A Hierarchical Prestack Seismic Inversion Scheme for VTI Media Based on the Exact Reflection Coefficient

Cong Luo, Jing Ba, and José M. Carcione

***Abstract*— Seismic exploration of unconventional hydrocarbon reservoirs (e.g., shale rocks) must take into account the transverse isotropy with a vertical axis of symmetry (VTI) characteristics. Prestack inversion for VTI media is more complex than the isotropic case, since the forward engine is highly nonlinear and more unknowns (five instead of three) are involved, which aggravate the ill-posedness of the inverse problem. Here, we propose a hierarchical inversion scheme to improve the estimation of the five parameters, where the exact reflection coefficient is used as forward engine (instead of the commonly-used approximations which lack accuracy). To handle the highly nonlinear inverse problem, we perform the prestack anisotropic inversion in two steps, namely, a preliminary linear result is used to reduce the search window and to formulate the constrain term and initial models of the subsequent nonlinear step. Specifically, for a reasonable preliminary estimation, we introduce a data-driven model building algorithm to provide reliable initial models, and employ the limited-memory Broyden–Fletcher–Goldfarb–Shanno combined with the momentum technique (LBFGS-MT) optimization to increase the convergence speed. We derive the Fréchet derivatives of the exact forward operator with respect to the parameters, i.e., the key factors of the linear part. Besides, we introduce a hybrid global optimization, the particle swarm optimization aided by very fast simulated annealing (PSO-VFSA) to enhance the accuracy and computational efficiency of the nonlinear stage. Synthetic tests demonstrate the effectiveness and accuracy of the proposed scheme. The field application shows that the method is capable to obtain reliable elastic information of shale reservoirs.**

***Index Terms*— Exact reflection coefficients, hierarchical inversion scheme, multi-parameter inversion, prestack seismic inversion, VTI medium.**

I. INTRODUCTION

PRESTACK seismic inversion using amplitude-versus-offset/angle (AVO/AVA) responses translates subsurface observations to elastic parameters [1]–[5], which enables the quantitative evaluation of oil/gas reservoirs and has been one of most effective techniques in geophysical exploration and petroleum industry. The traditional seismic inversions are based on isotropic media and achieve a great success in conventional sandstone explorations [6]–[8]. Nowadays, shale oil/gas, an unconventional resource with a great potential, is becoming one of the most important alternative resources. Most shales exhibit strong transverse isotropy with a vertical axis of symmetry (VTI) anisotropy due to the oriented clay particles and horizontal bedding structure [9], and the seismic response significantly differs from that of conventional sand-mudstone reservoirs. For conventional hydrocarbon reservoirs, seismic inversion based on the assumption of isotropic media is not applicable to unconventional reservoirs [10]–[13].

VTI caused by a combination of intrinsic mineral anisotropy and horizontal thin layers is one of the most common forms of effective anisotropy [14]. It is reported to have a significant effect on seismic amplitudes and AVO/AVA responses [15]–[17]. The exact expressions of the reflection coefficients for VTI media have been derived by several authors, e.g., Carcione [18], Daley and Hron [19], Graebner [20], Schoenberg and Protázio [21]. To characterize the anisotropy of VTI media, Thomsen defines three anisotropy parameters: ε, δ, and γ as a function of the elasticities, and gives simplified expressions for weak anisotropy [15]. Reflection coefficients can be parameterized using these parameters with different approximations [22]–[25]. Extension of the theory to stratified elastic/anelastic media is based on the reflectivity method [26]–[28]. However, this method is complex from a mathematical viewpoint for practical applications [20].

Many approximated reflection coefficients have been adopted in AVO/AVA inversion for VTI media [9], [29]–[31] due to their simple mathematical form. The most commonly used PP- and PS-reflection coefficients for VTI anisotropic prestack inversion are the Rüger approximation (RAI), Rüger [23], Lu *et al.* [31], Rüger [32] and its improved forms [9], [14]. Compared to isotropic three-parameter inversions, the VTI inverse problem has a higher ill-posedness and is more difficult in practice due to the higher number (five) of parameters and the smaller sensitivity of the anisotropy-related parameters to seismic amplitudes [9], [33].

In general, the existing prestack seismic AVO/AVA inversions for VTI media have several problems. The first is the simplified forward engines [9], [23], [31], [32] which are based on the assumption of weak anisotropy and weak impedance

Manuscript received August 27, 2021; revised November 27, 2021; accepted December 27, 2021. Date of publication January 4, 2022; date of current version March 8, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 42004111, in part by the China Postdoctoral Science Foundation under Grant 2019M661716 and Grant 2021T140175, and in part by the Jiangsu Innovation and Entrepreneurship Plan, and the Outstanding Youth Foundation of Jiangsu Province of China. *(Corresponding author: Jing Ba.)*

Cong Luo and Jing Ba are with the School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China (e-mail: jba@hhu.edu.cn).

José M. Carcione is with the National Institute of Oceanography and Applied Geophysics (OGS), 34010 Trieste, Italy, and also with the School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China.

Digital Object Identifier 10.1109/TGRS.2021.3140133

contrasts, and have a low accuracy at large offsets, where the anisotropy parameters exhibit more sensitivity to seismic responses [33]. Since the forward engine is the key factor of inversion, it significantly affects the accuracy of the inverted results. To overcome this problem, the isotropic [34]–[37] and anisotropy reflectivity methods [38] were introduced into AVO/AVA inversion algorithms using linear [35]–[37] and nonlinear schemes [34], [38]. Although the effectiveness and success of the reflectivity-method-based inversion have been verified, its application is limited because of the mathematical complexity and high-computational cost. Therefore, the study for VTI inversion still focuses on forward modeling based on the single interface assumption. More parameters make the inversion highly nonlinear and more difficult to achieve an optimal solution, which is the second problem of VTI inversion. Some studies try to reduce the number of parameters [9] or propose two-step strategies [14], [33] to improve the performance. However, these methods cannot estimate the five parameters simultaneously, and in most cases additional assumptions and approximations are needed to convert the results into the anisotropy parameters, which introduces additional errors. Besides, it is pointed out that there are large sensitivity differences between velocity- (P- and S-wave velocities) and anisotropy-related (Thomsen's ε and δ) parameters [33], leading to an unstable inversion algorithm. Using a linear optimization, anisotropic multi-parameter inversion will give local-minima solutions which need to be solved by nonlinear schemes [31], [39]–[41]. Lu *et al.* [31] propose a hierarchical inversion strategy by combining linear and nonlinear optimizations to improve the stability and accuracy in VTI media. Although using RAI as forward modeling has the restriction of weak contrasts and moderate anisotropy, their work provides a suitable approach for anisotropic multi-parameter estimation.

We propose a hierarchical inversion strategy that improves the precision to simultaneously estimate the five elastic parameters, namely, P-wave velocity α, S-wave velocity β, density ρ, and Thomsen's anisotropy parameters ε and δ. The commonly-used simplified forms limit the accuracy for large angles, strong impedance contrasts, and strong anisotropy, which inevitably accounts for the ill-posedness of the VTI inverse problem. To overcome such shortcomings, the exact reflection coefficient by Graebner [20] is used as forward modeling, which makes the proposed inversion applicable to arbitrary contrasts and large offsets. Besides, five inverted parameters cause a high ill-posed and serious local minima problem which is difficult to solve with the commonly-used linear optimization. The hierarchical inversion scheme based on the exact Graebner equation (EGHI) is proposed to solve the local-minima problem. It consists of two steps, the linear and nonlinear inversions. In the first step, we obtain a preliminary estimation of the parameters by a linear anisotropic inversion. The result is used to reduce the search windows, set up the constraint term, and define the start models for the subsequent nonlinear inversion. For an effective first step, a data-driven model building algorithm is used to obtain reliable initial models. The LBFGS combined with a momentum technique (LBFGS-MT) optimization is adopted to damp oscillations and increase the convergence. The Fréchet derivatives of the exact Graebner modeling (EG) with respect to the five parameters are derived. For nonlinear inversion, a hybrid global optimization method, i.e., particle swarm optimization, aided by very fast simulated annealing (PSO-VFSA) is introduced to enhance the accuracy and reduce the computational cost of the nonlinear stage.

Compared with existing approaches, the proposed scheme has the following advantages. 1) setting the exact reflection coefficient equation as the forward engine avoid the hypotheses and shortcomings of the approximations to improve the performance and applicability of the inversion and 2) in the hierarchical scheme, the nonlinear inversion aided by the linear results can effectively solve the local-minima problem and achieve good estimations with limited computational costs.

This article is organized as follows. In Section II, we first review the theory of the exact reflection coefficient of the VTI media. Then, the proposed hierarchical inversion scheme is given, including the exact-equation-based linear, and preliminary-result-aided nonlinear inversions. In Sections III and IV, we demonstrate the effectiveness and feasibility of the proposed hierarchical anisotropic inversion scheme with synthetic and field seismic data.

II. Theory and Methodology

A. Modeling Using the Exact Reflection Coefficient

One can represent a VTI medium using five elastic stiffness parameters, where c_{11}, c_{13}, c_{33}, and c_{55} determine the velocities of the qP and qSV waves, and c_{66} that of the SH wave. Considering a qP wave incident on an interface separating two transversely-isotropic homogeneous media, which lie on the xy plane, Graebner [20] gives the exact analytical expression of the reflection coefficients to describe the AVA characteristics of VTI media with four elasticities as a function of the incidence angle θ. The reflection/transmission coefficients are

$$\mathbf{r} = \mathbf{A}^{-1} \cdot \mathbf{b} \tag{1}$$

$$\mathbf{r} = \left(r_{\mathrm{PP}}\ r_{\mathrm{PS}}\ t_{\mathrm{PP}}\ t_{\mathrm{PS}} \right)^{\mathrm{T}} \tag{2}$$

whose components are the reflection coefficients PP (r_{PP}) and PS (r_{PS}), and transmission coefficients PP (t_{PP}) and PS (t_{PS}) (3), as shown at the bottom of the next page, and

$$\mathbf{b} = \left(-\ell_P^U\ n_P^U\ c_{55}^U\left(s_P^U \ell_P^U + p n_P^U\right)\ -p\ell_P^U c_{13}^U - s_P^U n_P^U c_{33}^U \right)^{\mathrm{T}} \tag{4}$$

where the subscripts U and L denote the upper and lower layers and the subscripts P and S refer to the P- and S-wave modes, respectively. The direction cosines are

$$\ell_P = \left[\left(c_{33}s_P^2 + c_{55}p^2 - \rho\right)/g_P\right]^{1/2}, \tag{5a}$$

$$n_P = \left[\left(c_{55}s_P^2 + c_{11}p^2 - \rho\right)/g_P\right]^{1/2} \tag{5b}$$

$$\ell_S = \left[\left(c_{55}s_S^2 + c_{11}p^2 - \rho\right)/g_S\right]^{1/2}, \tag{5c}$$

$$n_S = \left[\left(c_{33}s_S^2 + c_{55}p^2 - \rho\right)/g_S\right]^{1/2} \tag{5d}$$

$$g_{\#=P,S} = (c_{55} + c_{33})s_\#^2 + (c_{11} + c_{55})p^2 - 2\rho \tag{5e}$$

where s denotes the vertical slowness, which is a function of the horizontal slowness p and stiffness moduli

$$s = \left(1/\sqrt{2}\right) \cdot \left[K_1 \pm \left(K_1^2 - 4K_2K_3\right)^{1/2}\right]^{1/2}. \tag{6a}$$

The signs in s correspond to: $(-,-)$ the vertical slowness of the P-wave, and $(+,-)$ the vertical slowness of the S-wave. In terms of elasticities, we have

$$K_1 = \frac{\rho}{c_{33}} + \frac{\rho}{c_{55}} - \left(\frac{c_{11}}{c_{55}} + \frac{c_{55}}{c_{33}} - \frac{(c_{13}+c_{55})^2}{c_{33}c_{55}}\right)p^2 \tag{6b}$$

$$K_2 = (c_{11}/c_{33})p^2 - \rho/c_{33}, \quad K_3 = p^2 - \rho/c_{55}. \tag{6c}$$

According to Thomsen [15] and Carcione [28], the horizontal slowness is

$$p = \sqrt{2\rho^U}\sin\theta\left(c_{55}^U + c_{11}^U\sin^2\theta + c_{33}^U\cos^2\theta + \psi\right)^{-1/2} \tag{7a}$$

$$\begin{aligned}\psi = \Big\{&\left(c_{33}^U - c_{55}^U\right)^2 \\ &+\left[4\left(c_{13}^U + c_{55}^U\right)^2 - 2\left(c_{33}^U - c_{55}^U\right)\left(c_{11}^U + c_{33}^U - 2c_{55}^U\right)\right] \\ &\times \sin^2\theta + \left[\left(c_{11}^U + c_{33}^U - 2c_{55}^U\right)^2 - 4\left(c_{13}^U + c_{55}^U\right)^2\right] \\ &\sin^4\theta\Big\}^{1/2}\end{aligned} \tag{7b}$$

where θ denotes the incidence angle. Equations (1)–(7) give the exact expression of the reflection coefficient in terms of (c_{11}, c_{13}, c_{33}, c_{55}) and ρ. To intuitively characterize the degree of anisotropy, Thomsen [15] replaced the stiffness parameters with a new set of coefficients involving two velocities (α and β) and two non-dimensional "anisotropies" ε and δ

$$c_{33} = \alpha^2\rho, \quad c_{55} = \beta^2\rho, \ c_{11} = (2\varepsilon+1)\alpha^2\rho, c_{13} = \rho\eta - \beta^2\rho \tag{8a}$$

$$\eta = \left[(\varepsilon+\delta+1)\alpha^4 - (\varepsilon+2)\alpha^2\beta^2 - \beta^4\right]^{1/2}. \tag{8b}$$

The above expressions are exact and valid for arbitrary (not only weak) anisotropy [15]. Then, (8) allows us to express the reflection coefficient in terms of α, β, ρ, ε, and δ.

B. Hierarchical Anisotropic Inversion Scheme

A hierarchical inversion strategy is adopted based on an exact expression of the reflection coefficients. In the first step, we perform a linear anisotropic AVA inversion to derive five preliminary parameters. The results of the linear inversion deviate from the true models to some extent. We use these preliminary results to setup initial particles and search windows to accelerate and constrain the subsequent nonlinear inversion.

1) Objective Function: The seismic angle gather $\mathbf{d}^*$ is a matrix of dimension, Nt $\times$ Na where Nt and Na denote the number of sampling points and incidence angles, respectively. It can be expressed as

$$\mathbf{d}^* = \begin{bmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,\mathrm{Na}} \\ \vdots & \vdots & \ddots & \vdots \\ d_{\mathrm{Nt},1} & d_{\mathrm{Nt},2} & \cdots & d_{\mathrm{Nt},\mathrm{Na}} \end{bmatrix}_{\mathrm{Nt}\times\mathrm{Na}} \tag{9a}$$

where $d_{i,j}$ with $i = 1,\ldots,\mathrm{Nt}$ and $j = 1,\ldots,\mathrm{Na}$. In the inversion process, the observed vector $\mathbf{d}^*$ is discretized with a Nt $\times$ Na $\times$ 1 dimension, as

$$\mathbf{d} = \left[d_{1,1},\ldots,d_{\mathrm{Nt},1}, d_{1,2},\ldots,d_{\mathrm{Nt},2},\ldots,d_{1,\mathrm{Na}},\ldots,d_{\mathrm{Nt},\mathrm{Na}}\right]^{\mathrm{T}}. \tag{9b}$$

The forward model can be expressed as

$$\mathbf{d} = G_{\mathrm{PP}}(\mathbf{m}) + \mathbf{e} = w * r_{\mathrm{PP}}(\mathbf{m}) + \mathbf{e} = \mathbf{W}\cdot r_{\mathrm{PP}}(\mathbf{m}) + \mathbf{e} \tag{10}$$

where G_{PP}, the forward engine nonlinearly related to the model $\mathbf{m}$, is the convolution of the source wavelet w with the reflection coefficient $r_{\mathrm{PP}}(\mathbf{m})$, and $\mathbf{e}$ is a vector of random noise. The model vector contains 5Nt unknowns

$$\begin{aligned}\mathbf{m} = [&\alpha_1,\ldots,\alpha_{\mathrm{Nt}}, \beta_1,\ldots,\beta_{\mathrm{Nt}}, \rho_1,\ldots,\rho_{\mathrm{Nt}}, \\ &\varepsilon_1,\ldots,\varepsilon_{\mathrm{Nt}}, \delta_1,\ldots,\delta_{\mathrm{Nt}}].\end{aligned} \tag{11}$$

Due to the presence of noise, the problem is ill-posed. Thus, regularization techniques are commonly used. The objective function is usually written as the weighted sum of the data misfit term $\Gamma(\mathbf{m})$ and the model constraint term $\Phi(\mathbf{m})$

$$S(\mathbf{m}) = \Gamma(\mathbf{m}) + \lambda\Phi(\mathbf{m}). \tag{12a}$$

The most widely-used data misfit function is

$$\Gamma(\mathbf{m}) = [\mathbf{d} - G_{\mathrm{PP}}(\mathbf{m})]^{\mathrm{T}}[\mathbf{d} - G_{\mathrm{PP}}(\mathbf{m})]. \tag{12b}$$

The regularization constraint $\Phi(\mathbf{m})$ should be set according to the distribution of the subsurface properties. In regions where no sharp changes exist, $\Phi(\mathbf{m})$ could be defined as a smooth L_2-based regularization. Then, the objective function is

$$S(\mathbf{m}) = [\mathbf{d} - G_{\mathrm{PP}}(\mathbf{m})]^{\mathrm{T}}[\mathbf{d} - G_{\mathrm{PP}}(\mathbf{m})] + \lambda(\mathbf{m}-\mathbf{u})^{\mathrm{T}}\mathbf{C}_m^{-1}(\mathbf{m}-\mathbf{u}) \tag{13}$$

where $\mathbf{u}$ is the expectation of $\mathbf{m}$. $\mathbf{C}_m$, the covariance matrix of the parameter vector, is introduced to improve the stability of the inversion [42].

$$\mathbf{A} = \begin{pmatrix} \ell_P^U & n_S^U & -\ell_P^L & -n_S^L \\ n_P^U & -\ell_S^U & n_P^L & -\ell_S^L \\ c_{55}^U\left(s_P^U\ell_P^U + pn_P^U\right) & c_{55}^U\left(s_S^Un_S^U - p\ell_S^U\right) & c_{55}^L\left(s_P^L\ell_P^L + pn_P^L\right) & c_{55}^L\left(s_S^Ln_S^L - p\ell_S^L\right) \\ p\ell_P^Uc_{13}^U + s_P^Un_P^Uc_{33}^U & pn_S^Uc_{13}^U - s_S^U\ell_S^Uc_{33}^U & -p\ell_P^Lc_{13}^L - s_P^Ln_P^Lc_{33}^L & -pn_S^Lc_{13}^L + s_S^L\ell_S^Lc_{33}^L \end{pmatrix} \tag{3}$$

2) Linear Anisotropic AVA Inversion: For linear inversion, the gradient-based optimization method is usually used to solve the objective function. To overcome the issue of numerous local minima, damp oscillations, and increase the convergence speed, the momentum technique (MT) [43] has been introduced into the quasi-Newton optimization approaches [44]. By combining the LBFGS and MT algorithms, given the current estimation $\mathbf{m}_k$ at the kth iteration (from $k = 1$), the iterative method of the LBFGS-MT algorithm can be expressed as

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \mathbf{v}_{k+1} \tag{14a}$$

$$\mathbf{v}_{k+1} = \varphi_k \mathbf{v}_k - a_k \mathbf{g}(\mathbf{m}_k). \tag{14b}$$

Compared with the traditional update form, the kernel idea of MT algorithm is to add a fraction of the previous update vector $\mathbf{v}_k$ to the current one $\mathbf{v}_{k+1}$. In particular, $\mathbf{v}_1$ is set to zero. φ_k, the momentum coefficient, satisfies $0 < \phi_k < 1$ and $\phi_k \underset{k\to 0}{\longrightarrow} 0$, and is set as

$$\varphi_k = F^k \varphi_0, \quad 0 < F < 1 \tag{15}$$

where φ_0 is the initial coefficient. In (14), a_k denotes the step size of the kth iteration computed with the strong Wolfe line search algorithm [45], and $-\mathbf{g}(\mathbf{m}_k)$ represents the descent direction of the kth iteration, which is determined by

$$-\mathbf{g}(\mathbf{m}_k) = -H(\mathbf{m}_k) \cdot J(\mathbf{m}_k). \tag{16}$$

The Jacobian matrix $J(\mathbf{m}_k)$ for the kth iteration denotes the first-order derivative of the objective function

$$J(\mathbf{m}_k) = \left[\frac{\partial G_{\mathrm{PP}}(\mathbf{m}_k)}{\partial \mathbf{m}}\right]^{\mathrm{T}} [\mathbf{d} - G_{\mathrm{PP}}(\mathbf{m}_k)] + \lambda \mathbf{C}_m^{-1}(\mathbf{m}_k - \mathbf{u}). \tag{17}$$

In the LBFGS algorithm, $H(\mathbf{m}_k)$, the pseudo Hessian matrix of the objective function, is given by an iterative algorithm

$$\begin{cases} H_i(\mathbf{m}_k)^{-1} = \mathbf{Q}_{i-1}^{\mathrm{T}} H_{i-1}(\mathbf{m}_k)^{-1} \mathbf{Q}_{i-1} + \Gamma_{i-1}\mathbf{s}_{i-1}\mathbf{s}_{i-1}^{\mathrm{T}} \\ \Gamma_{i-1} = 1/\left(\mathbf{y}_{i-1}^{\mathrm{T}}\mathbf{s}_{i-1}\right), \mathbf{Q}_{i-1} = \mathbf{I} - \Gamma_{i-1}\mathbf{y}_{i-1}\mathbf{s}_{i-1}^{\mathrm{T}} \qquad i = 2, \ldots, k \\ \mathbf{s}_{i-1} = \mathbf{m}_i - \mathbf{m}_{i-1}, \mathbf{y}_{i-1} = J(\mathbf{m}_i) - J(\mathbf{m}_{i-1}) \end{cases} \tag{18}$$

we set

$$H_1(\mathbf{m}_k) = \left(\left(\frac{\partial G_{\mathrm{PP}}(\mathbf{m}_k)}{\partial \mathbf{m}}\right)^{\mathrm{T}} \frac{\partial G_{\mathrm{PP}}(\mathbf{m}_k)}{\partial \mathbf{m}} + \lambda \mathbf{C}_m^{-1}\right). \tag{19}$$

In (17) and (19), $\partial G_{\mathrm{PP}}(\mathbf{m}_k)/\partial \mathbf{m}$ is the Fréchet derivative of the nonlinear forward operator G_{PP} with respect to $\mathbf{m}$ in the kth iteration. Based on the convolution model, the Fréchet derivatives $\partial G_{\mathrm{PP}}(\mathbf{m})/\partial \mathbf{m}$, a matrix of dimension $(\mathrm{Nt}\cdot\mathrm{Na}) \times (5\mathrm{Nt})$, can be computed in terms of the derivatives of the reflection coefficient and wavelet matrix

$$\frac{\partial G_{\mathrm{PP}}(\mathbf{m})}{\partial \mathbf{m}} = \begin{bmatrix} \mathbf{W}(\theta_1)\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_1)}{\partial \boldsymbol{\alpha}} & \mathbf{W}(\theta_1)\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_1)}{\partial \boldsymbol{\beta}} & \mathbf{W}(\theta_1)\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_1)}{\partial \boldsymbol{\rho}} & \mathbf{W}(\theta_1)\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_1)}{\partial \boldsymbol{\varepsilon}} & \mathbf{W}(\theta_1)\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_1)}{\partial \boldsymbol{\delta}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{W}(\theta_{\mathrm{Na}})\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_{\mathrm{Na}})}{\partial \boldsymbol{\alpha}} & \mathbf{W}(\theta_{\mathrm{Na}})\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_{\mathrm{Na}})}{\partial \boldsymbol{\beta}} & \mathbf{W}(\theta_{\mathrm{Na}})\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_{\mathrm{Na}})}{\partial \boldsymbol{\rho}} & \mathbf{W}(\theta_{\mathrm{Na}})\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_{\mathrm{Na}})}{\partial \boldsymbol{\varepsilon}} & \mathbf{W}(\theta_{\mathrm{Na}})\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_{\mathrm{Na}})}{\partial \boldsymbol{\delta}} \end{bmatrix}_{(\mathrm{Nt}\cdot\mathrm{Na})\times(5\mathrm{Nt})} \tag{20}$$

where $\partial \mathbf{r}_{\mathrm{PP}}(\theta_i)/\partial \mathbf{m}_*$, with $* = \alpha, \beta, \rho, \varepsilon, \delta$; $\mathbf{W}(\theta_i)$ is the wavelet matrix with dimension $\mathrm{Nt} \times \mathrm{Nt}$. If α is considered, we have

$$\frac{\partial \mathbf{r}_{\mathrm{PP}}(\theta_i)}{\partial \boldsymbol{\alpha}} = \begin{bmatrix} \frac{\partial r_{\mathrm{PP}}(\theta_i, t_1)}{\partial \boldsymbol{\alpha}_1} & \frac{\partial r_{\mathrm{PP}}(\theta_i, t_1)}{\partial \boldsymbol{\alpha}_2} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{\partial r_{\mathrm{PP}}(\theta_i, t_2)}{\partial \boldsymbol{\alpha}_2} & \frac{\partial r_{PP}(\theta_i, t_2)}{\partial \boldsymbol{\alpha}_3} & 0 & \vdots & \vdots \\ \vdots & 0 & & \ddots & & 0 \\ \vdots & \vdots & \vdots & 0 & \frac{\partial r_{\mathrm{PP}}(\theta_i, t_{\mathrm{Nt}-1})}{\partial \boldsymbol{\alpha}_{\mathrm{Nt}-1}} & \frac{\partial r_{\mathrm{PP}}(\theta_i, t_{\mathrm{Nt}-1})}{\partial \boldsymbol{\alpha}_{Nt}} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{\partial r_{\mathrm{PP}}(\theta_i, t_{\mathrm{Nt}})}{\partial \boldsymbol{\alpha}_{\mathrm{Nt}}} \end{bmatrix}_{\mathrm{Nt}\times\mathrm{Nt}} \tag{21}$$

where $\partial r_{\mathrm{PP}}(\theta_i, t_j)/\partial \alpha_k$ denotes the derivative of r_{PP}(jth interface and angle θ_i) with respect to α of the jth layer.

Then the general forms of the derivatives of the reflection/transmission coefficients are given. For a given sampling point and a given incidence angle, differentiating both sides of (1), we obtain

$$\frac{\partial \mathbf{r}}{\partial \mathbf{M}} = \mathbf{A}^{-1}\frac{\partial \mathbf{b}}{\partial \mathbf{M}} - \mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial \mathbf{M}}\mathbf{r}. \tag{22}$$

$\mathbf{M}$ denotes a vector which contains the model parameters across a certain reflecting interface (m^U and m^L). The superscripts U and L represent the upper and lower layers, respectively. Since $\partial \mathbf{b}/\partial m_*^L = 0$, (22) can be expressed as

$$\frac{\partial \mathbf{r}}{\partial m_*^U} = \mathbf{A}^{-1}\frac{\partial \mathbf{b}}{\partial m_*^U} - \mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial m_*^U}\mathbf{r} \tag{23}$$

$$\frac{\partial \mathbf{r}}{\partial m_*^L} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial m_*^L}\mathbf{r}. \tag{24}$$

Solving (23) and (24), we can obtain $\partial \mathbf{r}/\partial m_*^\# = [\partial r_{\mathrm{PP}}/\partial m_*^\# \;\; \partial r_{\mathrm{PS}}/\partial m_*^\# \;\; \partial t_{\mathrm{PP}}/\partial m_*^\# \;\; \partial t_{\mathrm{PS}}/\partial m_*^\#]^{\mathrm{T}}$ with $* = \alpha, \beta, \rho, \varepsilon, \delta$ and $\# = U, L$, and one can extract the derivatives $\partial r_{\mathrm{PP}}/\partial m_*^\#$ or $\partial r_{\mathrm{PS}}/\partial m_*^\#$ as required in practice. In (22) and (23), $\partial \mathbf{A}/\partial \mathbf{M}$ and $\partial \mathbf{b}/\partial \mathbf{M}$ are the partial derivatives of the intermediate matrices $\mathbf{A}$ and $\mathbf{b}$ with respect to $\mathbf{M}$, including $\partial \mathbf{A}/\partial m_*^U$, $\partial \mathbf{A}/\partial m_*^L$, $\partial \mathbf{b}/\partial m_*^U$, and $\partial \mathbf{b}/\partial m_*^L$, whose expressions are given in Appendix A. By calculating all $\partial \mathbf{r}/\partial \mathbf{M}$ at different sampling points and incidence angles by (22)–(24) and extracting all elements of (21), the Fréchet-derivative matrix $\partial G_{\mathrm{PP}}(\mathbf{m})/\partial \mathbf{m}$ is obtained.

3) Nonlinear Anisotropic AVA Inversion: The five-parameter anisotropic inversion requires a high computational efficiency and accuracy for the nonlinear optimization approach. Therefore, the PSO-VFSA algorithm is used [46]–[49], which incorporates the probabilistic jumping mechanism of the FSA into the PSO, thus reducing the computational cost of the FSA, as well as avoiding the premature convergence of the PSO [47], [49].

To initialize the standard PSO [50], n particles in Nt $\times$ 1 dimension are defined, each denoting a potential solution to the inverse problem. The ith particle is $\mathbf{m}_i = [m_{i,1}, \ldots, m_{i,j}, \ldots m_{i,5Nt}]^{\mathrm{T}}, i = 1, \ldots, n$ and $j = 1, \ldots, 5Nt$. The iteration format is

$$\mathbf{m}_{i,k} = \mathbf{m}_{i,k-1} + \kappa_1 \mathbf{V}_{i,k-1} \tag{25}$$

where $\mathbf{m}_{i,k}$ represents the ith particle in the kth iteration, and κ_1 is the inertia coefficient. $\mathbf{V}_{i,k-1}$, the updated velocity associated with the ith particle in the $(k-1)$th iteration, is given by

$$\mathbf{V}_{i,k} = \mathbf{V}_{i,k-1} + \mathrm{rand}_1(\cdot)\kappa_2\left(\mathbf{m}_{i,k-1}^{\mathrm{pbest}} - \mathbf{m}_{i,k-1}\right) + \mathrm{rand}_2(\cdot)\kappa_3\left(\mathbf{m}_{k-1}^{\mathrm{gbest}} - \mathbf{m}_{i,k-1}\right) \tag{26}$$

where $\mathbf{m}_{i,k-1}^{\mathrm{pbest}}$ and $\mathbf{m}_{k-1}^{\mathrm{gbest}}$ are the personal best position for the ith particle and the global best position for all the particles, respectively, all of which are selected by comparing each particles' fitness value computed according to the objective function (13). $\mathrm{rand}_1(\cdot)$ and $\mathrm{rand}_2(\cdot)$ are independent random numbers within [0, 1], and κ_2 and κ_3 are two constants to control the step size of update.

Conventionally, $\mathbf{m}_k^{\mathrm{gbest}}$ is one vector selected from all $\mathbf{m}_{i,k}^{\mathrm{pbest}}$ and plays a leading role in the final solution. However, if $\mathbf{m}_k^{\mathrm{gbest}}$ is close to a local minimum, the solution will deviate from the global solution [47]. To this end, in the PSO-VFSA algorithm, $\mathbf{m}_k^{\mathrm{gbest}}$ is no longer selected simply from the best one of $\mathbf{m}_{i,k}^{\mathrm{pbest}}$, but selected from all the particles $\mathbf{m}_{i,k}$ according to the accepted probability $P_{i,k}$ inspired by the VFSA algorithm [39], [51]

$$P_{i,k} = Z_r\left(\mathbf{m}_{i,k}, \mathbf{m}_k^{\mathrm{gbest}}, T_k\right) \Big/ \sum Z_r\left(\mathbf{m}_{i,k}, \mathbf{m}_k^{\mathrm{gbest}}, T_k\right) \tag{27}$$

with

$$Z_r\left(\mathbf{m}_{i,k}, \mathbf{m}_k^{\mathrm{gbest}}, T_k\right) = \left[1 - (1-\tau) \cdot \left(\Delta S_{m_{i,k}} / T_k\right)\right]^{1/(1-k)} \tag{28a}$$

$$\Delta S_{m_{i,k}} = S\left(\mathbf{m}_{i,k}\right) - S\left(\mathbf{m}_k^{\mathrm{gbest}}\right) \tag{28b}$$

where $P_{i,k}$ denotes the accepted probability of the ith particle in the kth iteration, T_k represents the current temperature in the kth iteration, τ is a constant, and $S(\mathbf{m}_{i,k})$ and $S(\mathbf{m}_k^{\mathrm{gbest}})$ denote the objective fitness values by substituting $\mathbf{m}_{i,k}$ and $\mathbf{m}_k^{\mathrm{gbest}}$ into the objective function (13). Z_r is the replacing fitness value, used to describe the possibility of replacing $\mathbf{m}_k^{\mathrm{gbest}}$ by $\mathbf{m}_{i,k}$. When the $P_{i,k}$ of all the particles is obtained, the roulette wheel selection [47] is employed to probabilistically determine which particle is the new $\mathbf{m}_k^{\mathrm{gbest}}$.

To accelerate the nonlinear inversion, another key factor is to setup a suitable search window denoted by $[\mathbf{m}^{\min}\mathbf{m}^{\max}]$ with $\mathbf{m}^{\min} = [m_1^{\min}, \ldots, m_j^{\min}, \ldots, m_{5\mathrm{Nt}}^{\min}]$ and $\mathbf{m}^{\max} = [m_1^{\max}, \ldots, m_j^{\max}, \ldots, m_{5\mathrm{Nt}}^{\max}]$, of which size is expected to be small while covering feasible solution space. Based on the linear inversion results, the search window is set as

$$\begin{cases} m_j^{\min} = m_j^{\mathrm{Li}} - 2b \cdot m_j^{\mathrm{Li}} \\ m_j^{\max} = m_j^{\mathrm{Li}} + 2b \cdot m_j^{\mathrm{Li}}, \end{cases} \quad \text{if } \left|m_j^{\mathrm{Re}} - m_j^{\mathrm{Li}}\right| < b \cdot m_j^{\mathrm{Li}} \tag{29a}$$

$$\begin{cases} m_j^{\min} = m_j^{\mathrm{Li}} - (1+2b) \cdot \left|m_j^{\mathrm{Re}} - m_j^{\mathrm{Li}}\right| \\ m_j^{\max} = m_j^{\mathrm{Li}} + (1+2b) \cdot \left|m_j^{\mathrm{Re}} - m_j^{\mathrm{Li}}\right|, \end{cases} \quad \text{if } \left|m_j^{\mathrm{Re}} - m_j^{\mathrm{Li}}\right| \geq b \cdot m_j^{\mathrm{Li}} \tag{29b}$$

where $j = 1, \ldots, 5Nt \cdot m_j^{\mathrm{Li}}$ and m_j^{Re} denote the preliminary result of the linear inversion and a reference true value extracted from well data. The size of the search window is determined by the accuracy of m^{Li}, i.e., when too many anomalies occur, (29b) is useful to generate a wider search window. b is a scale coefficient which varies with the relative change of the parameters.

Besides reducing the search windows, the linear stage is used to define the constraint in (13). We compute the covariance matrix of the constrain term with the preliminary results to provide a reliable correlation between parameters for the nonlinear inversion. Although, in general, initial models are not necessary (randomly-generated models are usually used), a good initial model helps in improving the estimations and convergence speed. Therefore, we set the initial models (initial particles) for the nonlinear inversion according to the linear outcome, i.e., the initial particles are randomly generated within the reduced search window and the linear result is also kept as one of the initial particles.

C. Some Practical Aspects of the Technique

1) Rock-Physics-Based ε and δ Estimation: Since the EGHI algorithm is a model-based inversion algorithm, the target parameters at the well position, including α, β, ρ, and the two anisotropy parameters ε and δ, should at least be provided. However, it is difficult to directly measure the anisotropy parameters in a borehole, especially for a VTI medium. An effective rock-physics-model-based method [9], [52] is used to predict them, which contains two main parts, the rock-physics model for shale rocks and the anisotropy parameter prediction method. Based on the shale model, the prediction procedure for ε and δ is achieved by solving an inverse problem, where two new parameters, the pore aspect ratio (asp) and the lamination index (LI), are considered since they affect the anisotropy. The objective function of the inversion process is as follows:

$$\arg\min \left\|\mathbf{c}_{33}^{\mathrm{real}} - \mathbf{c}_{33}^{\mathrm{pre}}(\mathrm{asp}, \mathrm{LI})\right\| + \left\|\mathbf{c}_{55}^{\mathrm{real}} - \mathbf{c}_{55}^{\mathrm{pre}}(\mathrm{asp}, \mathrm{LI})\right\| \tag{30}$$

where $\mathbf{c}_{33}^{\mathrm{real}}$ and $\mathbf{c}_{55}^{\mathrm{real}}$ denote the elastic constant vectors calculated by the logging data α, β, and ρ; $\mathbf{c}_{33}^{\mathrm{pre}}(\mathrm{asp}, \mathrm{LI})$ and

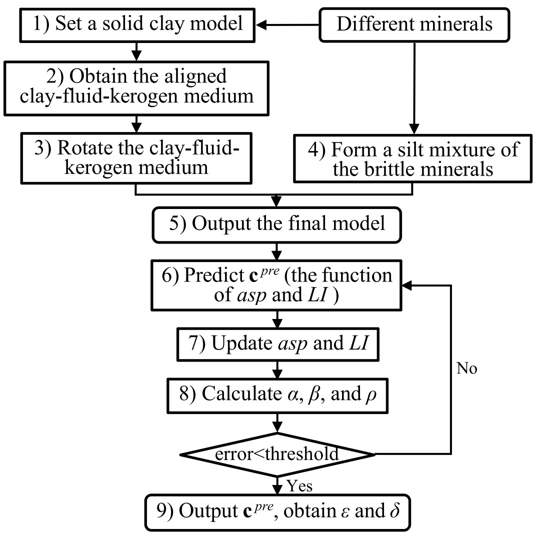

Fig. 1. Flowchart of the rock-physics-model-based prediction method. 1) Set a solid clay model consisting of different minerals by Backus average [53]. 2) Add the saturated clay-related pores and kerogen to the solid clay in 1) based on SCA and DEM [53] to obtain the aligned clay-fluid-kerogen medium. 3) Aided by a combination of the Bond transform and the VRH average [53], we have a rotated clay-fluid-kerogen medium. 4) Form a silt mixture of the brittle minerals by the Hashin–Shtrikman–Walpole averaging [52]. 5) Add the silt mixture to 3) by the anisotropic SCA to get the final model. 6) Using the model in 5), predict the elastic constants $\mathbf{c}^{\mathrm{pre}} = [\mathbf{c}_{11}^{\mathrm{pre}}\ \mathbf{c}_{13}^{\mathrm{pre}}\ \mathbf{c}_{33}^{\mathrm{pre}}\ \mathbf{c}_{55}^{\mathrm{pre}}]$ which are the functions of asp and LI. 7) Minimize the objective function (30) to update asp and LI. 8) Based on the model in 5), calculate α, β, and ρ using the new asp and LI; if the errors are high, repeat steps 6) and 7). 9) When the errors are less than the preset thresholds, output $\mathbf{c}^{\mathrm{pre}}$; calculate ε and δ using according to (8).

$\mathbf{c}_{55}^{\mathrm{pre}}(\mathrm{asp}, \mathrm{LI})$, the elastic constants predicted by the rock-physics-model, are the functions of asp and LI. The procedure of the rock-physics-based prediction method is shown in Fig. 1.

2) Data-Driven Model Building: A good initial model can effectively improve the accuracy of the linear inversion. For target-oriented modeling, it is routine to build the initial model by interpolating well logs along artificially interpreted geological horizons, which requires intensive human interpretation and only a small number of geological elements are considered as constraints. A data-driven model building algorithm by combining the seismic slope of geological structures and well-log interpolation provides a remarkable improvement [54]. Based on plane-wave decomposition (PWD) [55], the seismic slope attribute is estimated from migration profiles to replace the artificial interpreted horizons as lateral constraints [3]. Then, high-confidence initial models can be obtained by solving a reshaping-regularized inverse problem, instead of being computed by the traditional interpolation method.

For a work area, sparsely distributed wells can be used to sample a subsurface model via the sampling operator $\mathbf{\Theta}$ [54] as

$$\mathbf{\Theta X} = \mathbf{Y} \tag{31}$$

where $\mathbf{X}$ indicates property models and $\mathbf{Y}$ is related to well data. Assuming that there are two wells (at the first and ith grid points) in a 2-D working area and taking α as the

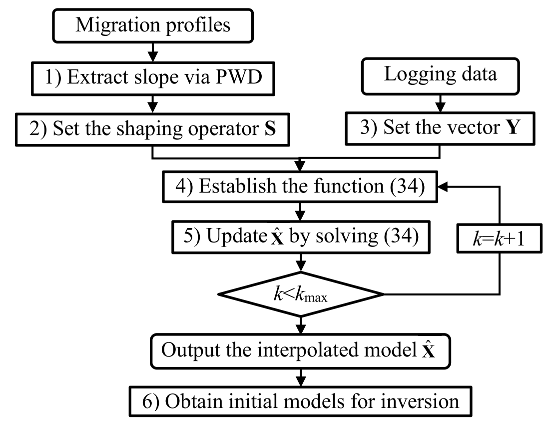

Fig. 2. Flowchart of the data-driven model building algorithm. 1) Extract the seismic slope attribute from input migration profiles by PWD. 2) Set the shaping operator $\mathbf{S}$ based on the slope attribute in 1). 3) Set the vector $\mathbf{Y}$ using the logging data. 4) Establish the shaping-regularized function (34) based on 2) and 3). 5) Solve (34) to update $\hat{\mathbf{X}}$ via the conjugate gradient algorithm. 6) After Backus averaging [53], the initial models can be obtained.

example, (31) takes the form

$$\begin{bmatrix} \mathbf{I} & & & & & \\ & \mathbf{0} & & & & \\ & & \ddots & & & \\ & & & \mathbf{I} & & \\ & & & & \mathbf{0} & \\ & & & & & \ddots \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_i \\ \boldsymbol{\alpha}_{i+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \mathbf{0} \\ \vdots \\ \boldsymbol{\alpha}_i \\ \mathbf{0} \\ \vdots \end{bmatrix} \tag{32}$$

where $\mathbf{X} = [\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \cdots\ \boldsymbol{\alpha}_i\ \boldsymbol{\alpha}_{i+1}\ \cdots]^{\mathrm{T}}$ stands for a vectorized model of $\boldsymbol{\alpha}$, which is reshaped from a 2-D matrix to a 1-D vector; $\mathbf{Y} = [\boldsymbol{\alpha}_1\ \mathbf{0}\ \cdots\ \boldsymbol{\alpha}_i\ \mathbf{0}\ \cdots]^{\mathrm{T}}$ is constructed from the logging data. In the sampling matrix Θ, the diagonal blocks corresponding to the well positions are set to the identity matrix, the others are zeros. To solve the inverse problem (31) to get the reconstructed model $\mathbf{X}$, the common way is to formulate it as a Tikhonov-regularized optimization problem

$$\hat{\mathbf{X}} = \arg\min \|\mathbf{\Theta X} - \mathbf{Y}\|_2^2 + \lambda \|\mathbf{X}\|_2^2. \tag{33}$$

The shaping regularization is introduced into the inverse problem to add lateral constraints provided by seismic data. In this framework, a shaping operator $\mathbf{S}$ constrains the model based on prior information [56], and we have (the derivation process is shown in Appendix B)

$$\hat{\mathbf{X}} = \mathbf{PH}\left[\upsilon^2 \mathbf{I} + \mathbf{H}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\left(\Theta^{\mathrm{T}}\Theta - \upsilon^2 \mathbf{I}\right)\mathbf{PH}\right]^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\Theta^{\mathrm{T}}\mathbf{Y} \tag{34}$$

where $\hat{\mathbf{X}}$ represents the shaping-regularized least-square estimation. Then, one can reconstruct the model by solving (34) via the conjugate gradient algorithm [57]. In building an initial model, a robust local slope extracted from seismic data by PWD is used to obtain the structural shaping operator $\mathbf{S}$. For practical applications, the final output $\hat{\mathbf{X}}$ are the interpolated 2-D/3-D models, including $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\rho}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\delta}$. Fig. 2 shows the workflow of this model building algorithm.

By applying the proposed EGHI scheme to estimate the elastic parameters, the procedure can be described as Fig. 3.

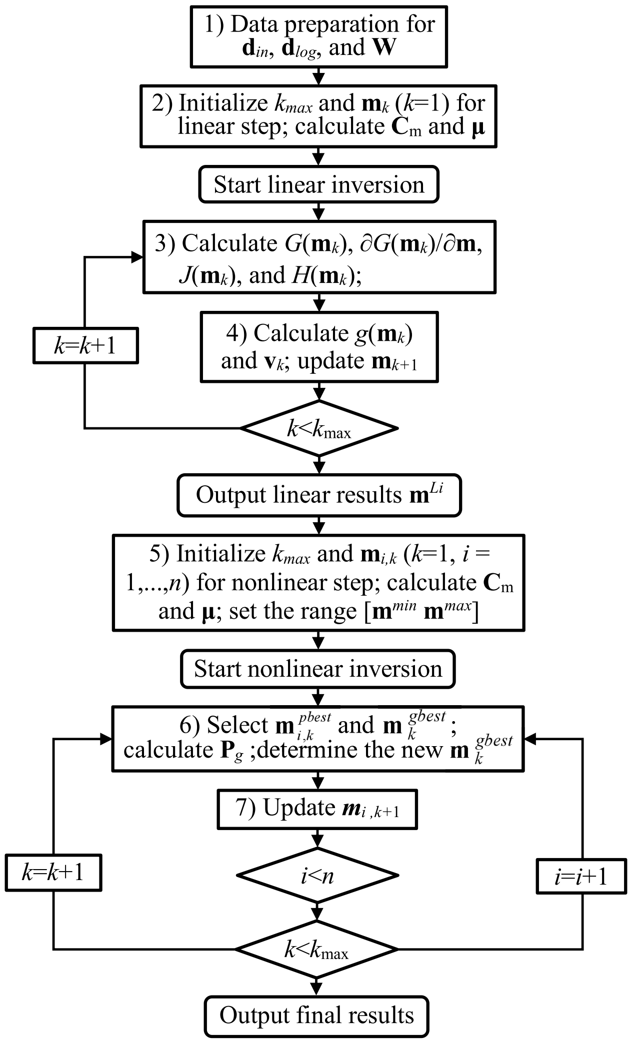

Fig. 3. Flowchart of the EGHI. The implementation of the scheme is as follows. 1) Reshape $\mathbf{d}_{obs}$ to $\mathbf{d}_{in}$ with an (Nt·Na) $\times$ 1 dimension according to (9); predict the logs of ε and δ according to the rock-physics-based method; estimate the angle-dependent statistical wavelets from seismic data to set wavelet matrix $\mathbf{W}$. 2) For the linear inversion, initialize k_{max}, and set the initial models $\mathbf{m}_k$ ($k = 0$) using the data-driven model building algorithm; calculate $\mathbf{C}_m$ and μ. 3) Based on $\mathbf{m}_k$, calculate $G(\mathbf{m}_k)$ using (10) and (1), $\partial G(\mathbf{m}_k)/\partial \mathbf{m}$ according to (20)–(24); then calculate $J(\mathbf{m}_k)$ and $H(\mathbf{m}_k)$ with (17)–(19), respectively. 4) Calculate $g(\mathbf{m}_k)$ using (16) and update $\mathbf{m}_{k+1}$ according to (14). 5) For the nonlinear inversion, initialize k_{max}; based on the linear results $\mathbf{m}^{Li}$, set the start particles $\mathbf{m}_{i,k}$ ($k = 1, i = 1, \ldots, n$), calculate $\mathbf{C}_m$ and $\boldsymbol{\mu}$, and set the search range [$\mathbf{m}^{min}$ $\mathbf{m}^{max}$] according to (29). 6) Apply the objective function (13) to all the particles, select $\mathbf{m}_{i,k}^{pbest}$ and $\mathbf{m}_k^{gbest}$; calculate P_i of all the particles with (27) and (28) to determine the new $\mathbf{m}_k^{gbest}$. 7) Calculate $\mathbf{v}_{i,k}$ using (26) and update $\mathbf{m}_{i,k+1}$ according to (25).

III. Synthetic Data Test

A. Test a

We consider a two-layer model to test the sensitivity of the PP reflection coefficient using the AVA variation with respect to the five parameters. The base values of the upper and lower layers are 3383 m/s (α), 2438 m/s (β), 2.35 g/cm3 (ρ), 0.065 (ε), and 0.059 (δ), and 3688 m/s (α), 2774 m/s (β), 2.15 g/cm3 (ρ), 0.081 (ε), and 0.057 (δ), respectively. Several models can be obtained when one parameter varies and the others are constant. We compute the reflection coefficients as

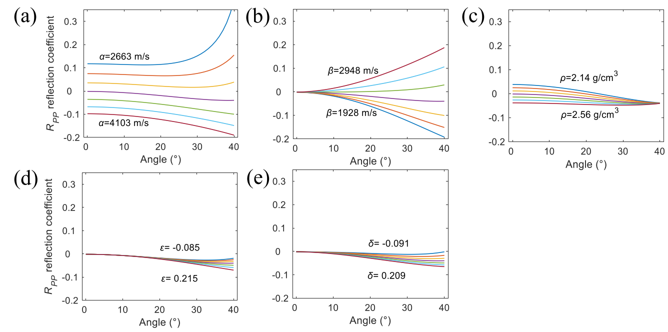

Fig. 4. PP reflection coefficients r_{pp} as a function of the incidence angle; r_{PP} variation with (a) α, (b) β, (c) ρ, (d) ε, and (e) δ. The colors correspond to different values of the parameters.

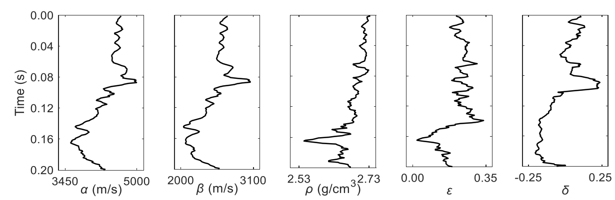

Fig. 5. Well logs for the synthetic data test after Backus averaging. From left to right the panels are α, β, ρ, and anisotropy parameters ε and δ.

a function of the incidence angle using the EGNI (shown in Fig. 4). The five parameters have different sensitivities, which affect the inversion dissimilarly. Coefficients ε and δ are sensitive mainly at mid and far offsets (angles), so that ensuring modeling accuracy at these ranges is important to improve the results. The exact equation is therefore recommended.

B. Test B

A well-log model is used to test the hierarchical inversion scheme with synthetic data. The inverted parameters are α, β, ρ, ε, and δ, and the true well models are shown in Fig. 5. A rock-physics-model-based method (Fig. 1) is used to obtain the anisotropy parameters ε and δ. The ranges of ε and δ are 0.02–0.3396 and -0.2039–0.2344, respectively, which includes weak, moderate, and strong anisotropy cases [15], [58]. We generate a input PP gather within angles ranging from 1° to 40° by convolving the EG reflection coefficient (1) with a Ricker wavelet of 35 Hz dominant frequency and zero phase. This synthetic example neglects multiples and noise. The input gather and the synthetic data of the inversion process use the same wavelet to eliminate the errors induced by an inaccurate wavelet. We compare the isotropic inversion based on exact Zoeppritz equation (EZI), the conventional direct inversion based on RAI, the nonlinear inversion based on the EGNI, and the proposed hierarchical inversion based on the EGHI.

The traditional isotropic inversion (EZI) is tested using the input data generated from the anisotropic model. The results are displayed in Fig. 6, including the P-wave velocity α, the S-wave velocity β, and the density ρ. Ignoring the anisotropic characteristics of the medium, the EZI fails to obtain acceptable results for the three elastic parameters, which can be effectively obtained if anisotropy is considered.

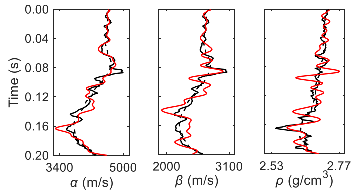

Fig. 6. Inversion results of the EZI assuming isotropy using the synthetic PP data generated from anisotropic media. From left to right the panels are α, β, and ρ. The red solid, black solid, and black dotted lines represent the results, true logs and initial models, respectively.

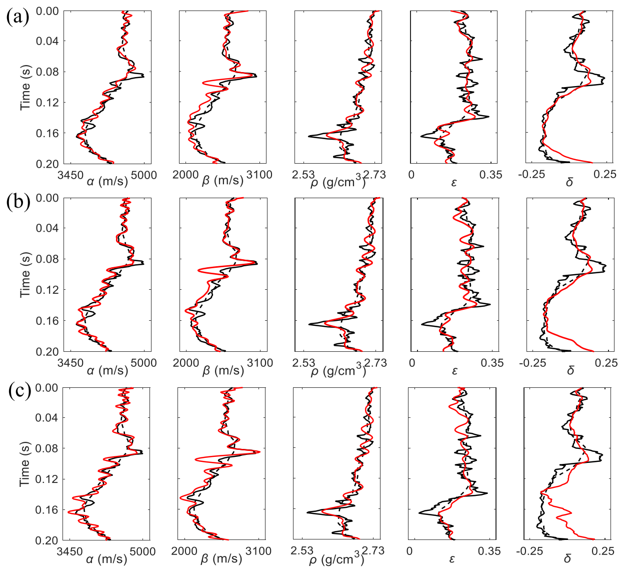

Fig. 7. Results of the RAI by synthetic PP data within different angle ranges. From left to right the panels are α, β, ρ, ε, and δ. From (a) 0° to 20°, (b) 0 to 30°, and (c) 0° to 40°. The red solid, black solid, and black dotted lines represent the inversion results, true logs, and initial models, respectively.

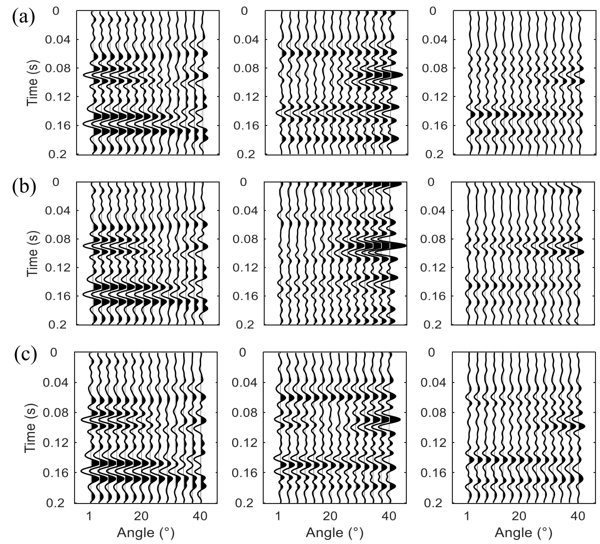

Fig. 8. (Left panels) Input gather, (Middle) synthetic gathers by the inversion, and (Right) corresponding residuals. Residuals using the RAI results in (a) Fig. 7(a), (b) Fig. 7(b), and (c) Fig. 7(c).

TABLE I
RRMSE AND CC BETWEEN THE INVERSION RESULTS IN FIG. 7 BY RAI AND THE TRUE MODELS

RAI results		α	β	ρ	ε	δ
0~20	RRMSE	2.3916	3.8148	0.5936	13.9848	29.0014
	CC	0.9621	0.9496	0.9016	0.8962	0.8730
0~30	RRMSE	3.0449	3.9488	0.5927	15.0158	40.2507
	CC	0.9406	0.9301	0.9024	0.8580	0.8258
0~40	RRMSE	2.7528	3.7796	0.5855	17.2613	67.3614
	CC	0.9593	0.9430	0.9050	0.8372	0.7697

The RAI requires the inversion of the reflectivities (including $\Delta\alpha/\alpha$, $\Delta\beta/\beta$, $\Delta\rho/\rho$, $\Delta\varepsilon$, and, $\Delta\delta$) first by a linear optimization, and then a conversion to α, β, ρ, ε, and δ by a trace integral algorithm [59]. We use the RAI method with input gathers of angles ranging from 0° to 20°, 0° to 30°, and 0° to 40°, and a fixed regularization parameter. The results and the residual profiles between the input gather and the synthetics generated by RAI results are shown in Figs. 7 and 8, respectively. Since the RAI is limited by weak anisotropy, the RAI does not yield accurate results, especially for the anisotropy parameters. Moreover, its low accuracy at large angles affects the estimation of ε and δ. The poor results are reflected in the low relative root-mean-square errors (RRMSE) and correlation coefficients (CC) between the true logs and the estimated results shown in Table I. Although the input gather within 0°–20° achieves the best estimations of ε and δ, unsatisfactory RRMSEs and CCs are obtained.

The results obtained with the proposed hierarchical scheme EGHI are shown in Fig. 9, including the preliminary linear inversion [Fig. 9(a)], the search windows generated with the linear constraints [Fig. 9(b)], and the results of the nonlinear stage [Fig. 9(c)]. Since one purpose of the linear inversion is to approximate the real model to reduce the length of the search window, a relatively large regularization parameter is adopted here to avoid the occurrence of anomalies. Compared with RAI, the estimation, especially the anisotropy parameters, is significantly improved by the linear inversion based on the EGLI, which is reflected by better RRMSEs and CCs in Table II. This indicates that the inversion method based on the exact reflection coefficient has a higher accuracy and is applicable for moderate to strong anisotropy. According to (29), we set the windows [in Fig. 9(b) (dashed pink lines)] to constrain the subsequent nonlinear inversion. Compared with the EGLI, the hierarchical scheme (EGHI) shows obviously better estimations of the CCs and RRMSEs (Table II). Density has the smallest RRMSE due to its smaller relative change. Besides, due to the lower sensitivity of δ to PP-wave data [33], the results for this parameter are worse than those of ε in both the linear and nonlinear steps. The red curves in Fig. 10(b) show the simultaneous results of EGNI (pure nonlinear inversion based on the PSO-VFSA and EGNI). The dashed lines in Fig. 10(a) define the search windows. The corresponding RRMSEs and CCs are shown in Table III, showing that the estimation of the parameters with the classical nonlinear inversion [Fig. 10(b)] are better than the preliminary results of EGLI [Fig. 9(a)] but worse than the final results of EGHI [Fig. 9(c)]. Fig. 11(a) and (b) shows the total RRMSEs

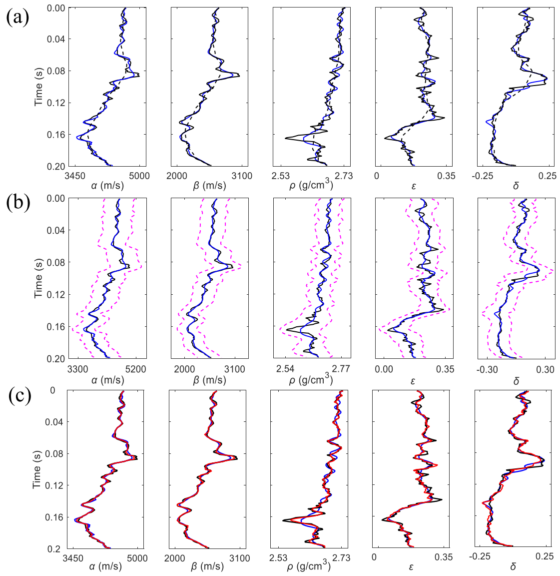

Fig. 9. Inversion results of the well model by the hierarchical inversion scheme. (a) Preliminary results by EGLI in the first step, (b) search windows used for the subsequent nonlinear inversion, and (c) final results of EGHI. The black solid and the black dotted lines are the real logs and the initial logs, respectively; pink dotted lines indicate the search windows; the blue and red curves are the results of the linear and nonlinear inversions, respectively.

TABLE II

RRMSE AND CC BETWEEN THE INVERSION RESULTS BY EGHI IN FIG. 9(A) AND (C) AND THE TRUE MODELS

Hierarchical inversion		α	β	ρ	ε	δ
EGLI	RRMSE	1.0778	1.1971	0.5024	8.3872	16.2988
	CC	0.9829	0.9801	0.9364	0.9411	0.9302
EGHI	RRMSE	0.5912	0.8768	0.2314	5.5062	12.6058
	CC	0.9975	0.9909	0.9806	0.9759	0.9683

TABLE III

RRMSE AND CC BETWEEN THE INVERSION RESULTS IN FIG. 10(B) BY EGNI AND THE TRUE MODELS

EGNI results	α	β	ρ	ε	δ
RRMSE	0.9933	1.0956	0.4114	7.9146	14.7061
CC	0.9897	0.9859	0.9510	0.9526	0.9414

of the iterations/time of the two steps in EGHI. Fig. 12 shows the residual errors between the input and synthetic gathers using the results of EGLI, EGHI, and EGNI.

Besides, EGHI has a higher computational efficiency, as illustrated by the convergence curves in Fig. 11 (they show the total RRMSEs of all five inverted parameters varying with the iteration/time). Compared with the pure nonlinear method EGNI [Fig. 11(b) (blue curve)], the EGHI aided by the linear stage yields better results and uses less time (EGHI employs 692 s with 3500 iterations, while EGNI requires 1705 s and 8000 iterations). Based on the linear results [Fig. 9(a)], the convergences are tested using the EGHI with two searching windows, including that in Fig. 9(b) (limited) and 1.25 times that window [the comparison is shown in Fig. 11(c)], which indicates that reducing the search range can effectively improve the results and save time (the latter needs 812 s and 4000 iterations). From the normalized RRMSEs of all three inversions [Fig. 11(d)], it can be seen that EGHI yields the faster convergence.

TABLE IV

RRMSE AND CC BETWEEN THE INVERSION RESULTS IN FIG. 13 AND TRUE MODELS

EGHI results		α	β	ρ	ε	δ
25Hz	RRMSE	1.2647	1.6463	0.4385	8.3069	15.463
	CC	0.9882	0.9846	0.9513	0.9480	0.9411
45Hz	RRMSE	1.1693	1.6936	0.4252	8.2612	15.251
	CC	0.9897	0.9872	0.9536	0.9496	0.9409
-5°	RRMSE	1.6554	1.9904	0.4716	10.108	20.708
	CC	0.9821	0.9783	0.9405	0.9358	0.9302
5°	RRMSE	1.6399	2.0119	0.4705	9.9225	21.056
	CC	0.9832	0.9754	0.9412	0.9374	0.9310

C. Test C

Since the forward operator of the linear and nonlinear methods is based on the convolution model, the wavelet affects the inversion results. In this test, by considering different wavelets with different (inaccurate) dominant frequency and phase, we analyze how this affects the results of the proposed EGHI algorithm. The input gather is the same to Test A and the actual dominant frequency and phase of the wavelet are 35 Hz and zero, respectively. Fig. 13(a) and (b) show the inversion results when the dominant frequencies are 25 and 45 Hz, deviating from the real one by -28.6% and 28.6%, respectively. Compared with the results using the accurate wavelet [Fig. 9(c)], an incorrect dominant frequency affects the accuracy of the results, especially for ρ, ε, and δ, which are quantified with the RRMSEs and CCs listed in Table IV. Fig. 13(c) and (d) show the inversion results of EGHI using phases of $-5°$ and $5°$. The corresponding RRMSEs and CCs are shown in Table IV, which indicate that, similar to the dominant frequency, phase inaccuracy worsens the inversion. Compared to the velocities, ρ, ε, and δ are more affected. Then, to ensure reliable estimations of the anisotropy parameters, an accurate wavelet should be considered.

Next, we investigate the effect of Gaussian random noise on gathers processed by EGHI. Fig. 14(a)–(c) show results with signal-to-noise ratio (SNR) $= 10, 5, 3$, respectively. The RRMSEs and CCs of the inverted parameters are shown in Table V. Fig. 14(c) shows that the quality of α and β are satisfactory even using the noisy gather with SNR $= 3$. Compared with the velocities, the errors of the other three parameters increase rapidly with decreasing SNR. Among them, two Thomsen's parameters are more affected by noise, especially δ. Given the data with SNR $= 3$, the estimation of δ has a relatively poor RRMSE (17.439) and CC (0.9377). The quality of the input gathers should be good enough to obtain relatively acceptable estimations of two Thomsen's parameters.

IV. FIELD DATA EXAMPLE

Prestack seismic data from a shale-gas reservoir survey is used to verify the feasibility of the proposed method. The target area is a set of shale layers from the Sichuan Basin

TABLE V
RRMSE AND CC BETWEEN THE INVERSION RESULTS IN FIG. 14 AND THE TRUE MODELS

EGHI results		α	β	ρ	ε	δ
S/N=10	RRMSE	0.7748	1.1493	0.3196	7.7002	14.477
	CC	0.9962	0.9898	0.9737	0.9523	0.9457
S/N=5	RRMSE	0.7742	1.1498	0.3656	8.1910	15.405
	CC	0.9960	0.9894	0.9654	0.9493	0.9412
S/N=3	RRMSE	0.7653	1.1728	0.4021	8.9346	17.439
	CC	0.9956	0.9883	0.9490	0.9404	0.9377

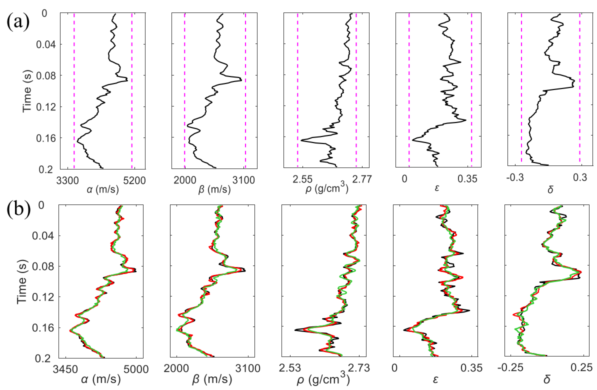

Fig. 10. Results corresponding to the well model by pure nonlinear inversion (EGNI) based on the PSO-VFSA and the exact Graebner equation. (a) Search windows and (b) comparison between the log curves (black) and the results by EGNI (green) and EGHI (red) [see Fig. 9(c)]. The dashed lines in (a) indicate the search windows.

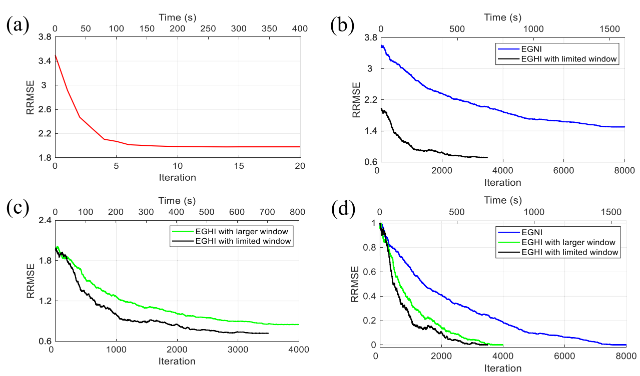

Fig. 11. RRMSEs variation with the iterations/time cost of the different inversion methods. (a) Linear stage of EGHI [EGLI in Fig. 9(a)]. (b) Comparison between the nonlinear stages of EGHI [Fig. 9(c)] with the limited window [Fig. 9(b)] and the pure nonlinear EGNI [Fig. 10(c)]. (c) Comparison between the nonlinear stage of EGHI with the limited window [Fig. 9(b)] and that with a relatively larger window. (d) Normalized RRMSEs of the results in (b) and (c).

of southwest China. The input seismic data are a set of common-image-point gathers after normal moveout (NMO) correction and offset-to-angle transform. To account for anisotropy, appropriate methods, such as nonhyperbolic moveout and double scanning, have been used for the NMO correction [9], [60], [61]. Other procedures for conventional prestack isotropic inversions, e.g., offset-to-angle transform, angle-dependent statistical wavelet estimation, and well ties are used in this inversion and implemented with the

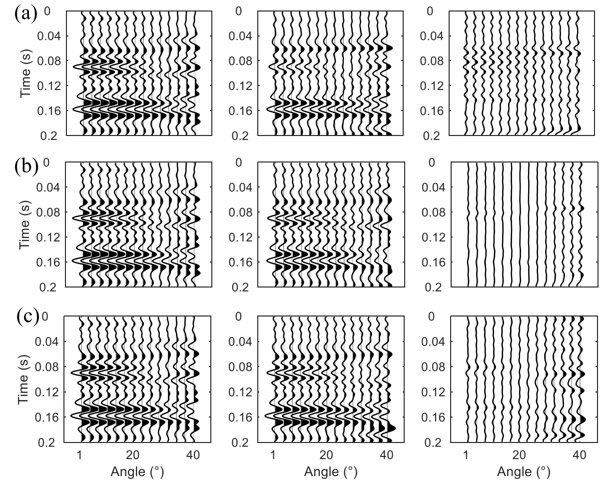

Fig. 12. (Left panels) Input gather, (Middle) synthetic gathers using the inversion results, and (Right) corresponding the residuals. (a) Residuals using the results of EGLI in Fig. 9(a). (b) EGHI in Fig. 9(c). (c) EGNI in Fig. 10(b).

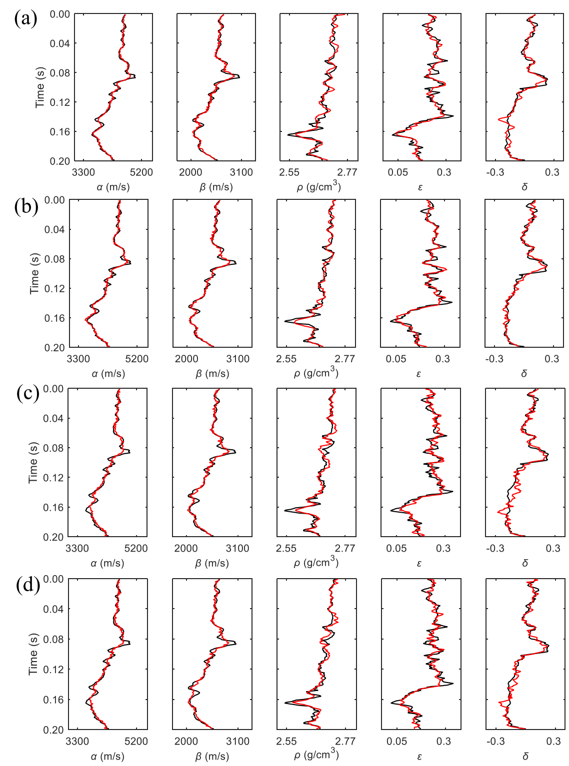

Fig. 13. Comparison of the inversion results of the EGHI algorithm using the wavelet with inaccurate dominant frequency and phase. (a) and (b) Dominant frequencies of 25 and 45 Hz, respectively, deviating from the real one (35 Hz) by -28.5% and 28.5%. (c) and (d) Phase of $-5°$ and 5°, respectively. The black and red curves are the real logs and the inversion results.

Hampson–Russell software (HRS). The well logs show that this formation is located between 2100 and 2350 m depth (1.25 and 1.48 s times) at CDP 50 and contains moderate clay content, a relatively large volume ratio of quartz and carbonates. The available logs are P- (α) and S-wave (β) velocities, and density (ρ). Since the anisotropy parameters cannot be measured directly, a rock-physics-based method is

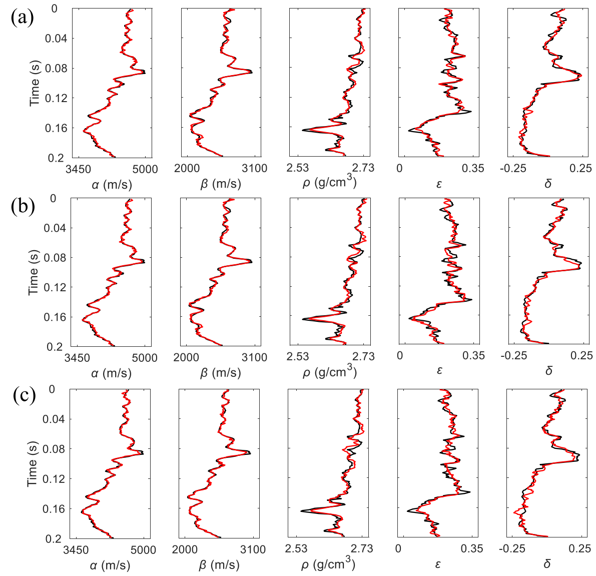

Fig. 14. Comparison of the inversion results of the EGHI algorithm using noisy gathers. (a) SNR = 10. (b) SNR = 5. (c) SNR = 3. The black and red curves correspond to the real logs and inversion results, respectively.

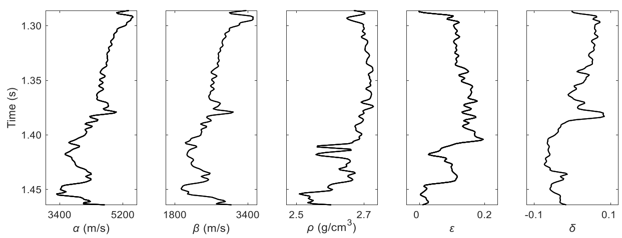

Fig. 15. Well logs at CDP 50. From left to right, the panels show the logs of α, β, ρ, and anisotropy parameters ε and δ.

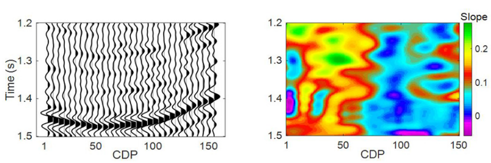

Fig. 16. (Left) Post-stack profile of the real seismic data and (Right) seismic slope estimated with the PWD algorithm.

adopted to predict the exact ε and δ at the well location. The feasibility of this method has been verified by model tests and real data applications [9]. The true parameters at the well location are displayed in Fig. 15.

In the proposed hierarchical scheme, the initial models significantly affect the accuracy of the linear inversion and the effectiveness of the nonlinear stage. Then a data-driven model-building algorithm is introduced to obtain reasonable initial models. It is known that the conventional model building method interpolates the well data laterally constrained by artificial interpretations of geological horizons. However, in the areas where there is a lack of interpreted horizons, the structural form of models will mainly be consistent with the upper or lower horizons. It is unreasonable, especially for the complex geological conditions. The data-driven model

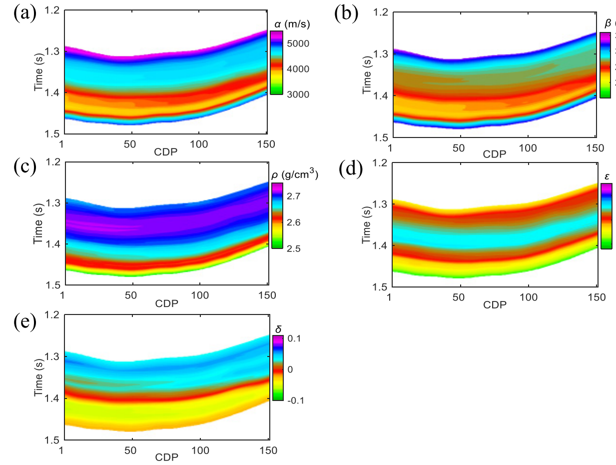

Fig. 17. Initial models of (a) α, (b) β, (c) ρ, (d) ε, and (e) δ obtained with the data-driven model building method.

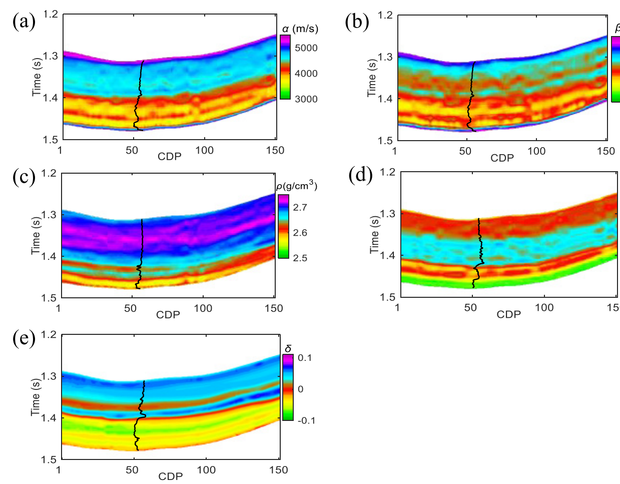

Fig. 18. Inversion results, namely, (a) α, (b) β, (c) ρ, (d) ε, and (e) δ, using the EGLI in the first stage. The corresponding well logs are given.

building needs no artificial interpretations and use the seismic slope attribute as lateral constraint. Using the PWD algorithm [55], [62], the seismic slope attribute [3] [Fig. 16(b)] is extracted from the post-stack seismic profile [Fig. 16(a)]. Using the slope attribute as lateral constraint, and a shaping-regularization-based inversion strategy [(34)] to interpolate the initial curves at the well position, we obtained the initial models shown in Fig. 17, which are consistent with the available geological information.

The input data is a set of PP-wave angle gathers. To suppress random noise, the gathers are processed into several constant angle sections (also called partial stacked sections) by stacking the data within a certain angle interval. We use these partial-stacked data to estimate the angle-dependent wavelets for the subsequent inversion. The statistical wavelets are estimated by HRS after amplitude normalization. In the first stage, the LBFGS-MT optimization is used to solve the objective function. To improve the nonlinear inversion, the linear results are used to reduce the search window and setup the initial models for the second stage, and the PSO-VFSA is adopted

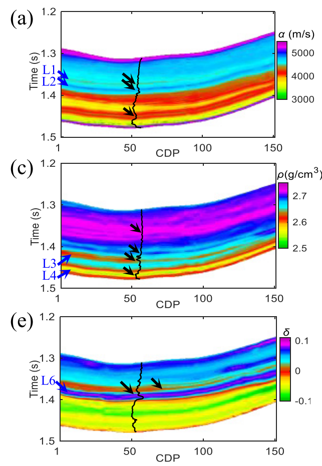

Fig. 19. Final inversion results, namely, (a) α, (b) β, (c) ρ, (d) ε, and (e) δ by the EGHI based on the results of Fig. 18. The corresponding well logs are given.

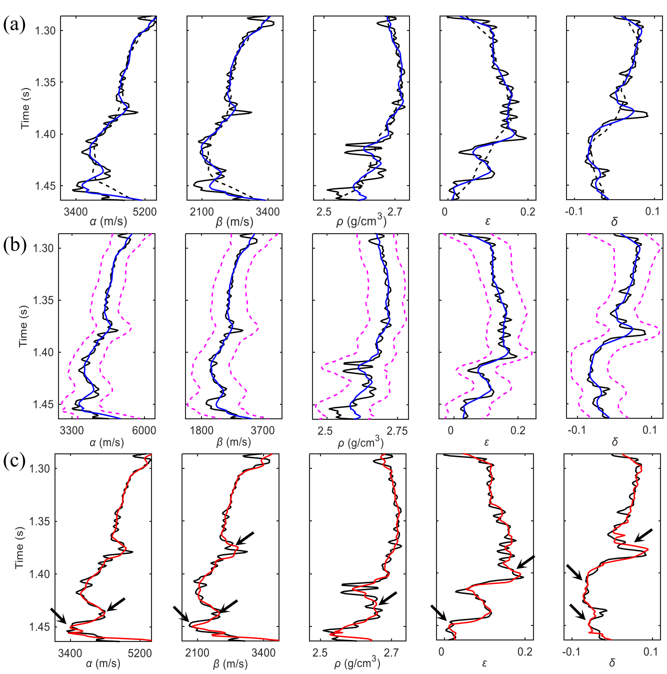

Fig. 20. Inversion results at the well location. (a) Preliminary results obtained with EGLI in the first stage. (b) Search windows used for the subsequent nonlinear inversion. (c) Final results of EGHI. The black solid and the black dotted lines denote the real logs and the initial logs, respectively; pink dotted lines indicate the search windows; the blue and red curves are the results of the linear and nonlinear inversions, respectively.

to achieve good convergence. A joint penta-variate Gaussian is used as the prior constraint to develop the stability of the multi-parameter inversion, and the covariance matrices of the linear and nonlinear inversion are computed from the initial models and the linear results, respectively.

Figs. 18 and 19 show the preliminary and final inverted profiles, respectively, for all CDPs (from 1 to 151) and the corresponding well profiles are included. Compared with the preliminary estimation (Fig. 18), the hierarchical inversion scheme gives higher resolution, better well consistency, and better lateral continuity (Fig. 19). The improved areas are marked with black arrows. The profiles with better continuity in Fig. 18 are in better agreement with the local geological structure. Moreover, the better resolution helps in highlighting several layers (marked with blue arrows), e.g., layers $L1$ and $L2$ in the P- and S-wave velocity sections [Fig. 19(a) and (b)], respectively, layers $L3$ and $L4$ in the density [Fig. 19(c)], layer $L5$ in ε [Fig. 19(d)], and layer $L6$ in δ [Fig. 19(e)].

The results for seismic traces near the borehole are shown in Fig. 20. Compared with the initial models (the black dashed curves extracted from Fig. 17), the inverted parameters of the first stage [Fig. 20(a) (blue lines)] are close to the true logs (black lines) to some extent but there are still errors. It can be seen from the RRMSEs and CCs, between the true and inverted parameters, shown in Table VI. According to (29), the search windows are determined by the linear results and represented by pink dashed lines in Fig. 20(b). The final results of the nonlinear inversion are shown in Fig. 20(c) (red lines). It can be seen that the second stage has a better consistency with well logs, and the improved areas are highlighted by black arrows in Fig. 20(c). Table VI shows that the EGHI scheme can effectively improve the inversion performance of the five parameters, especially for ρ, ε, and δ (see RRMSEs and CCs).

TABLE VI

RRMSE AND CC BETWEEN THE INVERSION RESULTS BY SEISMIC TRACES NEAR THE WELL AND THE REAL LOGS

Hierarchical inversion		α	β	ρ	ε	δ
EGLI	RRMSE	4.9202	5.8943	0.9529	13.963	25.306
	CC	0.9208	0.9109	0.8690	0.8911	0.8692
EGHI	RRMSE	3.2262	4.5901	0.7993	9.6345	16.176
	CC	0.9454	0.9395	0.9281	0.9238	0.9007

V. CONCLUSION

In this work, a hierarchical scheme that combines linear (first stage) and nonlinear prestack seismic anisotropic inversions (second stage) is proposed to improve the performance to estimate the velocities (α and β), density (ρ), and anisotropy parameters (ε and δ). The exact Graebner reflection coefficient is used as forward operator to improve the accuracy and applicability of the inversion to arbitrary contrast, arbitrary anisotropy, and large angles. To obtain reliable preliminary results in the first stage, a data-driven initial model building method and the LBFGS-MT optimization algorithm are adopted. The linear inversion results are used to reduce the search windows and constrain the nonlinear, second stage. A hybrid optimization PSO-VFSA is introduced to improve the estimations and save the computational time in nonlinear stage.

The proposed hierarchical scheme (EGHI) is tested with synthetic and field seismic data. The test shows that the inversion algorithm yields acceptable results for ε and δ in the moderate to strong anisotropy cases. Compared with the classical linear and nonlinear anisotropic inversion, EGHI improves the estimation and reduces the computation cost. Moreover, the inverted 2-D profiles and the results using seismic traces near the well validate the better accuracy and stability of the scheme.

Although a nonlinear optimization is used to obtain the final results, the accuracy and the computational efficiency of this inversion scheme relies on the provided initial models. If the initial model has a large deviation (from true model) or the low-frequency information is inaccurate, linear outcomes will deviate from the true models causing a large search window for the sub-sequential nonlinear inversion, which leads to unacceptable final results. Therefore, to ensure the accuracy of parameter estimations, a good initial model is still requested for the proposed hierarchical scheme. Besides, only PP seismic data is adopted in this work, which limits the improvement of the inverted properties. According to the previous studies [33], [63], [64], the converted PS data contains important information on anisotropy parameters, especially for δ. For better estimations, a joint PP-PS inversion is recommended in the future study for the hierarchical anisotropic inversion. It is generally accepted that the phenomenon of wave-induced fluid flow is dominant in subsurface rocks, and the dispersion caused by this loss mechanism affects the inversion results to some extent. In future work, we will consider the presence of attenuation and velocity dispersion based on this mechanism.

When using the proposed hierarchical inversion scheme, one should choose the forward operator (the approximated or exact reflection coefficient) according to the actual conditions. For moderate to strong anisotropy conditions and input data containing large offsets, the EGNI is recommended to improve the inversion accuracy. For weak anisotropy and small to moderate offsets, the RAI may be a choice due to its higher computational efficiency.

Appendix A

A. Derivations of Matrices $\mathbf{A}$ and $\mathbf{b}$

According to (3), the derivative of $\mathbf{A}$ with respect to the elastic parameters of the upper layer m_*^U can be expressed as

$$\frac{\partial \mathbf{A}}{\partial m_*^U} = \begin{bmatrix} \frac{\partial \ell_P^U}{\partial m_*^U} & \frac{\partial n_S^U}{\partial m_*^U} & 0 & 0 \\ \frac{\partial n_P^U}{\partial m_*^U} & -\frac{\partial \ell_S^U}{\partial m_*^U} & 0 & 0 \\ a^U & b^U & 0 & 0 \\ d^U & e^U & 0 & 0 \end{bmatrix} \tag{A-1}$$

with

$$\begin{aligned} a^U &= c_{55}^U\left(\frac{\partial s_P^U}{\partial m_*^U}\ell_P^U + s_P^U\frac{\partial \ell_P^U}{\partial m_*^U} + \frac{\partial p}{\partial m_*^U}n_P^U + p\frac{\partial n_P^U}{\partial m_*^U}\right) \\ &\quad + \frac{\partial c_{55}^U}{\partial m_*^U}\left(s_P^U\ell_P^U + pn_P^U\right) \end{aligned} \tag{A-2}$$

$$\begin{aligned} b^U &= c_{55}^U\left(\frac{\partial s_S^U}{\partial m_*^U}n_S^U + s_S^U\frac{\partial n_S^U}{\partial m_*^U} - \frac{\partial p}{\partial m_*^U}\ell_S^U - p\frac{\partial \ell_S^U}{\partial m_*^U}\right) \\ &\quad + \frac{\partial c_{55}^U}{\partial m_*^U}\left(s_S^U n_S^U - p\ell_S^U\right) \end{aligned} \tag{A-3}$$

$$\begin{aligned} d^U &= \frac{\partial c_{13}^U}{\partial m_*^U}p\ell_P^U + c_{13}^U\left(\ell_P^U\frac{\partial p}{\partial m_*^U} + p\frac{\partial \ell_P^U}{\partial m_*^U}\right) \\ &\quad + \frac{\partial c_{33}^U}{\partial m_*^U}n_P^U s_P^U + c_{33}^U\left(n_P^U\frac{\partial s_P^U}{\partial m_*^U} + s_P^U\frac{\partial n_P^U}{\partial m_*^U}\right) \end{aligned} \tag{A-4}$$

$$\begin{aligned} e^U &= \frac{\partial c_{13}^U}{\partial m_*^U}pn_S^U + c_{13}^U\left(n_S^U\frac{\partial p}{\partial m_*^U} + p\frac{\partial n_S^U}{\partial m_*^U}\right) \\ &\quad - \frac{\partial c_{33}^U}{\partial m_*^U}s_S^U\ell_S^U - c_{33}^U\left(\ell_S^U\frac{\partial s_S^U}{\partial m_*^U} + s_S^U\frac{\partial \ell_S^U}{\partial m_*^U}\right). \end{aligned} \tag{A-5}$$

The derivative of vector $\mathbf{b}$ can be obtained by differentiating both sides of (4) with respect to m_*^U

$$\frac{\partial \mathbf{b}}{\partial m_*^U} = \left(-\frac{\partial \ell_P^U}{\partial m_*^U} \;\; \frac{\partial n_P^U}{\partial m_*^U} \;\; a^U \;\; f^U\right)^{\mathrm{T}} \tag{A-6}$$

with

$$\begin{aligned} f^U &= -c_{13}^U\ell_P^U\frac{\partial p}{\partial m_*^U} - c_{13}^U p\frac{\partial \ell_P^U}{\partial m_*^U} - p\ell_P^U\frac{\partial c_{13}^U}{\partial m_*^U} \\ &\quad - c_{33}^U n_P^U\frac{\partial s_P^U}{\partial m_*^U} - c_{33}^U s_P^U\frac{\partial n_P^U}{\partial m_*^U} - s_P^U n_P^U\frac{\partial c_{33}^U}{\partial m_*^U}. \end{aligned} \tag{A-7}$$

Regarding the parameters of the lower layers m_*^L, the derivative $\partial\mathbf{A}/\partial m_*^L$ can be expressed as

$$\frac{\partial \mathbf{A}}{\partial m_*^L} = \begin{pmatrix} 0 & 0 & -\frac{\partial \ell_P^L}{\partial m_*^L} & -\frac{\partial n_S^L}{\partial m_*^L} \\ 0 & 0 & \frac{\partial n_P^L}{\partial m_*^L} & -\frac{\partial \ell_S^L}{\partial m_*^L} \\ 0 & 0 & a^L & b^L \\ 0 & 0 & -d^L & -e^L \end{pmatrix} \tag{A-8}$$

where a^L, b^L, d^L, and e^L have the same forms as those of the upper layer [see (A-2)–(A-5)]. From (5), we obtain the derivatives $\partial\ell_P/\partial m_*$, $\partial n_P/\partial m_*$, $\partial\ell_S/\partial m_*$, and $\partial n_S/\partial m_*$ in (A-1)–(A-8)

$$\begin{aligned} \frac{\partial \ell_P}{\partial m_*} &= \frac{1}{2\ell_P g_P^2} \\ &\quad \times\left[\left(s_P^2\frac{\partial c_{33}}{\partial m_*} + p^2\frac{\partial c_{55}}{\partial m_*} + 2c_{33}s_P\frac{\partial s_P}{\partial m_*} - \frac{\partial \rho}{\partial m_*} + 2c_{55}p \right.\right. \\ &\quad \left.\left. \times\frac{\partial p}{\partial m_*}\right)g_P - \left(c_{33}s_P^2 + c_{55}p^2 - \rho\right)\frac{\partial g_P}{\partial m_*}\right] \end{aligned} \tag{A-9}$$

$$\begin{aligned} \frac{\partial n_P}{\partial m_*} &= \frac{1}{2n_P g_P^2} \\ &\quad \times\left[\left(s_P^2\frac{\partial c_{55}}{\partial m_*} + p^2\frac{\partial c_{11}}{\partial m_*} + 2c_{55}s_P\frac{\partial s_P}{\partial m_*} - \frac{\partial \rho}{\partial m_*} + 2c_{11} \right.\right. \\ &\quad \left.\left. \times p\frac{\partial p}{\partial m_*}\right)g_P - \left(c_{55}s_P^2 + c_{11}p^2 - \rho\right)\frac{\partial g_P}{\partial m_*}\right] \end{aligned} \tag{A-10}$$

$$\begin{aligned} \frac{\partial \ell_S}{\partial m_*} &= \frac{1}{2\ell_S g_S^2} \\ &\quad \times\left[\left(s_S^2\frac{\partial c_{55}}{\partial m_*} + p^2\frac{\partial c_{11}}{\partial m_*} + 2c_{55}s_S\frac{\partial s_S}{\partial m_*} - \frac{\partial \rho}{\partial m_*} + 2c_{11} \right.\right. \\ &\quad \left.\left. \times p\frac{\partial p}{\partial m_*}\right)g_S - \left(c_{55}s_S^2 + c_{11}p^2 - \rho\right)\frac{\partial g_S}{\partial m_*}\right] \end{aligned} \tag{A-11}$$

$$\begin{aligned} \frac{\partial n_S}{\partial m_*} &= \frac{1}{2n_S g_S^2} \\ &\quad \times\left[\left(s_S^2\frac{\partial c_{33}}{\partial m_*} + p^2\frac{\partial c_{55}}{\partial m_*} + 2c_{33}s_S\frac{\partial s_S}{\partial m_*} - \frac{\partial \rho}{\partial m_*} + 2c_{55} \right.\right. \\ &\quad \left.\left. p\frac{\partial p}{\partial m_*}\right)g_S - \left(c_{33}s_S^2 + c_{55}p^2 - \rho\right)\frac{\partial g_S}{\partial m_*}\right] \end{aligned} \tag{A-12}$$

with

$$\frac{\partial g_P}{\partial m_*} = \left(s_P^2 + p^2\right)\frac{\partial c_{55}}{\partial m_*} + s_P^2\frac{\partial c_{33}}{\partial m_*} + 2s_P(c_{33}+c_{55})\frac{\partial s_P}{\partial m_*} + p^2\frac{\partial c_{11}}{\partial m_*} + 2p(c_{11}+c_{55})\frac{\partial p}{\partial m_*} - 2\frac{\partial \rho}{\partial m_*} \quad \text{(A-13)}$$

$$\frac{\partial g_S}{\partial m_*} = \left(s_s^2 + p^2\right)\frac{\partial c_{55}}{\partial m_*} + s_s^2\frac{\partial c_{33}}{\partial m_*} + 2s_S(c_{33}+c_{55})\frac{\partial s_S}{\partial m_*} + p^2\frac{\partial c_{11}}{\partial m_*} + 2p(c_{11}+c_{55})\frac{\partial p}{\partial m_*} - 2\frac{\partial \rho}{\partial m_*}. \quad \text{(A-14)}$$

Differentiating (6) with respect to m_*, we obtain

$$\frac{\partial s_P}{\partial m_*} = \frac{1}{4s_P}\left[\frac{\partial K_1}{\partial m_*} - \left(K_1\frac{\partial K_1}{\partial m_*} - 2K_3\frac{\partial K_2}{\partial m_*} - 2K_2\frac{\partial K_3}{\partial m_*}\right) \cdot \frac{1}{\sqrt{K_1^2 - 4K_2K_3}}\right] \quad \text{(A-15)}$$

$$\frac{\partial s_S}{\partial m_*} = \frac{1}{4s_S}\left[\frac{\partial K_1}{\partial m_*} + \left(K_1\frac{\partial K_1}{\partial m_*} - 2K_3\frac{\partial K_2}{\partial m_*} - 2K_2\frac{\partial K_3}{\partial m_*}\right) \cdot \frac{1}{\sqrt{K_1^2 - 4K_2K_3}}\right] \quad \text{(A-16)}$$

with

$$\begin{aligned}\frac{\partial K_1}{\partial m_*} &= \left[-\frac{\rho}{c_{33}^2} - \frac{\left(c_{13}^2 + 2c_{13}c_{55}\right)p^2}{c_{33}^2 c_{55}}\right]\frac{\partial c_{33}}{\partial m_*} + \left(\frac{1}{c_{33}} + \frac{1}{c_{55}}\right)\frac{\partial \rho}{\partial m_*} \\ &\quad + \left[\frac{\left(c_{33}c_{11} - c_{13}^2\right)p^2}{c_{33}c_{55}^2} - \frac{\rho}{c_{55}^2}\right]\frac{\partial c_{55}}{\partial m_*} \\ &\quad - \frac{p^2}{c_{55}}\frac{\partial c_{11}}{\partial m_*} + \frac{2p}{c_{55}}\left[\frac{c_{13}}{c_{33}}(c_{13}+2c_{55}) - c_{11}\right]\frac{\partial p}{\partial m_*} \\ &\quad + \frac{2c_{33}(c_{13}+c_{55})p^2}{c_{33}^2 c_{55}}\frac{\partial c_{13}}{\partial m_*} \end{aligned} \quad \text{(A-17)}$$

$$\frac{\partial K_2}{\partial m_*} = \frac{p^2}{c_{33}}\frac{\partial c_{11}}{\partial m_*} + \frac{\rho - c_{11}p^2}{c_{33}^2}\frac{\partial c_{33}}{\partial m_*} - \frac{1}{c_{33}}\frac{\partial \rho}{\partial m_*} + \frac{2c_{11}p}{c_{33}}\frac{\partial p}{\partial m_*} \quad \text{(A-18)}$$

$$\frac{\partial K_3}{\partial m_*} = \frac{\rho}{c_{55}^2}\frac{\partial c_{55}}{\partial m_*} + 2p\frac{\partial p}{\partial m_*} - \frac{1}{c_{55}}\frac{\partial \rho}{\partial m_*}. \quad \text{(A-19)}$$

According to (7), we have

$$\frac{\partial p}{\partial m_*^U} = \frac{p}{2\rho^U}\frac{\partial \rho^U}{\partial m_*^U} - \frac{p^3}{4\rho^U \sin^2\theta} \times \left(\frac{\partial c_{55}^U}{\partial m_*^U} + \sin^2\theta\frac{\partial c_{11}^U}{\partial m_*^U} + \cos^2\theta\frac{\partial c_{33}^U}{\partial m_*^U} + \frac{\partial \psi}{\partial m_*^U}\right) \quad \text{(A-20)}$$

$$\begin{aligned}\frac{\partial \psi^U}{\partial m_*^U} &= \frac{1}{\psi^U} \times \Bigg\{\left(c_{33}^U - c_{55}^U\right)\left(\frac{\partial c_{33}^U}{\partial m_*^U} - \frac{\partial c_{55}^U}{\partial m_*^U}\right) + \sin^2\theta \\ &\quad \times \Bigg[\left(c_{55}^U - c_{33}^U\right)\frac{\partial c_{11}^U}{\partial m_*^U} + 4\left(c_{13}^U + c_{55}^U\right)\frac{\partial c_{13}^U}{\partial m_*^U} \\ &\quad - \left(c_{11}^U + 2c_{33}^U - 3c_{55}^U\right)\frac{\partial c_{33}^U}{\partial m_*^U} + \left(c_{11}^U + 4c_{13}^U + 3c_{33}^U\right)\frac{\partial c_{55}^U}{\partial m_*^U}\Bigg] \\ &\quad + \sin^4\theta\Bigg[\left(c_{11}^U + c_{33}^U - 2c_{55}^U\right)\left(\frac{\partial c_{11}^U}{\partial m_*^U} + \frac{\partial c_{33}^U}{\partial m_*^U}\right) - 4\left(c_{13}^U + c_{55}^U\right) \\ &\quad \times \frac{\partial c_{13}^U}{\partial m_*^U} - 2\left(c_{11}^U + c_{33}^U + 2c_{13}^U\right)\frac{\partial c_{55}^U}{\partial m_*^U}\Bigg]\Bigg\} \end{aligned} \quad \text{(A-21)}$$

and

$$\frac{\partial p}{\partial m_*^L} = 0. \quad \text{(A-22)}$$

The elasticities and density are functions of the model parameters m^*, and the corresponding derivatives are derived according to (8)

$$\frac{\partial c_{33}}{\partial \alpha} = 2\alpha\rho, \quad \frac{\partial c_{33}}{\partial \rho} = \alpha^2, \quad \frac{\partial c_{33}}{\partial \beta} = \frac{\partial c_{33}}{\partial \varepsilon} = \frac{\partial c_{33}}{\partial \delta} = 0 \quad \text{(A-23a)}$$

$$\frac{\partial c_{55}}{\partial \alpha} = \frac{\partial c_{55}}{\partial \varepsilon} = \frac{\partial c_{55}}{\partial \delta} = 0, \quad \frac{\partial c_{55}}{\partial \beta} = 2\beta\rho, \quad \frac{\partial c_{55}}{\partial \rho} = \beta^2 \quad \text{(A-23b)}$$

$$\frac{\partial c_{11}}{\partial \alpha} = 2\alpha\rho(2\varepsilon+1), \frac{\partial c_{11}}{\partial \rho} = \alpha^2(2\varepsilon+1), \quad \frac{\partial c_{11}}{\partial \beta} = \frac{\partial c_{11}}{\partial \delta} = 0 \quad \text{(A-23c)}$$

$$\frac{\partial c_{11}}{\partial \varepsilon} = 2\alpha^2\rho, \quad \frac{\partial c_{13}}{\partial \alpha} = \frac{\alpha\rho}{\eta}\left[2\alpha^2(\varepsilon+\delta+1) - \beta^2(\varepsilon+2)\right] \quad \text{(A-23d)}$$

$$\frac{\partial c_{13}}{\partial \beta} = -\frac{\beta\rho}{\eta}\left[\alpha^2(\varepsilon+2) + 2\beta^2\right] - 2\beta\rho, \quad \frac{\partial c_{13}}{\partial \rho} = \eta - \beta^2, \quad \frac{\partial c_{13}}{\partial \delta} = \frac{\alpha^4\rho}{2\eta} \quad \text{(A-23e)}$$

$$\frac{\partial c_{13}}{\partial \varepsilon} = \frac{\alpha^2\rho}{2\eta}\left(\alpha^2 - \beta^2\right), \quad \frac{\partial \rho}{\partial \alpha} = \frac{\partial \rho}{\partial \beta} = \frac{\partial \rho}{\partial \varepsilon} = \frac{\partial \rho}{\partial \delta} = 0, \quad \frac{\partial \rho}{\partial \rho} = 1. \quad \text{(A-23f)}$$

Based on the derivatives above, we obtain the partial derivatives of the reflection coefficients and then the Fréchet derivatives of the forward modeling.

Appendix B

According to Chen *et al.* [57] and Huang *et al.* [62], the shaping regularization method solves the inverse problem (31) via the following iteration

$$\mathbf{X}_{n+1} = \mathbf{S}\left[\mathbf{X}_n + \lambda\Theta^{\mathrm{T}}(\mathbf{Y} - \Theta\mathbf{X}_n)\right] \quad \text{(B-1)}$$

where $\mathbf{S}$ is the shaping operator and Θ^{T} is the adjoint of the forward operator. When $\mathbf{S}$ and Θ are both linear operators, we have

$$\hat{\mathbf{X}} = \left(\mathbf{I} - \mathbf{S} + \lambda \mathbf{S}\Theta^{\mathrm{T}}\Theta\right)^{-1} \lambda \mathbf{S}\Theta^{\mathrm{T}}\mathbf{Y}. \tag{B-2}$$

The structural smoothness shaping operator is

$$\mathbf{S} = \mathbf{P}\mathbf{H}\mathbf{H}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} \tag{B-3}$$

where $\mathbf{H}$ is a triangle smoothing operator [57], [62], and $\mathbf{P}$ is a summation operator. Since $\mathbf{PH}$, $\mathbf{P}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}$, and $\Theta^{\mathrm{T}}\Theta$ are diagonal square matrices, substituting (B-3) into (B-2) and setting $\lambda = 1/\upsilon^2$, we can obtain (34).

References

[1] Z. Zong, X. Yin, and G. Wu, "Geofluid discrimination incorporating poroelasticity and seismic reflection inversion," *Surv. Geophys.*, vol. 36, no. 5, pp. 659–681, Jul. 2015.

[2] Z. Gao, Z. Pan, and J. Gao, "Multimutation differential evolution algorithm and its application to seismic inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3626–3636, Jun. 2016.

[3] G. Huang *et al.*, "The slope-attribute-regularized high-resolution prestack seismic inversion," *Surv. Geophys.*, vol. 42, no. 3, pp. 625–671, May 2021, doi: 10.1007/s10712-021-09636-6.

[4] G. Huang, X. Chen, C. Luo, and Y. Chen, "Mesoscopic wave-induced fluid flow effect extraction by using frequency-dependent prestack waveform inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6510–6524, Aug. 2021, doi: 10.1109/TGRS.2020.3028032.

[5] C. Luo, G. Huang, X. Chen, and Y. Chen, "Registration-free multicomponent joint AVA inversion using optimal transport," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022, doi: 10.1109/TGRS.2021.3063271.

[6] Z. Zong and L. Ji, "Model parameterization and amplitude variation with angle and azimuthal inversion in orthotropic media," *Geophysics*, vol. 86, no. 1, pp. R1–R14, Jan. 2021.

[7] Z. Zong, Y. Wang, K. Li, and X. Yin, "Broadband seismic inversion for low-frequency component of the model parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5177–5184, Sep. 2018.

[8] Z. Gao, Z. Pan, C. Zuo, J. Gao, and Z. Xu, "An optimized deep network representation of multimutation differential evolution and its application in seismic inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4720–4734, Jul. 2019.

[9] F. Zhang, T. Zhang, and X. Li, "Seismic amplitude inversion for the transversely isotropic media with vertical axis of symmetry," *Geophys. Prospecting*, vol. 67, no. 9, pp. 2368–2385, Nov. 2019.

[10] J. M. Carcione and P. Avseth, "Rock-physics templates for clay-rich source rocks," *Geophysics*, vol. 80, no. 5, pp. D481–D500, Sep. 2015.

[11] J. Ba, W. Xu, L.-Y. Fu, J. M. Carcione, and L. Zhang, "Rock anelasticity due to patchy saturation and fabric heterogeneity: A double double-porosity model of wave propagation," *J. Geophys. Res., Solid Earth*, vol. 122, no. 3, pp. 1949–1976, 2017.

[12] P. Ding, D. Wang, G. Di, and X. Li, "Investigation of the effects of fracture orientation and saturation on the V_p/V_s ratio and their implications," *Rock Mech. Rock Eng.*, vol. 52, no. 9, pp. 3293–3304, Sep. 2019.

[13] P. Ding, D. Wang, and X.-Y. Li, "An experimental study on scale-dependent velocity and anisotropy in fractured media based on artificial rocks with controlled fracture geometries," *Rock Mech. Rock Eng.*, vol. 53, no. 7, pp. 3149–3159, Jul. 2020.

[14] F. Zhang, L. Wang, and X. Li, "Characterization of a shale-gas reservoir based on a seismic AVO inversion for VTI media and quantitative seismic interpretation," *Interpretation*, vol. 8, no. 1, p. SA11, 2020.

[15] L. Thomsen, "Weak elastic anisotropy," *Geophysics*, vol. 51, no. 10, pp. 1954–1966, Oct. 1986.

[16] J. Wright, "The effects of transverse isotropy on reflection amplitude versus offset," *Geophysics*, vol. 52, no. 4, pp. 564–567, Apr. 1987.

[17] A. Rüger, "P-wave reflection coefficients for transversely isotropic models with vertical and horizontal axis of symmetry," *Geophysics*, vol. 62, no. 3, pp. 713–722, 1997.

[18] J. M. Carcione, "Reflection and transmission of *qP-qS* plane waves at a plane boundary between viscoelastic transversely isotropic media," *Geophys. J. Int.*, vol. 129, no. 3, pp. 669–680, Jun. 1997.

[19] P. F. Daley and F. Hron, "Reflection and transmission coefficients for transversely isotropic media," *Bull. Seismol. Soc. Amer.*, vol. 67, no. 3, pp. 661–675, 1977.

[20] M. Graebner, "Plane-wave reflection and transmission coefficients for a transversely isotropic solid," *Geophysics*, vol. 57, no. 11, pp. 1512–1519, Nov. 1992.

[21] M. Schoenberg and J. Protazio, "'Zoeppritz' rationalized and generalized to anisotropy," *J. Seismic Explor.*, vol. 1, no. 2, pp. 125–144, 1992.

[22] B. Ursin and G. Haugen, "Weak-contrast approximation of the elastic scattering matrix in anisotropic media," *Pure Appl. Geophys.*, vol. 148, no. 3, pp. 686–714, 1996.

[23] A. Rüger, "Variation of P-wave reflectivity with offset and azimuth in anisotropic media," *Geophysics*, vol. 63, no. 3, pp. 935–947, 1998.

[24] V. Vavryčuk and I. Pšenčík, "PP-wave reflection coefficients in weakly anisotropic elastic media," *Geophysics*, vol. 63, no. 6, pp. 2129–2141, Nov. 1998.

[25] M. Zillmer, D. Gajewski, and B. M. Kashtan, "Anisotropic reflection coefficients for a weak-contrast interface," *Geophys. J. Int.*, vol. 132, no. 1, pp. 159–166, Jan. 1998.

[26] D. C. Booth and S. Crampin, "The anisotropic reflectivity technique: Theory," *Geophys. J. Int.*, vol. 72, no. 3, pp. 755–766, 1983.

[27] S. Mallick and L. N. Frazer, "Computation of synthetic seismograms for stratified azimuthally anisotropic media," *J. Geophys. Res.*, vol. 95, pp. 8513–8526, Jun. 1990.

[28] J. Carcione, *Wave Fields in Real Media: Theory and Numerical Simulation of Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Media*. Amsterdam, The Netherlands: Elsevier, 2014.

[29] R.-E. Plessix and J. Bork, "Quantitative estimate of VTI parameters from AVA responses," *Geophys. Prospecting*, vol. 48, no. 1, pp. 87–108, Jan. 2000.

[30] R. Lin and L. Thomsen, "Extracting polar anisotropy parameters from seismic data and well logs," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2013, pp. 310–313.

[31] J. Lu, Y. Wang, J. Chen, and Y. An, "Joint anisotropic amplitude variation with offset inversion of PP and PS seismic data," *Geophysics*, vol. 83, no. 2, pp. N31–N50, Mar. 2018.

[32] A. Rüger, "Variation of P-wave reflectivity with offset and azimuth in anisotropic media," *Geophysics*, vol. 63, no. 3, pp. 713–722, 1997.

[33] C. Luo, J. Ba, J. M. Carcione, G. Huang, and Q. Guo, "Joint PP and PS pre-stack AVA inversion for VTI medium based on the exact Graebner equation," *J. Petroleum Sci. Eng.*, vol. 194, Nov. 2020, Art. no. 107416.

[34] M. Sen and P. Stoffa, "Nonlinear seismic waveform inversion in one dimension using simulated annealing," *Geophysics*, vol. 56, no. 10, pp. 1624–1638, 1991.

[35] C. Luo, X. Li, and G. Huang, "Pre-stack AVA inversion by using propagator matrix forward modeling," *Pure Appl. Geophys.*, vol. 176, no. 10, pp. 4445–4476, Oct. 2019.

[36] G. Huang, X. Chen, C. Luo, and X. Li, "Prestack waveform inversion by using an optimized linear inversion scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5716–5728, Aug. 2019.

[37] C. Luo, J. Ba, J. M. Carcione, G. Huang, and Q. Guo, "Joint PP and PS pre-stack seismic inversion for stratified models based on the propagator matrix forward engine," *Surv. Geophys.*, vol. 41, no. 5, pp. 987–1028, Sep. 2020.

[38] A. Padhi and S. Mallick, "Multicomponent pre-stack seismic waveform inversion in transversely isotropic media using a non-dominated sorting genetic algorithm," *Geophys. J. Int.*, vol. 5, no. 3, p. 589, 2014.

[39] Q. Guo, H. Zhang, F. Han, and Z. Shang, "Prestack seismic inversion based on anisotropic Markov random field," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1069–1079, Feb. 2018.

[40] Q. Guo, H. Zhang, H. Cao, W. Xiao, and F. Han, "Hybrid seismic inversion based on multi-order anisotropic Markov random field," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 407–420, Jan. 2020.

[41] Q. Guo, J. Ba, and C. Luo, "Prestack seismic inversion with data-driven MRF-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 7122–7136, Aug. 2021.

[42] W. Alemie and M. D. Sacchi, "High-resolution three-term AVO inversion by means of a trivariate Cauchy probability distribution," *Geophysics*, vol. 76, no. 3, pp. R43–R55, May 2011.

[43] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Netw.*, vol. 12, no. 1, pp. 145–151, 1999.

[44] B. She, A. Fournier, Y. Wang, and G. Hu, "Incorporating momentum acceleration techniques applied in deep learning into traditional optimization algorithms," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2019, pp. 2668–2673.

[45] P. Wolfe, "Convergence conditions for ascent methods. II: Some corrections," *SIAM Rev.*, vol. 13, no. 2, pp. 185–188, Apr. 1971.

[46] L. Ingber, "Very fast simulated re-annealing," *Math. Comput. Model.*, vol. 12, no. 8, pp. 967–973, 1989.

[47] L.-L. Li, L. Wang, and L.-H. Liu, "An effective hybrid PSOSA strategy for optimization and its application to parameter estimation," *Appl. Math. Comput.*, vol. 179, no. 1, pp. 135–146, Aug. 2006.

[48] M. Sen and L. Stoffa, *Global Optimization Methods in Geophysical Inversion*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[49] Q. Guo, J. Ba, C. Luo, and S. Xiao, "Stability-enhanced prestack seismic inversion using hybrid orthogonal learning particle swarm optimization," *J. Petroleum Sci. Eng.*, vol. 192, Sep. 2020, Art. no. 107313.

[50] R. Shaw and S. Srivastava, "Particle swarm optimization: A new tool to invert geophysical data," *Geophysics*, vol. 72, no. 2, pp. F75–F83, Mar. 2007.

[51] C. Tsallis, "Possible generalization of Boltzmann–Gibbs statistics," *J. Stat. Phys.*, vol. 52, nos. 1–2, pp. 479–487, 1988.

[52] F. Zhang, X.-Y. Li, and K. Qian, "Estimation of anisotropy parameters for shale based on an improved rock physics model. Part 1: Theory," *J. Geophys. Eng.*, vol. 14, no. 1, pp. 143–158, Feb. 2017.

[53] G. Mavko, T. Mukerji, and J. Dvorkin, *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[54] Y. Chen, H. Chen, K. Xiang, and X. Chen, "Geological structure guided well log interpolation for high-fidelity full waveform inversion," *Geophys. J. Int.*, vol. 207, no. 2, pp. 1313–1331, 2016.

[55] J. Claerbout, *Earth Soundings Analysis: Processing Versus Inversion*. Oxford, U.K.: Blackwell, 2004.

[56] S. Fomel, "Shaping regularization in geophysical-estimation problems," *Geophysics*, vol. 72, no. 2, pp. R29–R36, 2007.

[57] Y. Chen, X. Chen, Y. Wang, and S. Zu, "The interpolation of sparse geophysical data," *Surv. Geophys.*, vol. 40, no. 1, pp. 73–105, Jan. 2019.

[58] C. H. Sondergeld and C. S. Rai, "Elastic anisotropy of shales," *Lead. Edge*, vol. 30, no. 3, pp. 324–331, Mar. 2011.

[59] G. Huang, X. Chen, C. Luo, M. Bai, and Y. Chen, "Time-lapse seismic difference- and-joint prestack AVA inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9132–9143, Nov. 2021, doi: 10.1109/TGRS.2020.3038762.

[60] I. Tsvankin and L. Thomsen, "Nonhyperbolic reflection moveout in anisotropic media," *Geophysics*, vol. 59, no. 8, pp. 1290–1304, Aug. 1994.

[61] V. Grechka and I. Tsvankin, "Processing-induced anisotropy," *Geophysics*, vol. 67, no. 6, pp. 1920–1928, 2002.

[62] G. Huang, X. Chen, C. Luo, and Y. Chen, "Geological structure-guided initial model building for prestack AVO/AVA inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1784–1793, Feb. 2021, doi: 10.1109/TGRS.2020.2998044.

[63] V. Grechka, I. Tsvankin, A. Bakulin, C. Signer, and J. O. Hansen, "Anisotropic inversion and imaging of PP and PS reflection data in the North Sea," *Lead. Edge*, vol. 21, no. 5, pp. 90–97, 2002.

[64] V. Grechka, I. Pšenčík, I. Ravve, and I. Tsvankin, "Introduction to special section: Seismic anisotropy," *Geophysics*, vol. 82, no. 2, pp. 1–3, 2017.

Cong Luo received the B.S. and M.S. degrees in exploration geophysics from the China University of Petroleum (East China), Qingdao, China, in 2012 and 2015, respectively, and the Ph.D. degree in exploration geophysics from the China University of Petroleum (Beijing), Beijing, China, in 2019.

She is currently an Associate Professor with Hohai University, Nanjing, China. Her research interests include seismic inversion for isotropic and anisotropic media.

Jing Ba received the B.S. degree in geophysics from Yunnan University, Kunming, China, in 2001, the M.S. degree in geophysics from the Institute of Geology, China Earthquake Administration, Beijing, China, in 2004, and the Ph.D. degree in mechanics from Tsinghua University, Beijing, in 2008.

He is currently a Professor with Hohai University, Nanjing, China. His research interests include rock physics theory, seismic forwarding modeling, and reservoir identification.

José M. Carcione received the "Licenciado in Ciencias Físicas" degree from Buenos Aires University, Buenos Aires, Argentina, in 1978, the "Dottore in Fisica" degree from the University of Milan, Milan, Italy, in 1984, and the Ph.D. degree in geophysics from Tel-Aviv University, Tel Aviv, Israel, in 1987.

He is a Senior Geophysicist at the National Institute of Oceanography and Applied Geophysics-OGS, Trieste, Italy, and also a Professor with Hohai University, Nanjing, China. His research deals with numerical modeling, the theory of wave propagation in acoustic and electromagnetic media, and their application to geophysics.