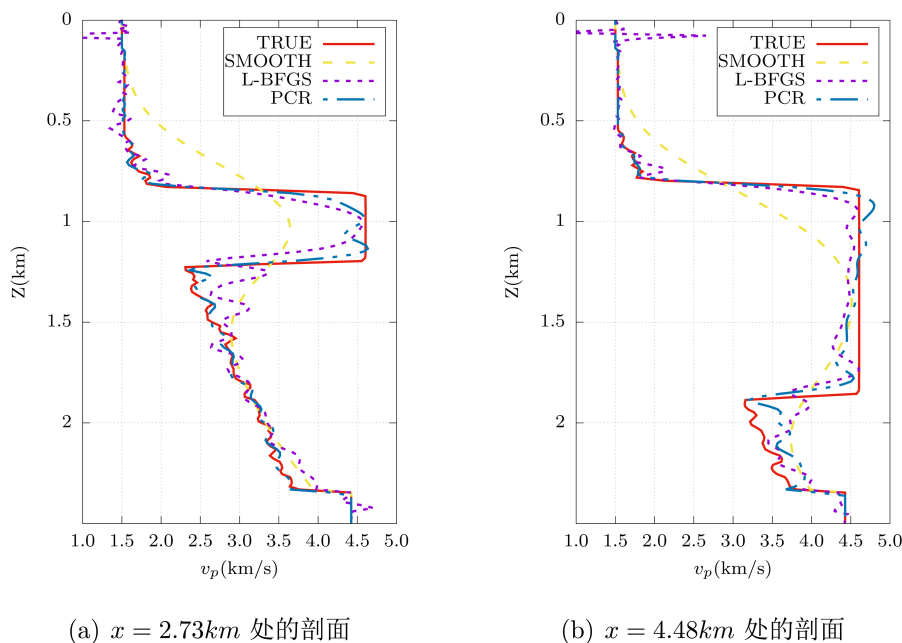


Figure 7. The inversion results of Sigsbee model

图 7. Sigsbee 模型的反演结果

## 4. 总结

从数值的角度来看, FWI是一个大规模不适定的强非线性问题, 如何建立一个准确高效的最小化方法成为一个难题。长期以来, 基于梯度的方法, 如最速下降算法或非线性共轭梯度算法,



**Figure 8.** Longitudinal profile of Sigsbee model parameters

**图 8.** Sigsbee 模型参数纵向剖面

以及拟牛顿方法L-BFGS, 已经被用来解决这个问题。但由于其二阶梯度信息的重要性, 本文使用了截断牛顿法。为了提高截断牛顿法的计算效率, 引入了预处理CR方法和信赖域全局化策略, 提出了PCRTRUST-NEWTON全波形反演方法。基于二维2004 BP模型和Sigsbee模型进行数值实验, 并与L-BFGS方法进行比较, 从数值角度验证了PCRTRUST-NEWTON反演方法的收敛性和高效性。实验结果表明PCRTRUST-NEWTON方法在收敛速度、计算效率以及重构分辨率方面均优于L-BFGS方法。本文提出的方法在复杂弹性介质中的表现还有待进一步的研究。

## 基金项目

国家自然科学基金(12261021)资助。

## 参考文献

- [1] Lailly, P. (1983) The Seismic Inverse Problem as a Sequence of before Stack Migrations. In: Bednar, J.B., Robinson, E. and Weglein, A., Eds., *Conference on Inverse Scattering—Theory and Application*, SIAM, Philadelphia, 206-220.
- [2] Tarantola, A. (1984) Inversion of Seismic Reflection Data in the Acoustic Approximation. *Geophysics*, **49**, 1259-1266. <https://doi.org/10.1190/1.1441754>
- [3] Tarantola, A. (1986) A Strategy for Nonlinear Elastic Inversion of Seismic Reflection Data. *Geophysics*, **51**, 1893-1903. <https://doi.org/10.1190/1.1442046>

- 
- [4] Pratt, R.G. (1990) Frequency-Domain Elastic Wave Modeling by Finite Differences: A Tool or Crosshole Seismic Imaging. *Geophysics*, **55**, 626-632. <https://doi.org/10.1190/1.1442874>
- [5] Mora, P. (1989) Inversion Migration Tomography. *Geophysics*, **54**, 1575-1586. <https://doi.org/10.1190/1.1442625>
- [6] Sirgue, L. and Pratt, R.G. (2004) Efficient Waveform Inversion and Imaging: A Strategy for Selecting Temporal Frequencies. *Geophysics*, **69**, 231-248. <https://doi.org/10.1190/1.1649391>
- [7] Choi, Y., Shin, C., Min, D.-J. and Ha, T. (2005) Efficient Calculation of the Steepest Descent Direction for Source-Independent Seismic Waveform Inversion: An Amplitude Approach. *Journal of Computational Physics*, **208**, 455-468. <https://doi.org/10.1016/j.jcp.2004.09.019>
- [8] Pratt, R.G., Shin, C. and Hick, G. (1998) Gauss-Newton and Full Newton Methods in Frequency-Space Seismic Waveform Inversion. *Geophysical Journal International*, **133**, 341-362. <https://doi.org/10.1046/j.1365-246X.1998.00498.x>
- [9] Brossier, R., Operto, S. and Virieux, J. (2009) Seismic Imaging of Complex Onshore Structures by 2D Elastic Frequency-Domain Full-Waveform Inversion. *Geophysics*, **74**, WCC105-WCC118. <https://doi.org/10.1190/1.3215771>
- [10] Métivier, L., Brossier, R., Virieux, J. and Operto, S. (2013) Full Waveform Inversion and the Truncated Newton Method. *SIAM Journal on Scientific Computing*, **35**, B401-B437. <https://doi.org/10.1137/120877854>
- [11] Wang, Y. and Yuan, Y. (2005) Convergence and Regularity of Trust Region Methods for Nonlinear Ill-Posed Inverse Problems. *Inverse Problems*, **21**, 821. <https://doi.org/10.1088/0266-5611/21/3/003>
- [12] Hestenes, M.R., Stiefel, E., *et al.* (1952) Methods of Conjugate Gradients for Solving Linear Systems. *Journal of research of the National Bureau of Standards*, **49**, 409-436. <https://doi.org/10.6028/jres.049.044>
- [13] Fong, D.C.-L. and Saunders, M. (2012) CG versus MINRES: An Empirical Comparison. *Sultan Qaboos University Journal for Science*, **17**, 44-62. <https://doi.org/10.24200/squjs.vol17iss1pp44-62>
- [14] Luenberger, D.G. (1970) The Conjugate Residual Method for Constrained Minimization Problems. *SIAM Journal on Numerical Analysis*, **7**, 390-398. <https://doi.org/10.1137/0707032>
- [15] Dahito, M.-A. and Orban, D. (2019) The Conjugate Residual Method in Linesearch and Trust-Region Methods. *SIAM Journal on Optimization*, **29**, 1988-2025. <https://doi.org/10.1137/18M1204255>
- [16] Shin, C., Jang, S. and Min, D.-J. (2001) Improved Amplitude Preservation for Prestack Depth Migration by Inverse Scattering Theory. *Geophysical Prospecting*, **49**, 592-606. <https://doi.org/10.1046/j.1365-2478.2001.00279.x>

- 
- [17] Hustedt, B., Operto, S. and Virieux, J. (2004) Mixed-Grid and Staggered-Grid Finite-Difference Methods for Frequency-Domain Acoustic Wave Modelling. *Geophysical Journal International*, **157**, 1269-1296. <https://doi.org/10.1111/j.1365-246X.2004.02289.x>
- [18] Berenger, J.-P. (1994) A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, 114, 185-200. <https://doi.org/10.1006/jcph.1994.1159>
- [19] Ghysels, P., Li, X.S., Rouet, F.-H., Williams, S. and Napov, A. (2016) An Efficient Multicore Implementation of a Novel HSS-Structured Multifrontal Solver Using Randomized Sampling. *SIAM Journal on Scientific Computing*, **38**, S358-S384. <https://doi.org/10.1137/15M1010117>
- [20] Billette, F. and Brandsberg-Dahl, S. (2005) The 2004 BP Velocity Benchmark. *67th EAGE Conference Exhibition, European Association of Geoscientists Engineers*, Madrid, 13-16 June 2005, cp-1-00513. <https://doi.org/10.3997/2214-4609-pdb.1.B035>
- [21] Paffenholz, J., McLain, B., Zaske, J. and Keliher, P.J. (2002) Subsalt Multiple Attenuation and Imaging: Observations from the Sigsbee2B Synthetic Dataset. *SEG International Exposition and Annual Meeting*, Salt Lake City, 6-11 October 2002, 2122-2125. <https://doi.org/10.1190/1.1817123>