

Original Paper

Incremental dimensionality reduction for efficiently solving Bayesian inverse problems

Qing-Qing Li^a, Bo Yu^{b,*}, Jia-Liang Xu^c, Ning Wang^b, Shi-Chao Wang^b, Hui Zhou^d

^a State Key Laboratory of Deep Oil and Gas, China University of Petroleum (East China), Qingdao, 266580, Shandong, China

^b The School of Earth Sciences, Northeast Petroleum University, Daqing, 163318, Heilongjiang, China

^c Sanya Offshore Oil and Gas Research Institute, Northeast Petroleum University, Sanya, 572025, Hainan, China

^d State Key Laboratory of Petroleum Resources and Engineering, CNPC Key Lab of Geophysical Exploration, China University of Petroleum (Beijing), Beijing, 102249, China

ARTICLE INFO

Article history:

Received 7 April 2025

Received in revised form

29 June 2025

Accepted 5 August 2025

Available online 8 August 2025

Edited by Meng-Jiao Zhou

Keywords:

Dimension reduction

Seismic inversion

Discrete cosine transform

ABSTRACT

The inversion of large sparse matrices poses a major challenge in geophysics, particularly in Bayesian seismic inversion, significantly limiting computational efficiency and practical applicability to large-scale datasets. Existing dimensionality reduction methods have achieved partial success in addressing this issue. However, they remain limited in terms of the achievable degree of dimensionality reduction. An incremental deep dimensionality reduction approach is proposed herein to significantly reduce matrix size and is applied to Bayesian linearized inversion (BLI), a stochastic seismic inversion approach that heavily depends on large sparse matrices inversion. The proposed method first employs a linear transformation based on the discrete cosine transform (DCT) to extract the matrix's essential information and eliminate redundant components, forming the foundation of the dimensionality reduction framework. Subsequently, an innovative iterative DCT-based dimensionality reduction process is applied, where the reduction magnitude is carefully calibrated at each iteration to incrementally reduce dimensionality, thereby effectively eliminating matrix redundancy in depth. This process is referred to as the incremental discrete cosine transform (IDCT). Ultimately, a linear IDCT-based reduction operator is constructed and applied to the kernel matrix inversion in BLI, resulting in a more efficient BLI framework. The proposed method was evaluated through synthetic and field data tests and compared with conventional dimensionality reduction methods. The IDCT approach significantly improves the dimensionality reduction efficiency of the core inversion matrix while preserving inversion accuracy, demonstrating prominent advantages in solving Bayesian inverse problems more efficiently.

© 2025 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Seismic inversion is a critical technique in geophysical exploration, enabling the transformation of seismic reflection data into quantitative subsurface properties such as velocity, density, and lithology (Chen et al., 2024a; Wang et al., 2024a; Sun et al., 2025). To date, seismic inversion significantly advances the characterization of subsurface structures and fluid distributions, serving a critical role in hydrocarbon exploration, mineral prospecting, and carbon sequestration monitoring (Ding et al., 2021; Zhang et al.,

2022; Chen et al., 2024b). Over the past six decades, seismic inverse problems have predominantly been solved within the Bayesian framework, evolving into two principal categories: deterministic and stochastic inversion. Deterministic seismic inversion employs optimization techniques to derive the optimal solution for target parameters from seismic data, whereas stochastic inversion not only estimates subsurface parameters from observational data but also quantifies the associated uncertainty (Bosch et al., 2010; Alemie and Sacchi, 2011; Yang et al., 2023).

Due to the ill-posed nature of seismic inversion problems and the uncertainties introduced by errors and noise in various types of observed datasets, such as seismic and well log data, seismic inversion exhibits strong multimodality (Downton, 2005; Chen et al., 2025). Therefore, in addition to seeking the optimal solution, it is crucial to assess the uncertainty of the inversion solution.

* Corresponding author.

E-mail address: zgzyzcd@163.com (B. Yu).

Peer review under the responsibility of China University of Petroleum (Beijing).

The stochastic inversion strategy offers distinct advantages in this regard (Yu et al., 2021). The ultimate objective of stochastic inversion is not to obtain the optimal solution for the parameters, but rather to derive a set of solutions. This solution set is typically represented by the posterior probability distribution of the parameters within a statistical framework, capturing the distribution space of the inversion parameters and characterizing the uncertainty of the inversion process.

Generally, seismic stochastic inversion can be broadly categorized into two types: iterative geostatistical stochastic inversion and the Kalman stochastic inversion (Pereira et al., 2019). The foundational framework of geostatistical inversion comprises geostatistics and Monte Carlo optimization (Pereira et al., 2020), and it has been extensively studied in the field of petroleum geophysics (Penna and Lupinacci, 2024). Another type of stochastic inversion essentially can be defined as Kalman stochastic inversion, which mainly renews the model parameters using the Kalman gain calculated from the well-log and seismic data (Zhang and Oliver, 2011). The main advantages of Kalman inversion lie in its stability and efficiency compared with geostatistical inversion. Currently, Kalman inversion can also be categorized into linear and nonlinear types. The linear method is essentially Bayesian linearized inversion (BLI), which directly solves the posterior probability distribution of the parameters using a linearization approach within the Bayesian framework (Buland and Omre, 2003; Grana et al., 2022; Shi et al., 2024). Compared with linear Kalman inversion (BLI), nonlinear Kalman inversion like ensemble Kalman inversion is beneficial for solving complex nonlinear inverse problems (Cao et al., 2024), such as the inversion of complex reservoir physical properties and full-waveform inversion (Thurin and Brossier, 2019).

Although Kalman inversion combines high efficiency and stability, it is constrained by the inversion of large sparse matrices. Both linear and nonlinear Kalman inversions require solving large kernel matrices to compute the Kalman gain. Given the large scale of field data in practical applications, the difficulty of matrix inversion becomes significantly high, which severely impacts the practicality of Kalman inversion. Choosing an appropriate dimensionality reduction method can help alleviate this issue to some extent, reducing the limitations of large matrix dimensionality reduction on Kalman stochastic inversion.

Dimensionality reduction plays a crucial role in fields that involve massive data processing, such as geophysics, mathematical geology, computer science, and image processing (Marzouk and Najm, 2009; Esser et al., 2012). In recent years, with the rapid development of deep learning, dimensionality reduction techniques have been extensively studied. A series of linear and nonlinear reduction methods have been proposed, which can effectively improve the efficiency of deep learning algorithms (Cunningham and Ghahramani, 2015). Dimensionality reduction of ultra-large matrices is one of the key challenges in information reduction (Bunte et al., 2012). The solution of high-dimensional inverse problems essentially involves the inversion of ultra-large matrices, which poses significant computational challenges in practical applications. To address this, various dimensionality reduction inversion methods have been proposed. By discarding part of the redundant information, these approaches significantly improve the computational efficiency of solving high-dimensional nonlinear inverse problems without notably compromising inversion accuracy (Aravkin et al., 2012; Atkinson and Zabaras, 2019). Zahm et al. (2022) introduced dimensionality reduction strategies into the Bayesian framework to accelerate the solution of nonlinear inverse problems. As a typical class of inverse problems, seismic inversion methods also commonly require information reduction techniques (Jumah and Herrmann, 2014; Gao et al., 2020). Among various seismic inversion approaches, two

methods exhibit the highest demand for dimensionality reduction. The first is full waveform inversion (FWI), where dimensionality reduction strategies can effectively reduce the model space (Yin et al., 2025), thereby improving computational efficiency. The second is stochastic inversion, particularly the Kalman-based stochastic inversion methods discussed earlier.

Previous studies have attempted to address the challenges of large matrix inversion and high-dimensional parameter estimation in Kalman-based stochastic inversion by incorporating dimensionality reduction strategies. Grana et al. (2019) proposed a data dimensionality reduction method that is of significant importance for improving the efficiency of Kalman gain computation in Bayesian linearized inversion. Liu et al. (2022) also adopted this dimensionality reduction strategy to enhance the efficiency of Bayesian stochastic inversion. These methods are mainly based on principal component analysis (PCA). Although they alleviate the curse of dimensionality to some extent, they also introduce additional computational burden during the reduction process. Yu et al. (2024) proposed an efficient dimensionality reduction strategy for linearized Kalman inversion based on discrete cosine transform (DCT), which effectively alleviates the issue of large matrix dimensionality reduction, without increasing the computational cost of the inversion method itself.

The above dimensionality reduction methods have reduced the difficulty of inverting large sparse matrices to some extent. These methods are not only significant for seismic inversion but also have practical value for other geophysical processing and interpretation techniques that require huge sparse matrix inverse. However, the existing dimensionality reduction methods typically reduce the matrix dimension to around 40%–50% of the original matrix size (Yu et al., 2024), which does not completely resolve the issue of inverting extremely large matrices. For ultra-large matrices, reducing the dimension by 50% still results in a relatively large matrix, and the difficulty of inversion remains.

Based on the dimensionality reduction method based on DCT (Yu et al., 2024), this paper proposes an incremental discrete cosine transform (IDCT) strategy, and further presents an innovative matrix dimensionality reduction method. Through an iterative IDCT-based reduction process, the matrix dimensions are incrementally reduced, thereby achieving deep dimensionality reduction of large matrices. By introducing this dimensionality reduction method into the linearized Kalman inversion—BLI, a fast stochastic inversion method is formulated and defined as IDCT-BLI (Bayesian linearized inversion based on incremental discrete cosine transform). Finally, the effectiveness of the proposed method is verified by a synthetic and a field data tests. The comparisons in terms of computational efficiency and dimensionality reduction performance between the IDCT-BLI and DCT-BLI proposed by Yu et al. (2024) are also conducted in the tests.

The remaining sections of the paper are introduced briefly as following. First, the basic principles of DCT are introduced. Then, the DCT-based dimensionality reduction method for large sparse matrices is discussed. Next, the basic principles and advantages of the IDCT-based dimensionality reduction method are presented. Following that, the expression for IDCT-BLI is derived. Finally, synthetic and real data tests are conducted to verify the effectiveness of IDCT-BLI.

2. Theory

2.1. Traditional dimension reduction based on DCT

2.1.1. DCT for a matrix

The core theorem of this research is DCT, which can be described as the following equation for 1D cases:

$$H(u) = \sqrt{c_1(u)/N} \sum_{i=0}^{N-1} \epsilon(i) \cos[(i + 0.5)\pi u / N], \quad (1)$$

where ϵ is a signal such as the seismic data, and its sampling number is represented by N . c_1 is a correction parameter that takes a value of 1 when u is zero and 2 when u is non-zero in the case of 1D signals.

Yu et al. (2024) used DCT to reduce the core matrix size in Bayesian seismic inversion, where a core matrix Σ is taken as a two-dimensional signal and transformed by the 2D DCT as follows:

$$H(u, v) = \sqrt{[c_1(u)c_2(v)]/N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \Sigma(i, j) \cos[(i + 0.5)\pi u / N] \cos[(j + 0.5)\pi v / N], \quad (2)$$

where c_1 and c_2 are the correction parameters corresponding to the u and v when dealing with a 2D signal, c_2 should be 1 and 2 when v is zero and non-zero, respectively. Actually, c_1 and c_2 are the same variable but their subscripts just denote different direction of a 2D matrix (or signal). Eq. (2) can be denoted by the following matrix form (Yu et al., 2024),

$$\Sigma' = \mathbf{R}\Sigma\mathbf{R}^T, \quad (3)$$

where the elements in the transform operator $\mathbf{R}$ can be expressed as

$$\mathbf{R}(u, v) = c_1(u) \cos[(j + 0.5)\pi u / N], \quad (4)$$

and $\mathbf{R}^T\mathbf{R}$ equals to a unit matrix, i.e.,

$$\Sigma = \mathbf{R}^T\mathbf{R}\Sigma\mathbf{R}^T\mathbf{R} = \mathbf{R}^T\Sigma'\mathbf{R}, \quad (5)$$

where $\Sigma' = \mathbf{R}\Sigma\mathbf{R}^T$ is the form of Σ after DCT. Actually, after the 2D DCT, most of the significant coefficients in matrix Σ' are concentrated in the top-left corner, i.e., Σ' can represent the effective information in Σ .

2.1.2. Extraction of effective information from a matrix after DCT

Matrix Σ' is a transformed version of Σ that concentrates its significant information, while both matrices share the same dimensions. Next step, the effective information of Σ' need to be extracted. Assuming that the dimension of Σ or Σ' is n_0 , and the expected dimension after reduction is n_1 , a reduction operator $\mathbf{X}$ of n_0 column and n_1 row is defined to extract the valid information from the Σ' as the following manner (Yu et al., 2024):

$$\Sigma_1 = \mathbf{X}\Sigma'\mathbf{X}^T, \quad (6)$$

where $\mathbf{X}$ is actually a sampling operator. In Eq. (6), Σ_1 is also a square matrix. Since $\mathbf{X}$ is a linear operator that determines the matrix dimensions before and after the transformation, it is not a square matrix. Here, n_1/n_0 ($n_0 \geq n_1$) is defined as reduction coefficient (Rcoe). Therefore, the core matrix Σ can be converted into Σ_1 with a much smaller size n_1 once the Rcoe is given.

Overall, the effective part of Σ are concentrated in one corner of Σ' . Therefore, these principal components can be effectively extracted by operator $\mathbf{X}$ with a suitable size, and formulate the final reduced matrix Σ_1 whose size is smaller than Σ .

2.1.3. The expression of the DCT-based dimension reduction

In the above process, an initial matrix Σ can be converted onto a matrix Σ_1 with a smaller size by DCT (Yu et al., 2024), and this process can be expressed as

$$\begin{aligned} \Sigma &= \mathbf{R}^T\mathbf{R}\Sigma\mathbf{R}^T\mathbf{R} = \mathbf{R}^T\Sigma'\mathbf{R} = \mathbf{R}^T\mathbf{X}^{-1}\mathbf{X}\Sigma'\mathbf{X}^T(\mathbf{X}^T)^{-1} \\ \mathbf{R} &= \mathbf{R}^T\mathbf{X}^{-1}\Sigma_1(\mathbf{X}^T)^{-1}\mathbf{R}. \end{aligned} \quad (7)$$

For simplicity, Eq. (7) is rewritten in the following form:

$$\Sigma = \Gamma_1\Sigma_1\Gamma_1^T, \quad (8)$$

where the operator for dimensionality reduction is defined as

$$\Gamma_1 = \mathbf{R}^T\mathbf{X}^{-1}. \quad (9)$$

Based on the above derivation, the inverse calculation of Σ can be written as

$$\Sigma^{-1} = (\Gamma_1\Sigma_1\Gamma_1^T)^{-1} = (\Gamma_1^T)^{-1}(\Sigma_1)^{-1}\Gamma_1^{-1}. \quad (10)$$

The inverse of Σ is converted into the inversions of the above three matrices. It seems that the matrix inverse after the above conversion is more complicated than inverting Σ directly. However, Γ (or Γ^T) is constant for a single two- or three-dimensional seismic inversion test. Thus, only once inversion of Γ (or Γ^T) is enough for all the traces.

The aforementioned approach is the dimensionality reduction method proposed by Yu et al. (2024). This method enables the generation of reduced representations by assigning a specific value to the Rcoe parameter. As illustrated in Fig. 1(a), the method operates by concentrating the effective information of the original matrix via Eq. (3), and then extracting this information using Eq. (6), thereby achieving dimensionality reduction. In most cases, this approach can reduce the dimensionality of a matrix to approximately 40%–50%. Building upon this framework, the present study introduces an incremental reduction method, which applies iterative DCT-based reduction to the matrix, enabling a more thorough dimensionality reduction than the conventional DCT approach.

2.2. Incremental reduction based on DCT

The key to the IDCT is achieving a deeper reduction degree incrementally, by conducting DCT iteratively, namely conducting several rounds (more than one) of DCT. However, it is essential to reasonably select the relevant parameters during the iterative process to ensure that the progressive dimensionality reduction does not compromise the integrity of the effective matrix information. In the following subsection, two cases including two and l ($l > 2$) rounds of dimension reductions are listed herein to clarify the basic theorem of IDCT.

2.2.1. Derivation of two reduction iterations

First, a simple case with two reduction iterations is used to introduce the IDCT initially. This case, in another word, is first converting Σ to Σ_1 using Eq. (8), and further transform Σ_1 to Σ_2 with a size of n_2 ($n_0 > n_1 > n_2$). Notably, n_1/n_0 and n_2/n_1 herein are Rcoe values for the first and the second reduction iteration. Generally, the Rcoe values in different reduction rounds can be a same value for convenience. The above second reduction process can be expressed as

$$\Sigma_1 = \Gamma_2\Sigma_2\Gamma_2^T, \quad (11)$$

where the dimension reduction operator is

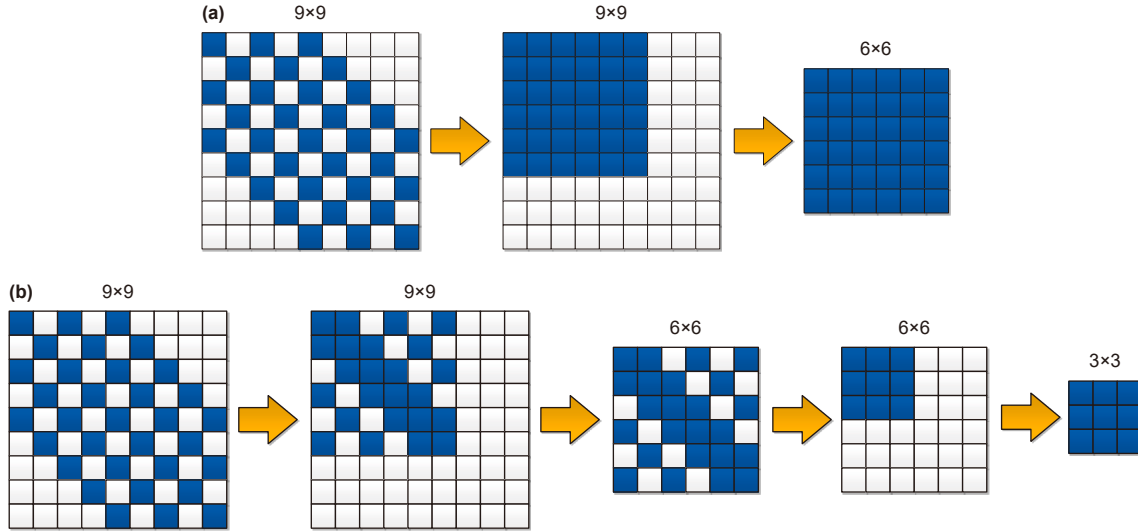


Fig. 1. The sketches of (a) DCT, and (b) IDCT with two rounds of reductions expressed in Eq. (13).

$$\Gamma_2 = \mathbf{R}_2^T \mathbf{X}_2^{-1}. \quad (12)$$

where the $\mathbf{R}_2$ and $\mathbf{X}_2$ are the DCT transform operator and sampling operator corresponding to the matrix size n_1 , which are introduced in Eqs. (3) and (6). Then, the relationship between Σ and Σ_2 can be written as

$$\Sigma = \Gamma_1 \Gamma_2 \Sigma_2 \Gamma_1^T \Gamma_2^T, \quad (13)$$

The inverse of Σ can be also expressed as

$$\Sigma^{-1} = (\Gamma_1 \Gamma_2 \Sigma_2 \Gamma_1^T \Gamma_2^T)^{-1} = (\Gamma_2^T)^{-1} (\Gamma_1^T)^{-1} \Sigma_2^{-1} \Gamma_1^{-1} \Gamma_2^{-1}. \quad (14)$$

Therefore, the inverse of Σ is transformed into the inverse of a much smaller matrix Σ_2 and a set of reduction operators, which means the inverse of Σ has a great possibility to be simplified.

This simple IDCT approach, which involves only two stages of dimensionality reduction, is illustrated in Fig. 1(b). The blue and white squares can respectively represent the effective and redundant elements within the matrix. First, the DCT method is applied to concentrate the effective information from the original 9-dimensional matrix. Then, the effective information (indicated by blue nodes) is extracted which results in a 6-dimensional matrix. However, this intermediate matrix still contains some redundant or non-informative components (represented by white nodes in Fig. 1). Therefore, a second DCT-based reduction is applied to further compress the data, ultimately yielding a three-dimensional matrix.

2.2.2. Derivation of l reduction iterations

According to the above derivation, Σ can be transformed into an extremely small matrix after l rounds of reductions, which can be expressed as following according to Eq. (13),

$$\Sigma = \Gamma_1 \Gamma_2 \cdots \Gamma_l \Sigma_l \Gamma_l^T \cdots \Gamma_2^T \Gamma_1^T = \Lambda \Sigma_l \Lambda^T, \quad (15)$$

for the convenience of the following inversion derivation, $\Gamma_1 \Gamma_2 \cdots \Gamma_l$ is defined as the final reduction operator Λ that responsible for incremental dimension reduction. The corresponding inverse of Σ after l times of reductions can be written as

$$\begin{aligned} \Sigma^{-1} &= (\Gamma_1 \Gamma_2 \cdots \Gamma_l \Sigma_l \Gamma_l^T \cdots \Gamma_2^T \Gamma_1^T)^{-1} = (\Gamma_l^T)^{-1} \cdots (\Gamma_2^T)^{-1} \\ &\quad (\Gamma_1^T)^{-1} \Sigma_l^{-1} (\Sigma_l)^{-1} \Gamma_1^{-1} \Gamma_2^{-1} \cdots \Gamma_l^{-1} \\ &= (\Lambda^T)^{-1} \Sigma_l^{-1} \Lambda^{-1}. \end{aligned} \quad (16)$$

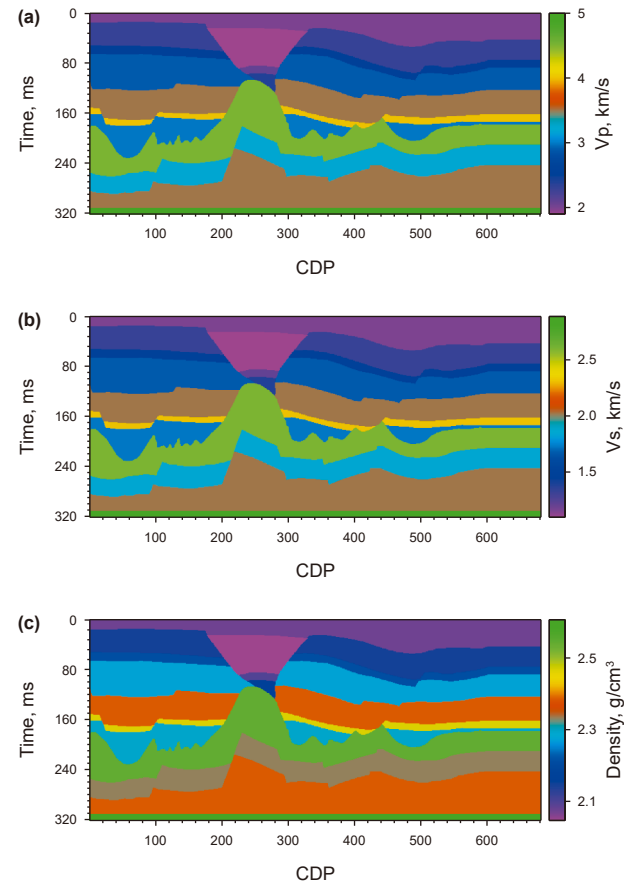


Fig. 2. Real models of (a) P-wave velocity (V_p), (b) S-wave velocity (V_s), and (c) density.

2.2.3. Parameters selection of IDCT

The above derivation proves that a large matrix can be transformed into a small matrix by IDCT, which is essentially conducting DCT iteratively on the large matrix. Although the mathematical theorem of IDCT is demonstrated, two critical issues must be clarified to keep the effectiveness of IDCT in geophysical applications.

- The Rcoe value should not be too large in IDCT, as it is a crucial factor in IDCT. In DCT, the reduction is performed only once, and Rcoe is typically set to 40%–60% (depending on the data type). However, DCT is relatively rough because such a small Rcoe may result in excessive loss of valid information. In each iteration of IDCT, a larger Rcoe (e.g., 70%–80%) is used to mitigate the loss of valid information. Consequently, the reduction per iteration in IDCT is smaller than that in DCT, but the overall reduction effect of IDCT is significantly more pronounced than that of DCT.
- Although IDCT can effectively reduce the size of a matrix, it also generates multiple reduction operators that need to be inverted. This may appear to introduce additional computational burden. However, in practical two-dimensional and three-dimensional applications of seismic inversion, the inversion of these reduction operators only needs to be performed for a single trace, and the same reduction operators can be applied to other traces.

Therefore, a 1D inversion test near a drilled well must be conducted before performing two-dimensional and three-

dimensional seismic inversion with IDCT for two purposes: First, the crucial sampling parameters including Rcoe and iteration count are determined; Second, the reduction operators like Γ_l^{-1} in Eq. (16) need to be pre-calculated and stored before the IDCT-based inversion application in large-scale datasets. Consequently, for all traces except the one used in the 1D test, the only matrix that needs to be inverted is actually the one after IDCT, e.g., the Σ_l in Eq. (16).

2.3. Bayesian linearized inversion with IDCT

BLI is a seismic inversion technique that integrates Bayesian inference with linearization methods. The fundamental principle is to update the posterior distribution of model parameters based on Bayesian theorem, given prior information and observed data. BLI is based on the Bayesian theory,

$$p(\mathbf{m}|\mathbf{d}) = p(\mathbf{d}|\mathbf{m})p(\mathbf{m}) / p(\mathbf{d}) \quad (17)$$

where $p(\mathbf{m})$ represents the prior probability distribution of the parameter $\mathbf{m}$, $p(\mathbf{d}|\mathbf{m})$ is the likelihood function, $p(\mathbf{m}|\mathbf{d})$ is the posterior probability distribution. The detailed information regarding calculating $p(\mathbf{m}|\mathbf{d})$ by BLI can be found in Appendix A.

Apparently, according to Appendix A, the inverse of a core matrix $\mathbf{C}_d$ is significant for both the estimation of the posterior means in Eq. (A-6) and covariance matrix in Eq. (A-6) of BLI. The inversion of large matrices is highly challenging, which significantly limits the practicality of BLI. Herein, the IDCT-based large matrix dimensionality reduction method mentioned above is

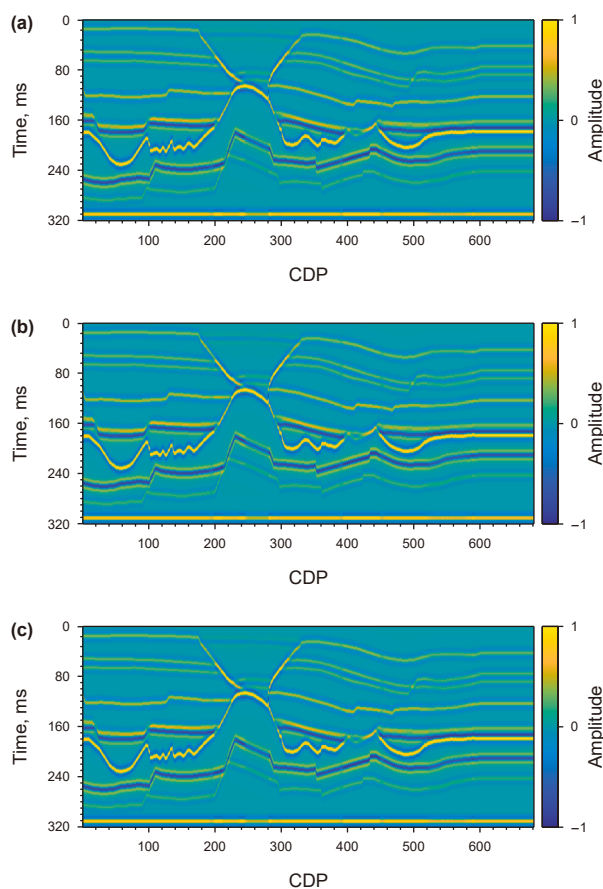


Fig. 3. Synthetic seismic profiles of partial angle stacked data with the incident angles of (a) 5°, (b) 15°, and (c) 25°.

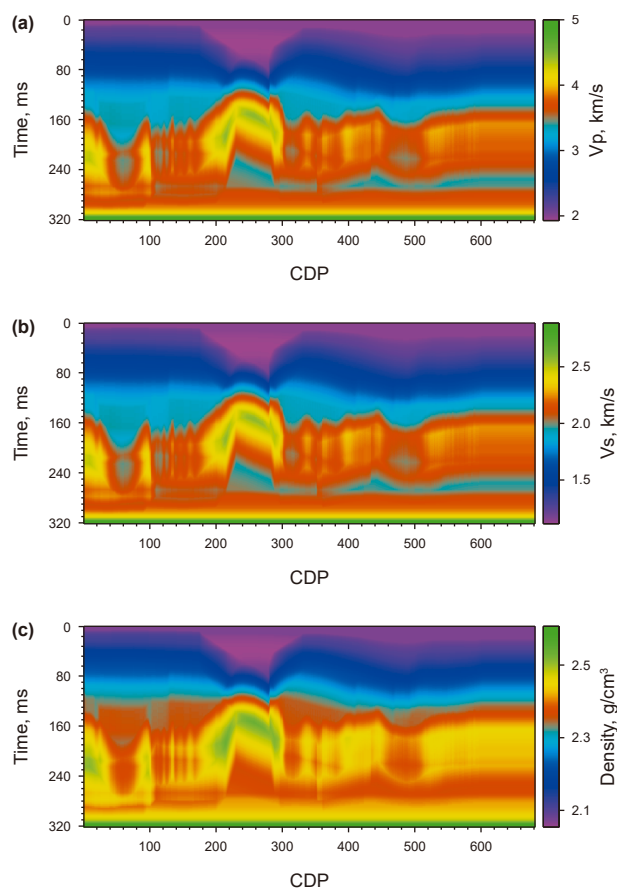


Fig. 4. The prior models of the (a) P-wave velocity, (b) S-wave velocity, and (c) density.

applied to solve the problem of traditional BLI. In the proposed method, the dimensionality reduction on the core matrix of BLI is conducted using the incremental dimensionality reduction operator Λ . As the manner shown in Eq. (15), the core matrix $\mathbf{C}_d$ (the subscript d represents the seismic data) of BLI, which is expressed by Eqs. (A-6) and (A-7) in Appendix A, is converted into the following form:

$$\mathbf{C}_d = \Lambda \mathbf{C}_d' \Lambda^T = \Gamma_1 \Gamma_2 \cdots \Gamma_l \mathbf{C}_d' \Gamma_l^T \cdots \Gamma_2^T \Gamma_1^T, \quad (18)$$

where $\mathbf{C}_d'$ is small matrix that has a much smaller size than $\mathbf{C}_d$ according to the IDCT theory, then inverse of $\mathbf{C}_d$ can be further expressed as

$$\mathbf{C}_d^{-1} = (\Gamma_1 \Gamma_2 \cdots \Gamma_l \mathbf{C}_d' \Gamma_l^T \cdots \Gamma_2^T \Gamma_1^T)^{-1} = (\Lambda^T)^{-1} (\mathbf{C}_d')^{-1} \Lambda^{-1}. \quad (19)$$

Therefore, using the IDCT strategy, the inverted posterior mean for BLI in Eq. (A-6) can be converted into the following expression:

$$\begin{aligned} \Psi_{m|d} &= \Psi_m + \mathbf{C}_m \mathbf{G}^T (\Gamma_1 \Gamma_2 \cdots \Gamma_l \mathbf{C}_d' \Gamma_l^T \cdots \Gamma_2^T \Gamma_1^T)^{-1} (\mathbf{d} - \mathbf{G} \Psi_m) \\ &= \Psi_m + \mathbf{C}_m \mathbf{G}^T (\Lambda^T)^{-1} (\mathbf{C}_d')^{-1} \Lambda^{-1} (\mathbf{d} - \mathbf{G} \Psi_m). \end{aligned} \quad (20)$$

Correspondingly, the inverted posterior covariance should be

$$\begin{aligned} \mathbf{C}_{m|d} &= \mathbf{C}_m - \mathbf{C}_m \mathbf{G}^T (\Gamma_1 \Gamma_2 \cdots \Gamma_l \mathbf{C}_d' \Gamma_l^T \cdots \Gamma_2^T \Gamma_1^T)^{-1} \mathbf{C}_m \\ &= \mathbf{C}_m - \mathbf{C}_m \mathbf{G}^T (\Lambda^T)^{-1} (\mathbf{C}_d')^{-1} \Lambda^{-1} \mathbf{C}_m. \end{aligned} \quad (21)$$

The physical meaning of the above parameters can be found in Appendix A. The subscript "m" denotes the model parameters. Eqs. (20) and (21) formulate the final expression of the IDCT-BLI, which only needs to solve a much smaller core matrix than the DCT-BLI. In the IDCT-BLI method, Λ and Λ^T are only related to the size of the seismic data, therefore, they are the same for all the traces in a two-dimensional and three-dimensional zone. Hence, only one inverse calculation for Λ and Λ^T is necessary in a two-dimensional and three-dimensional IDCT-BLI case. Therefore, in the whole IDCT-BLI process, the inverse of the small matrix $\mathbf{C}_d'$ is the core issue that need to be taken into consideration, which contributes to improving the efficiency of the traditional BLI.

3. Numerical examples

A synthetic test and a real data test of the above IDCT-BLI method is conducted to demonstrate the effectiveness of the IDCT-based dimensionality reduction method. In the two tests, the comparison between the IDCT-based dimensionality reduction method with the DCT-based method proposed by Yu et al. (2024) are also conducted to validate the advantages of the proposed IDCT-based dimensionality reduction approach.

3.1. Synthetic data test

A two-dimensional salt dome model is used herein to verify the performance of the proposed incremental reduction strategy in geophysical inversion. The two-dimensional real models of P- and S-wave velocities and density is displayed in Fig. 2, and these models includes 320 nodes in the vertical (time) direction and 681 nodes in the horizon (CDP) direction.

The noise-free synthetic prestack seismic data is obtained by a 30 Hz Ricker wavelet convolving with the PP-wave reflection coefficient from the Aki-Richards approximation formula. The

incident angles of the seismic profiles shown in Fig. 3(a)–(c) are 5°, 15° and 25°. Universally, prior models are crucial in seismic inversion, especially for the different BLI methods. In this theoretical model test, the prior models are constructed by smoothing the real models displayed in Fig. 2. Fig. 4 show the prior models of

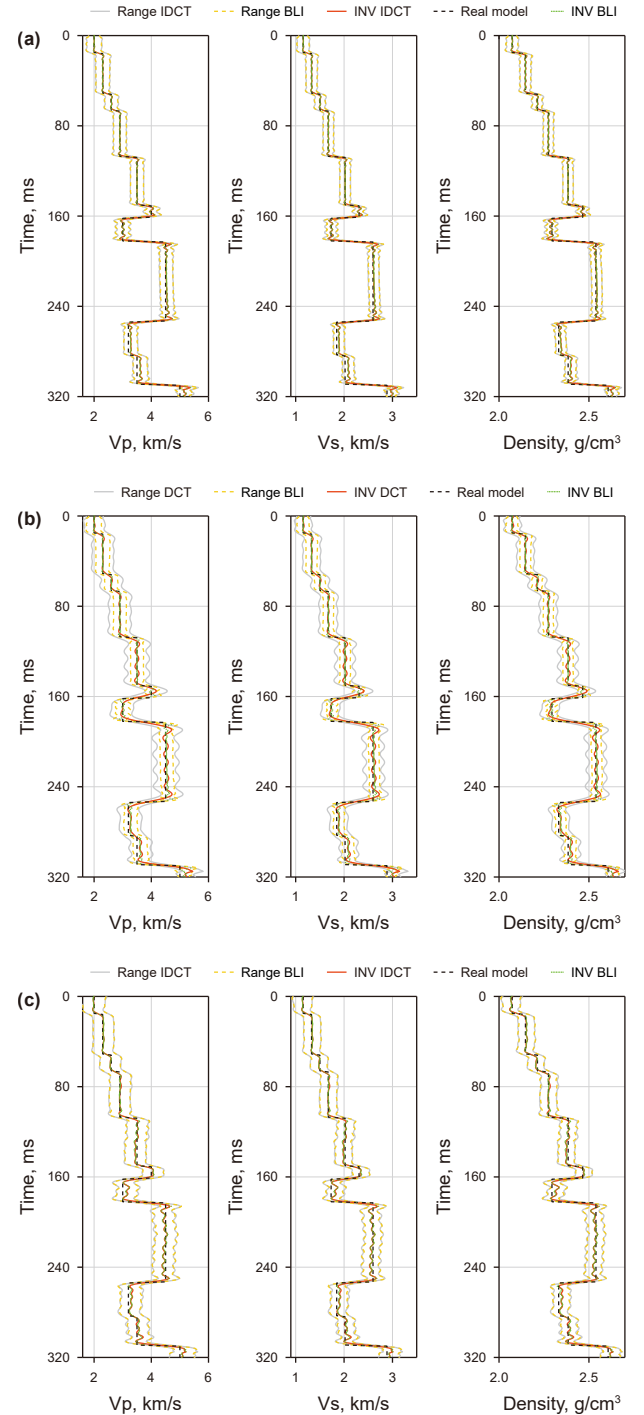


Fig. 5. The inversion accuracy comparison, (a) is the comparison between BLI and IDCT-BLI, (b) is the comparison between BLI and DCT-BLI, and (c) is the comparison between BLI and IDCT-BLI using the noisy seismic data. The black lines note the real models; the green lines are the inversion results of BLI; the red lines represent the posterior mean of the inversion results estimated by IDCT-BLI in (a) and DCT-BLI in (b); the orange and gray lines note the 95% confidence interval of the BLI and BLI with dimension reduction.

the P- and S-wave velocities, and density. The following tests of the traditional BLI, DCT-BLI, and IDCT-BLI use the same prior model.

The traditional BLI, DCT-BLI, and IDCT-BLI are conducted simultaneously for comparison. The first trace of the real model is used as an example, and the corresponding 1D inversion test results are displayed in Fig. 5(a) to prove the reduction effect of the IDCT. In Fig. 5(a), the black dashed lines present the real models of the P- and S-wave velocities, and density. The green dashed and

red lines denote the inversion results of BLI (without any dimension reduction) and IDCT-BLI. The orange dashed gray lines represent the 95% confidence interval of the 1D inversion results. The inversion results from IDCT-BLI keep good accordance with those of BLI. In addition, the boundaries of the confidence intervals estimated by BLI and IDCT-BLI are also the same with each other.

The core matrix size in the 1D BLI test is 1280, while the size in IDCT-BLI is 180, namely the percent of the dimension reduction of

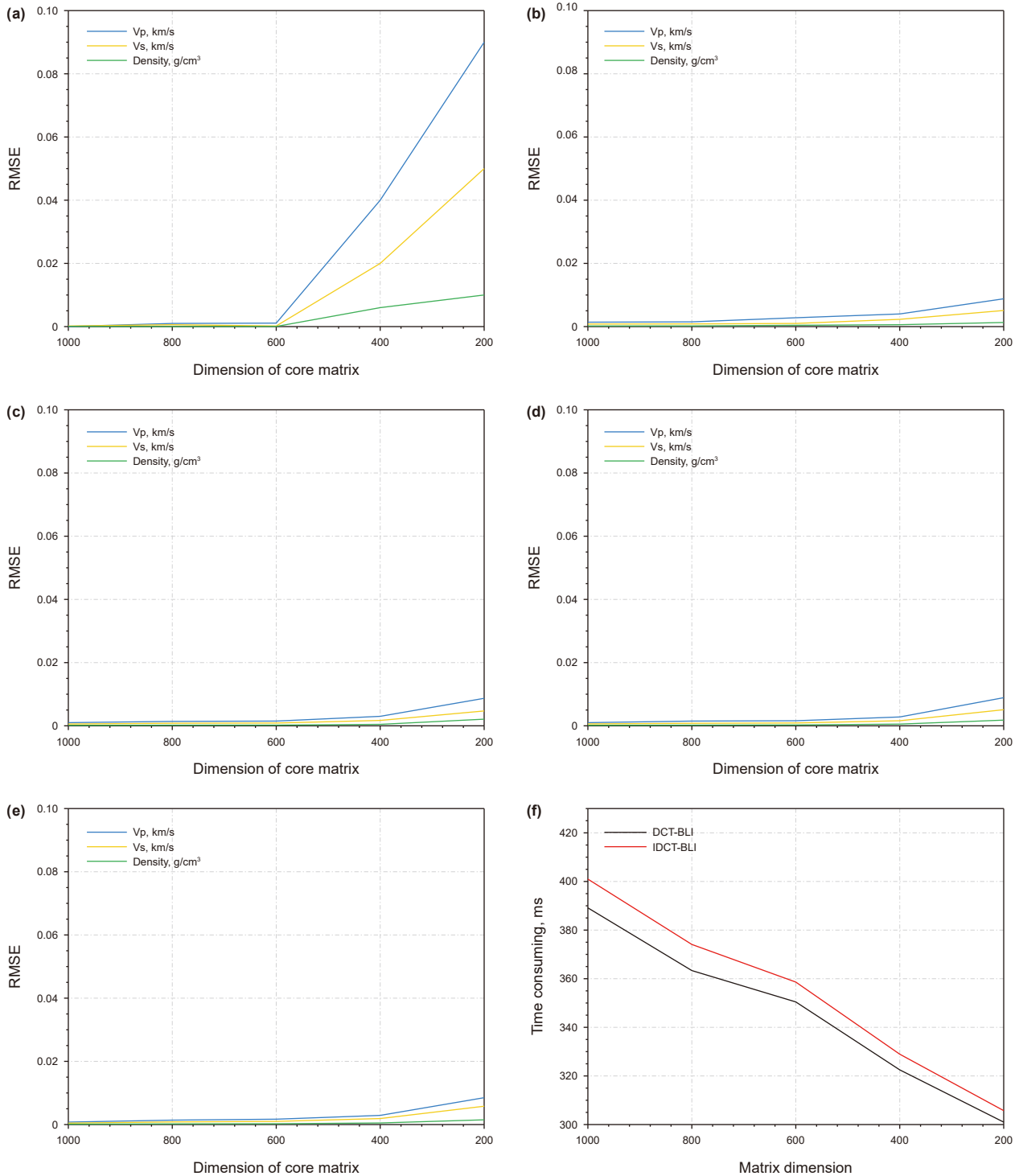


Fig. 6. The RMSEs of the three inverted model parameters from IDCT-BLI with different reduction iterations. (a)–(e) corresponds to the reduction iterations of 1–5. (f) Presents the comparison of the time consuming for IDCT-BLI and DCT-BLI solving different size of kernel matrices.

the IDCT strategy is more nearly 86% in this case. For comparison, the 1D DCT-BLI is also conducted using the same model, and the final core matrix dimension is also set as 180. Fig. 5(b) displays the corresponding inversion result, where the difference between the inversion results from BLI and DCT-BLI is more evident than that in Fig. 5(a), whatever for the inversion results and the confidence interval. Therefore, the reduction limit of the DCT-based method is over 180 in this case.

To further evaluate the noise robustness of the dimensionality reduction method, random noise is added to the one-dimensional seismic trace used in this test, resulting in a signal-to-noise ratio (SNR) of 7 dB. Fig. 5(c) presents a comparison between the traditional BLI and the IDCT-BLI methods under noisy seismic data. Compared to the noise-free inversion results in Fig. 5(a), both methods exhibit reduced inversion accuracy due to the presence of noise, and the confidence intervals become wider, indicating increased uncertainty in the inversion results. However, the inversion results and the corresponding confidence interval boundaries for both methods are nearly identical, suggesting that

Table 1
The Rcoe values corresponding to different target dimension in the five times of inversion tests.

Dimension iteration	1000	800	600	400	200
1	78%	62%	47%	31%	16%
2	89%	79%	69%	56%	40%
3	92%	86%	78%	68%	54%
4	94%	89%	83%	75%	63%
5	95%	91%	86%	79%	69%

the IDCT-based dimensionality reduction is unaffected by the added noise. Regardless of whether noise is present, the inversion results of pre-stack three-parameter estimation before and after IDCT-based dimensionality reduction remain consistent.

A set of inversion cases are conducted to discuss the reduction limit of DCT-based and IDCT-based reduction methods, and to prove the advantages of IDCT compared with DCT. In these tests, the size of the core matrix is 1280, and 5 reduced dimension are set as 1000, 800, 600, 400, and 200. As is introduced in the Theory section, IDCT-based reduction is actually an iterative DCT-based reduction strategy. IDCT equals DCT when the iteration round is 1. Herein, the iterations of IDCT-BLI vary from 1 to 5 in this synthetic data test. The root-mean-square errors (RMSEs) between inversion results of BLI and the IDCT-BLIs with different reduction rounds and different expected reduced dimensions are used as the evaluation criterion.

Fig. 6(a) displays the analysis result of IDCT-BLI with only 1 round of reduction (namely the DCT-BLI), and Fig. 6(b)–(e) show the analysis results of IDCT-BLI with the reduction rounds varying from 2 to 5. of dimension reduction. The blue, yellow, and green lines represent the RMSE curves for P- and S-wave velocities, and density.

After Eq. (7), the Rcoe is defined to characterize the dimensionality reduction ratio in each round of the dimension reduction. Table 1 presents the Rcoe values corresponding to the five inversion tests. For example, the value of 78% in the second row and second column of Table 1 indicates that when Rcoe is 78%, a single-step dimensionality reduction using the DCT-BLI method reduces the core matrix from 1280 dimensions to 1000 dimensions.

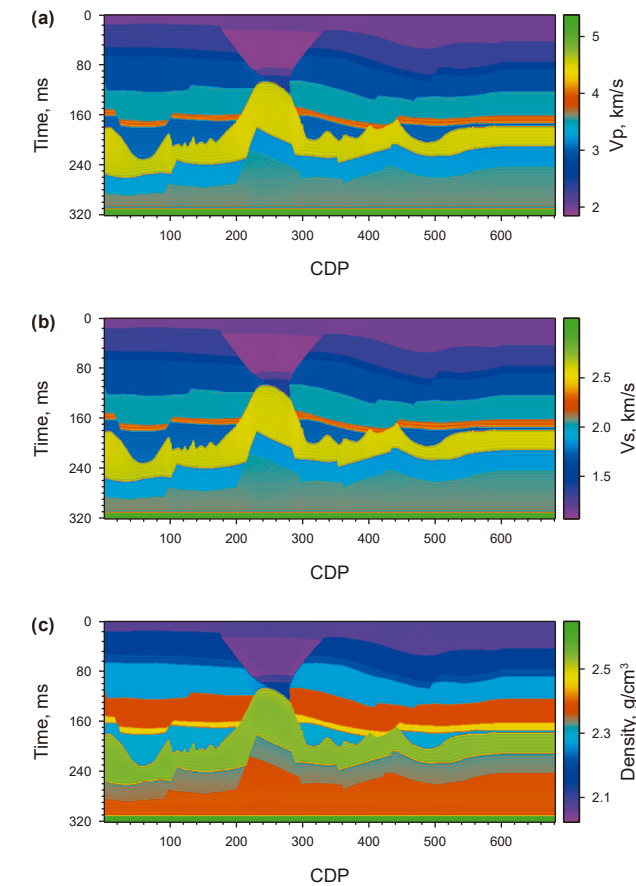


Fig. 7. The inversion results of (a) P-wave velocity, (b) S-wave velocity, and (c) density estimated by BLI.

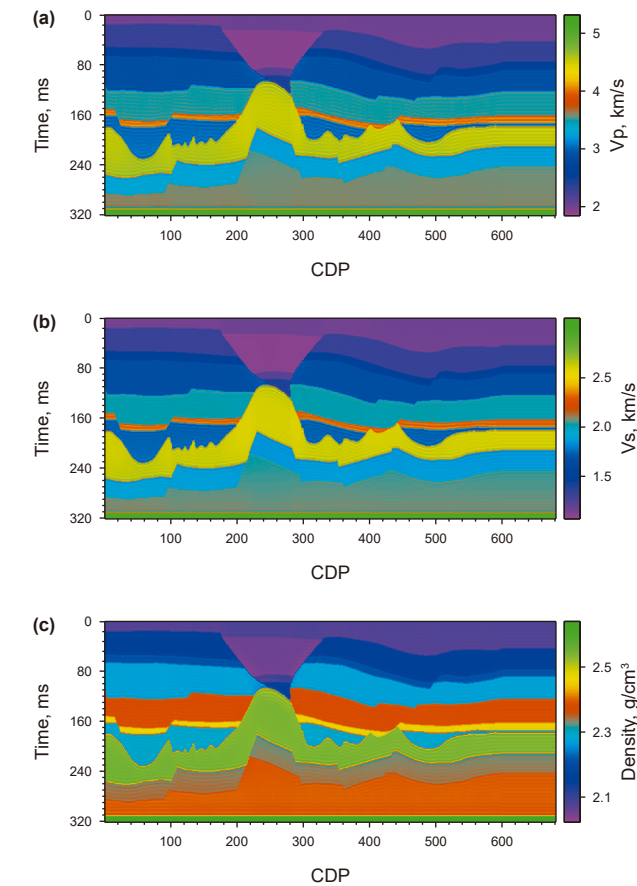


Fig. 8. The inversion results of (a) P-wave velocity, (b) S-wave velocity, and (c) density estimated by DCT-BLI.

Similarly, the value of 40% in the third row and sixth column means that when Rcoe is 40%, applying a two-step dimensionality reduction with the IDCT-BLI method can reduce the core matrix from 1280 dimensions to 200 dimensions.

It can be observed from Fig. 6 that when the inversion involves only a single-step dimensionality reduction, namely IDCT-BLI and DCT-BLI are equivalent, the RMSE of the inverted P-wave velocity, S-wave velocity, and density starts to increase significantly when the matrix dimension is reduced to 600, as shown in Fig. 6(a). However, when the number of dimensionality reduction steps in a single inversion exceeds two, as illustrated in Fig. 6(b)–(e), the RMSE of the three inversion parameters only shows a slight increase when the matrix dimension is reduced below 400.

When the dimensionality reduction exceeds one step, even if the matrix dimension is reduced to 200, the RMSE of the three model parameters does not increase obviously. When the matrix dimension is reduced to 200, the relative error of the three inversion parameters remains below 1% and has no visually noticeable impact on the inversion results for all Fig. 6(a)–(e). The specific effects can be referenced in Fig. 5(a). Furthermore, as shown in Fig. 6(b)–(e), although the number of dimensionality reduction steps increases from 2 to 5, the inversion accuracy does not improve too much after dimensionality reduction. Taking the P-wave velocity as an example, when the core matrix dimension is reduced to 200, the RMSE values of P-wave velocity in Fig. 6(b)–(e) are 0.0089, 0.0087, 0.0086, and 0.0084 km/s, respectively. Therefore, the optimal number of reduction iterations is two for the salt dome model in this test. A single-step reduction affects the

inversion accuracy after dimensionality reduction, while excessive reduction rounds can increase the calculation burden. Furthermore, an excessive number of dimensionality reduction steps may introduce computational instability due to the accumulation of iterative operations. Based on the results from the present model experiments, it is generally advisable to limit the number of reduction iterations to two or three for a balance between efficiency and stability.

In addition, a comparison of the computational time required by IDCT-BLI and DCT-BLI in the aforementioned reduction tests is presented in Fig. 6(f). Dimensionality reduction tests with target dimensions of 1000, 800, 600, 400, and 200 were performed on the 2D model (original matrix dimension: 1280). For each reduction level, inversion is carried out using both IDCT-BLI and DCT-BLI across all 681 seismic traces in the model, and the corresponding computational time is recorded. In the case of IDCT-BLI, the reduction is conducted using a two-step iterative approach, consistent with the tests in Fig. 6(b). The red and black curves in Fig. 6(f) represent the runtime of DCT-BLI and IDCT-BLI, respectively. Notably, the computational time for traditional BLI on the same 2D model is approximately 430 s. As shown in Fig. 6(f), a clear trend is observed: as the matrix dimensionality decreases, the inversion runtime also decreases. Compared to the 430 s required by traditional BLI, both DCT-BLI and IDCT-BLI achieve more than 30% runtime reduction at the maximum reduction level (dimension = 200), indicating the significant efficiency improvement brought by dimensionality reduction. Furthermore, although the runtime of IDCT-BLI is consistently slightly longer than that of

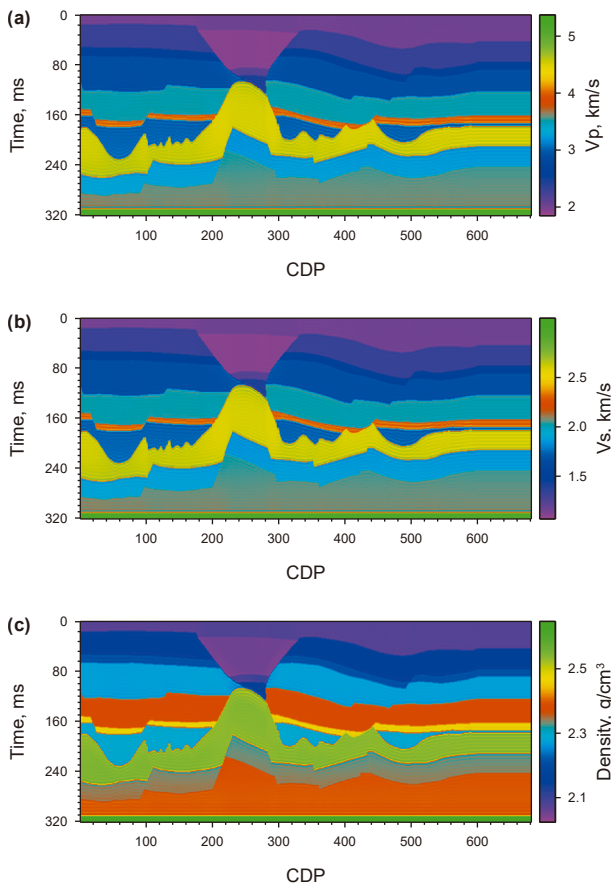


Fig. 9. The inversion results of (a) P-wave velocity, (b) S-wave velocity, and (c) density estimated by IDCT-BLI.

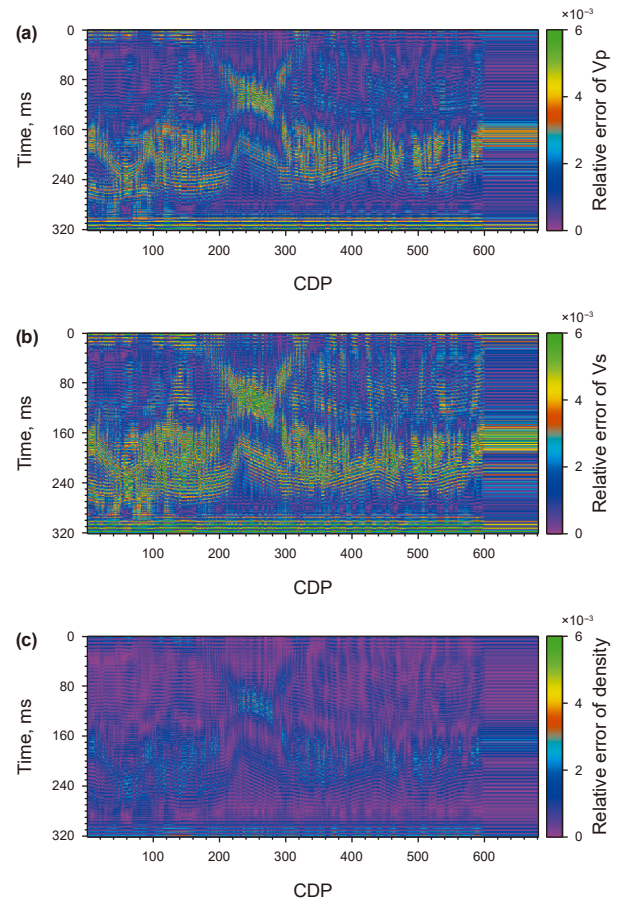


Fig. 10. The inversion error of (a) P-wave velocity, (b) S-wave velocity, and (c) density estimated by DCT-BLI.

DCT-BLI due to the additional matrix multiplications involved in the iterative reduction process, IDCT-BLI maintains high inversion accuracy even at a dimensionality of 200. In contrast, DCT-BLI fails to ensure inversion accuracy when the matrix dimension drops below 600. Therefore, despite a slight increase in computational cost, IDCT-BLI offers a more efficient and robust solution for inversion tasks, outperforming DCT-BLI in terms of both accuracy and practical applicability.

The above one-dimensional test demonstrates that IDCT-BLI can achieve inversion results with the same accuracy as traditional BLI and DCT-BLI, without compromising inversion precision. Therefore, the advantage of the IDCT-based reduction compared with the DCT-based method is also proved. A two-dimensional test will be conducted to further evaluate the effectiveness of IDCT-BLI. Based on the pre-stack seismic angle gathers and initial model shown in Figs. 3 and 4, the inversion results for the P- and S-wave velocities, and density are obtained.

Fig. 7 presents the inversion results obtained using the traditional BLI, while Figs. 8 and 9 show the three-parameter inversion results derived from DCT-BLI and IDCT-BLI, respectively. A direct comparison reveals that DCT-BLI and IDCT-BLI achieve the same inversion performance as traditional BLI, with the three-parameter inversion results closely matching the true model shown in Fig. 2. To further quantify the inversion performance, Fig. 10 illustrates the relative errors between the three-parameter inversion results of BLI and DCT-BLI, whereas Fig. 11 displays the relative errors between the three-parameter inversion results of BLI and IDCT-BLI. The relative

error equals the absolute difference between the inversion result of DCT-BLI or IDCT-BLI and that of BLI, divided by the BLI inversion result.

As observed from the above figures, both IDCT-BLI and DCT-BLI exhibit similar inversion error compared to traditional BLI. However, in Figs. 10 and 11, the relative errors of the P- and S-wave velocities, and density are mostly within 1%, indicating that both DCT-BLI and IDCT-BLI achieve high inversion accuracy. Notably, even when the core matrix dimension is reduced to 180, IDCT-BLI maintains a similar inversion accuracy to DCT-BLI, further demonstrating the effectiveness of the IDCT-based matrix dimensionality reduction method.

3.2. Field data test

The effectiveness of the proposed inversion method is evaluated using a field dataset from an oilfield in China. This dataset includes three partial-angle stacked two-dimensional seismic sections, corresponding to center incident angles of 5°, 15°, and 25° displayed in Fig. 12(a)–(c). The two-dimensional inversion grid consists of 900 cells in the horizontal direction and 400 cells in the vertical direction, corresponding to 900 seismic traces and a 400 ms time window for each trace. The grid cell sizes are 20 m horizontally and 1 ms vertically.

As mentioned in the theoretical model test, initial models are significant to BLI, and a reasonable initial model is necessary to verify the effectiveness of the IDCT-BLI. The initial model used in the real data test is constructed based on well-log data constrained by seismic horizons. Actually, a commercial software is used

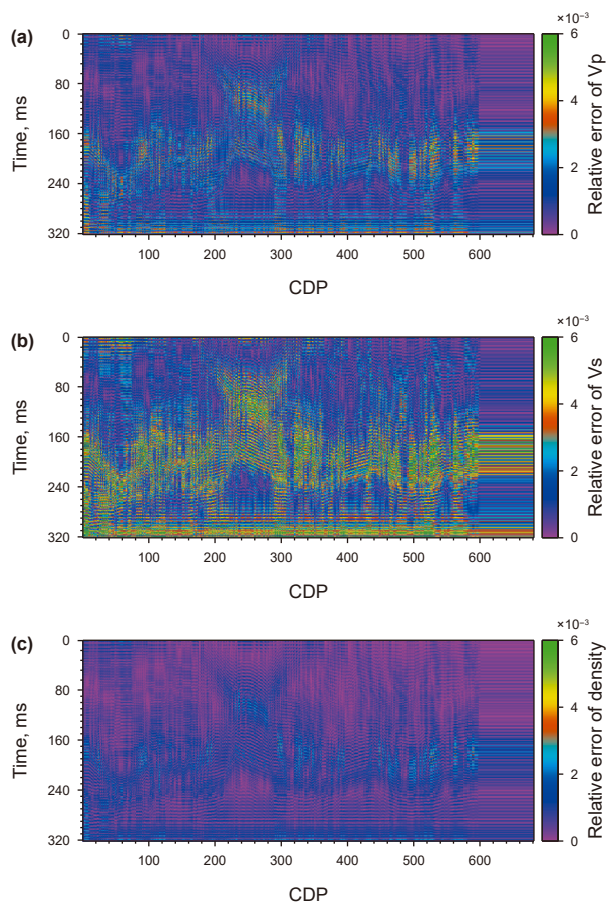


Fig. 11. The inversion error of (a) P-wave velocity, (b) S-wave velocity, and (c) density estimated by IDCT-BLI.

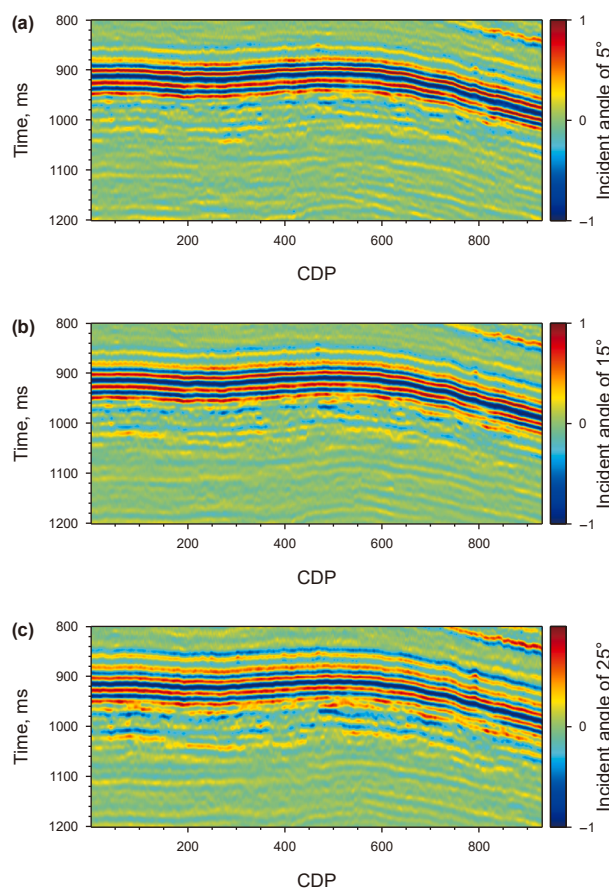


Fig. 12. Field prestack angle gathers with the incident angles of (a) 5°, (b) 15°, and (c) 25°.

herein to accomplish this work. The initial models for the P- and S-wave velocities, and density are shown in Fig. 13(a)–(c). Based on the prestack seismic data and the initial model described above, the inversion tests are conducted using BLI, DCT-BLI, and IDCT-BLI, respectively.

For a fair comparison of inversion performance, the three inversion tests are conducted using the same parameters, including the variogram, initial model, and noise variance. However, BLI does not employ the dimensionality reduction method proposed in this study, and the core matrix dimension remained at 1200. DCT-BLI utilizes the DCT-based dimensionality reduction method proposed by Yu et al. (2024). Through testing, the core matrix dimension is reduced from 1200 to 430, as further reduction would compromise inversion accuracy. IDCT-BLI that based on the dimensionality reduction method proposed in this study achieves a minimum core matrix dimension of 200.

Fig. 14(a)–(c) show the inversion results for the P- and S-wave velocities, and density corresponding to the traditional BLI. Fig. 15(a)–(c) present the inversion results for the three prestack elastic parameters obtained using DCT-BLI. Fig. 16(a)–(c) display the prestack three-parameter inversion results obtained using IDCT method. The inversion results from the above three methods exhibit satisfactory resolution and lateral continuity, effectively capturing the spatial distribution characteristics of the reservoir at 900 ms. Visually, the three prestack inversion results for the three elastic parameters shown in Figs. 14–16 are highly similar, indicating that the dimensionality reduction of the matrix has not compromised the inversion accuracy.

To quantitatively evaluate the effectiveness of the IDCT dimensionality reduction method, the relative errors between the inversion results of DCT-BLI, IDCT-BLI, and traditional BLI were computed, using the inversion results of traditional BLI as the reference. Fig. 17 presents the relative errors of the three parameters between DCT-BLI and traditional BLI, while Fig. 18 shows the relative errors of the three parameters between IDCT-BLI and traditional BLI. A comparison reveals that both DCT-BLI and IDCT-BLI produce inversion results for P-wave velocity, S-wave velocity, and density with relative errors within 1% compared to traditional BLI. This indicates that the inversion accuracy of DCT-BLI and IDCT-BLI is equivalent to that of traditional BLI, and the dimensionality reduction of the matrix does not affect the inversion accuracy.

To further quantitatively analyze the errors between the inversion results obtained by DCT-BLI and IDCT-BLI and those obtained by traditional BLI, the probability distribution curves corresponding to the errors shown in Figs. 17 and 18 are plotted and presented in Fig. 19. In Fig. 19(a), the error distribution curves of P-wave velocity are shown. The blue curve represents the error between the IDCT-BLI and traditional BLI results, while the red dashed curve represents the error between the DCT-BLI and traditional BLI results. A comparison reveals that, except for the region indicated by the black arrow where a slight discrepancy exists, the two curves almost completely overlap across the rest of the domain. Fig. 19(b) and (c) present the error distributions for S-wave velocity and density, respectively. These two figures exhibit similar patterns to Fig. 19(a), indicating that the inversion results from DCT-BLI and IDCT-BLI are nearly identical, and their

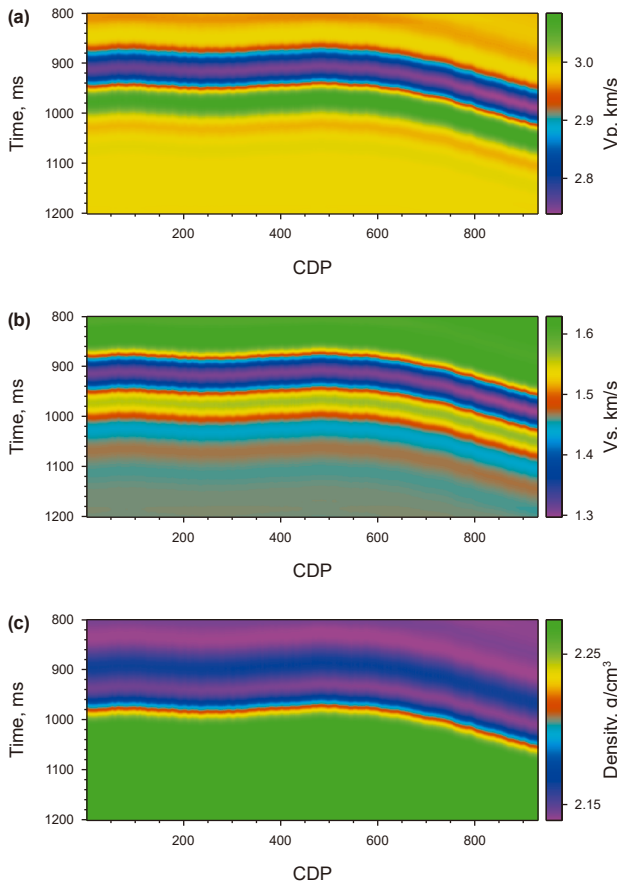


Fig. 13. Prior models of the (a) P-wave velocity, (b) S-wave velocity, and (c) density.

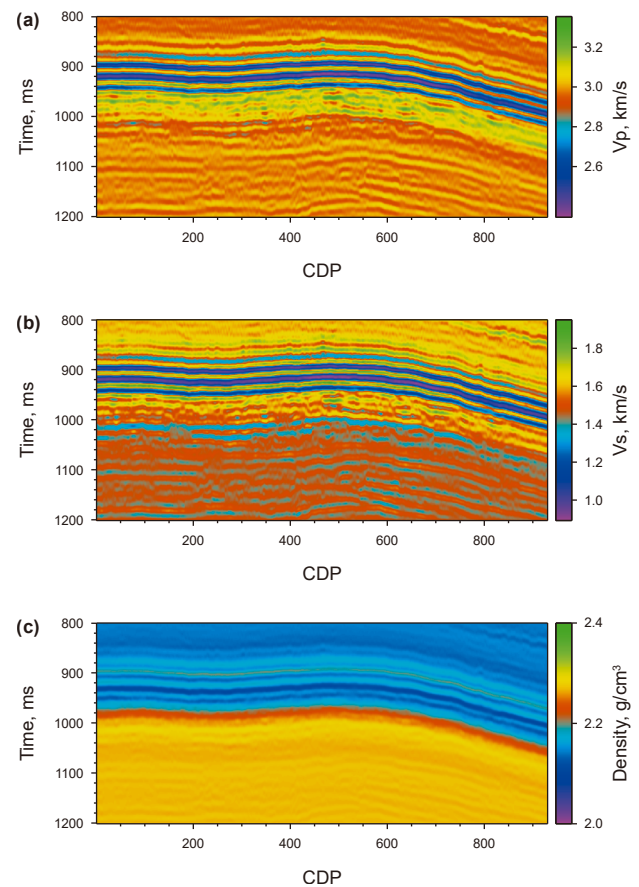


Fig. 14. BLI results of (a) P-wave velocity, (b) S-wave velocity, and (c) density.

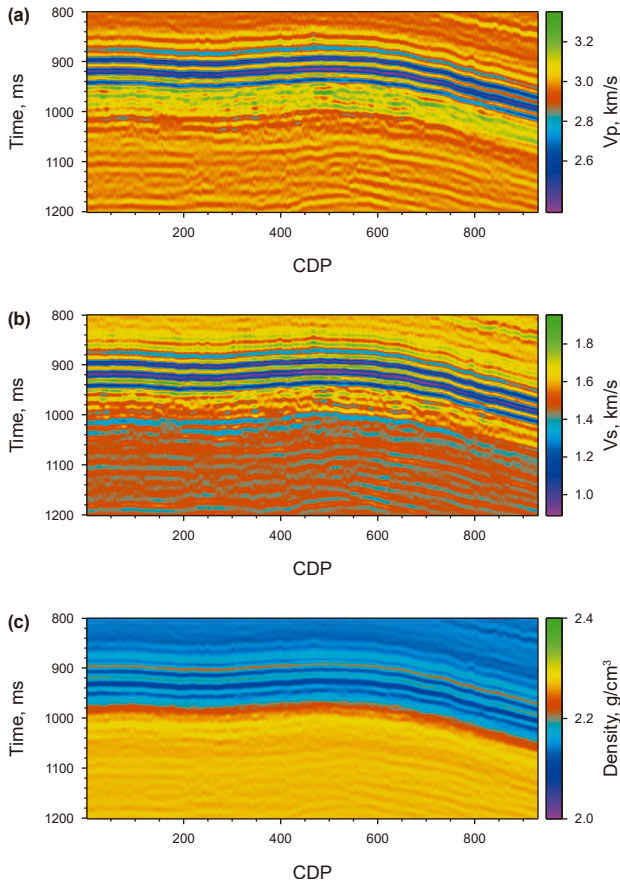


Fig. 15. DCT-BLI results of (a) P-wave velocity, (b) S-wave velocity, and (c) density.

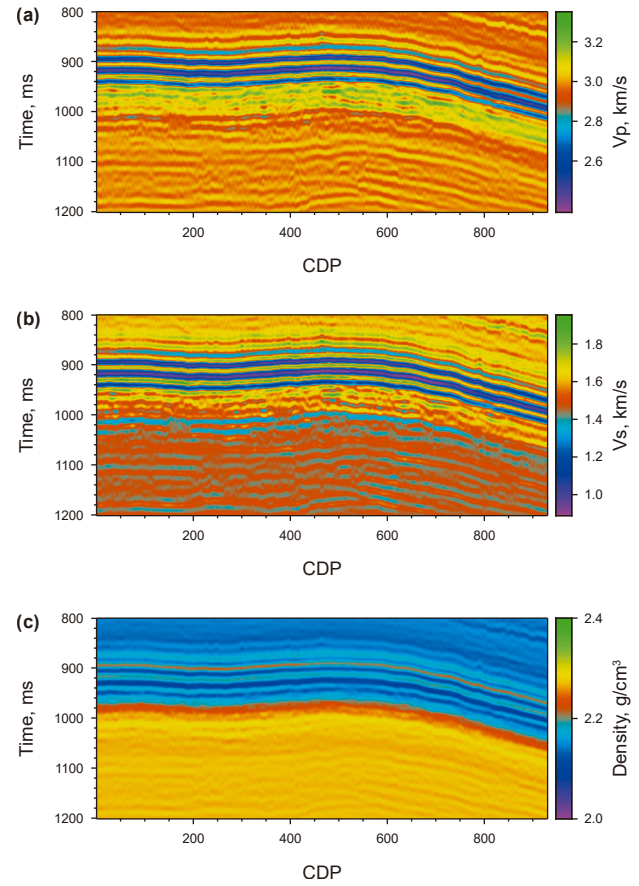


Fig. 16. IDCT-BLI results of (a) P-wave velocity, (b) S-wave velocity, and (c) density.

discrepancies from the traditional BLI results are essentially the same. Therefore, despite IDCT-BLI achieving a significantly higher level of dimensionality reduction compared to DCT-BLI, the inversion accuracy of both methods is nearly identical, demonstrating the effectiveness of the IDCT-based dimensionality reduction approach.

4. Discussion

The IDCT-based reduction method proposed in this study can significantly reduce the dimensionality of large, sparse kernel matrices in Kalman-based stochastic inversion, thereby addressing the computational challenges associated with matrix inversion in practical applications and substantially improving computational efficiency. This method builds on the theoretical foundation of the DCT-based dimensionality reduction approach (Yu et al., 2024), which has proven effective in compressing inversion kernel matrices to approximately 40%–50% of their original size without compromising inversion accuracy. In contrast, the proposed IDCT-based method offers the potential to compress the matrix dimensionality to as low as 15% (see the tests in Fig. 6). The advantage of the IDCT-based reduction lies in its iterative and progressive dimensionality reduction strategy. Essentially, matrix dimensionality reduction aims to retain critical information while discarding redundancies. The IDCT-based approach achieves this by performing small-scale reductions in each iteration, an incremental process that maximizes the preservation of critical matrix information throughout the reduction procedure.

This study conducted model tests with varying numbers of incremental dimensionality reduction steps. The results demonstrate that the maximum achievable reduction increases with the number of steps; however, beyond three iterations, the marginal benefit becomes negligible, as illustrated in Fig. 6. Moreover, excessive reduction steps introduce additional computational costs. Based on the findings of this study, the optimal number of iterations in the incremental reduction process is two to three. This strategy has also been validated through real data experiments. Current tests on this method suggest that two iterations are sufficient to achieve a significant reduction in matrix dimensionality for most cases.

The problem of inverting large sparse matrices is pervasive in fields such as geophysics and mathematical geology, with typical applications including inversion, deconvolution, denoising, and geological modeling. This presents non-negligible challenges to the practicality and scalability of the associated methods when applied to large-scale data. The proposed matrix information dimensionality reduction strategy offers an effective solution. However, existing reduction methods typically employ a single-step approach, resulting in limited reduction efficiency. This limitation arises because single-step dimensionality reduction removes non-informative components from the matrix in a single operation, which may compromise essential information. In contrast, the IDCT-based reduction iteratively extracts and progressively eliminates non-informative components, thereby maximizing the preservation of critical information while effectively discarding redundant data. In the IDCT method, the selection of Rcoe is

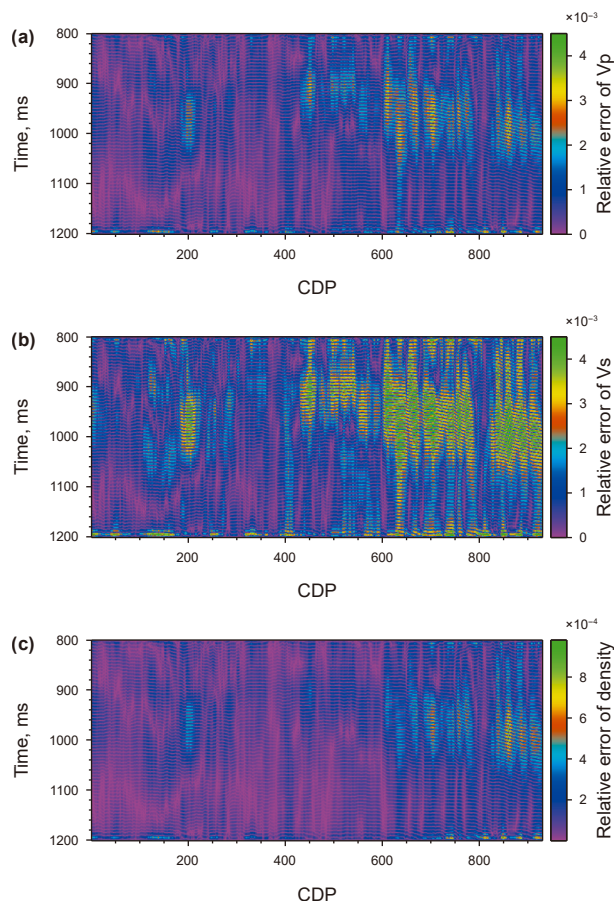


Fig. 17. The relative errors between the inversion results obtained by BLI and DCT-BLI for the (a) P-wave velocity, (b) S-wave velocity, and (c) density.

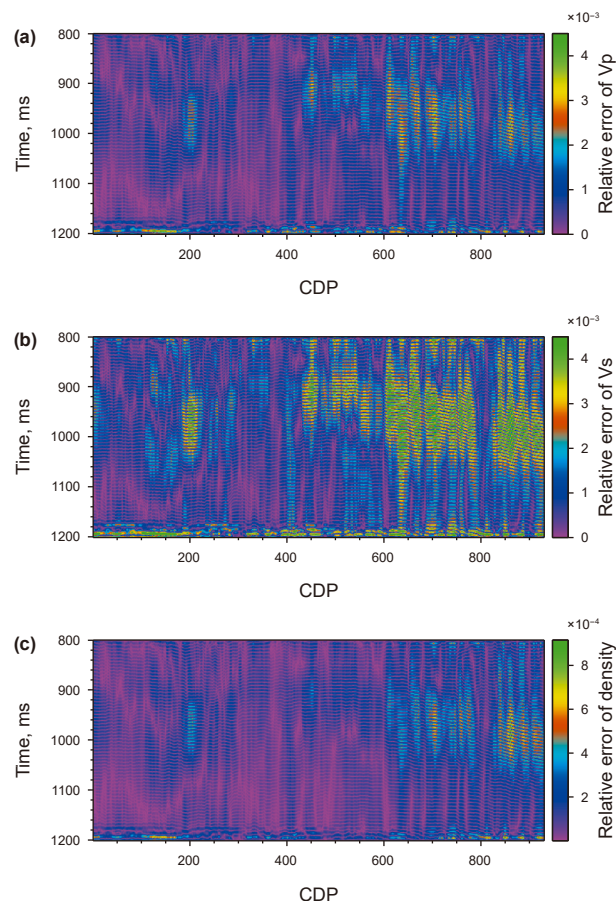


Fig. 18. The relative errors between the inversion results obtained by BLI and IDCT-BLI.

critical, which is determined through one-dimensional inversion tests currently. Another possible manner is choosing Rcoe by using certain optimization techniques.

Furthermore, existing dimensionality reduction methods applied to seismic inversion generally require computations such as principal component analysis, which inherently involve large matrix operations and consequently increase extra computational burden of seismic inversion. By comparison, the method proposed in this study employs the same dimensionality reduction operator for different seismic traces, thereby enhancing computational efficiency. Therefore, when performing two-dimensional or even three-dimensional dimensionality reduction inversion, the reduction operator only needs to be computed once during the inversion of the first trace, eliminating the need for recalculation in subsequent traces. This approach prevents the dimensionality reduction algorithm from introducing additional computational burden to the inversion process. Notably, this advantage is shared by both DCT-based and IDCT-based dimensionality reduction methods.

The complexity of geological structures also increases the difficulty of dimensionality reduction. As shown in the two-dimensional relative error maps in Figs. 17 and 18, the error is higher in regions with horizontal strata than in those with inclined strata, a phenomenon consistently observed in both numerical results and model validation tests. Nevertheless, the relative errors of the inversion results before and after dimensionality reduction remain within 1%, indicating that although complex geological features may reduce the accuracy of

dimensionality reduction, the inversion results after reduction are still reliable. Therefore, IDCT-based dimensionality reduction method can alleviate the challenge of large matrix inversion effectively. The advantages of IDCT become increasingly evident as the dimensionality of the matrix increases. Especially in scenarios involving large-scale geophysical data inversion and multi-channel data processing (Wang et al., 2024b), which necessitate the inversion of ultra-large matrices, IDCT is expected to offer substantial performance benefits.

5. Conclusions

A deep dimensionality reduction method named IDCT is presented to address the challenge of inverting large sparse matrices in Bayesian inverse problem. The method is fundamentally based on DCT and innovatively performs an iterative and incremental DCT-based dimensionality reduction on large sparse matrices inversion. The key to its effectiveness lies in calibrating a reasonable reduction magnitude in the iteration process, thereby preserving essential matrix information to the greatest extent. In other words, the reduction intensity in each iteration of this method remains moderate. The proposed method presents two notable advantages over traditional dimensionality reduction techniques. First, it achieves a substantially higher reduction ratio compared to conventional approaches. Second, it introduces no additional computational burden to the inversion process. This study conducts both synthetic and field data tests, comparing the proposed approach with traditional DCT-based dimensionality

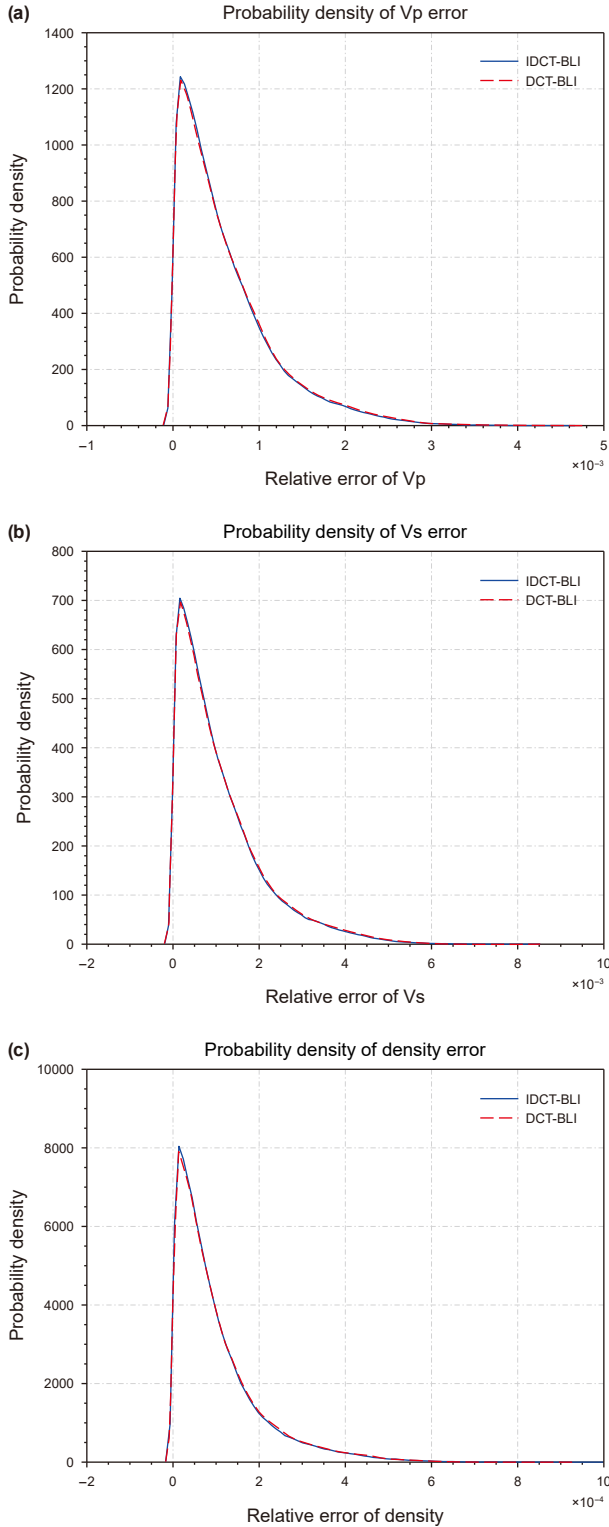


Fig. 19. The distributions of relative errors between the inversion results obtained by BLI and IDCT-BLI for the (a) P-wave velocity, (b) S-wave velocity, and (c) density.

reduction methods. The test results fully validate the effectiveness of the proposed method. This dimensionality reduction technique is particularly valuable for solving Bayesian inverse issue under the Kalman framework, where the inversion of large matrices is inevitable. Moreover, it has broader implications for geophysics

and mathematical geology, providing a valuable reference for addressing large sparse matrix inversion problems in these fields.

CRedit authorship contribution statement

Qing-Qing Li: Methodology, Conceptualization. **Bo Yu:** Methodology, Funding acquisition, Conceptualization. **Jia-Liang Xu:** Writing – original draft. **Ning Wang:** Visualization, Software. **Shi-Chao Wang:** Writing – original draft, Visualization. **Hui Zhou:** Writing – original draft, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is partly supported by Hainan Provincial Joint Project of Sanya Yazhou Bay Science and Technology City (2021JJLH0052); National Natural Science Foundation of China (42274154, 42304116), Natural Science Foundation of Heilongjiang Province, China (LH2024D013), and Heilongjiang Postdoctoral Fund (LBH-Z23103); Hainan Yazhou Bay Science and Technology City Jingying Talent Project (SKJC-JYRC-2024-05).

Appendix A

BLI is an excellent seismic stochastic inversion methodology, whose objective is estimating $p(\mathbf{m}|\mathbf{d})$ in Eq. (17) from the observed datasets. Generally, a Gaussian function is used herein to describe the prior distribution of the prestack elastic parameters including the P- and S-wave velocities and density:

$$\mathbf{m} \sim \mathcal{N}(\Psi_{\mathbf{m}}, \mathbf{C}_{\mathbf{m}}) \quad (\text{A-1})$$

where $\Psi_{\mathbf{m}}$ and $\mathbf{C}_{\mathbf{m}}$ are the mean value and covariance matrix.

The seismic forward modeling based on the convolution theorem is commonly expressed as the following formula,

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}, \quad (\text{A-2})$$

where $\mathbf{e}$ is the forward error. $\mathbf{G}$ is a forward operator. Generally, $\mathbf{G}$ can both represent the linear and non-linear forward relationships. However, this research focuses on the matrix dimension reduction, hence, the linear forward equation constructed by Aki-Richards is utilized herein to quantify the relationship between the model parameters and the observed seismic data (Aki and Richards, 2002). Considering the linear Gaussian theorem, the distribution of the seismic data $\mathbf{d}$ can be written as

$$\mathbf{d} \sim \mathcal{N}(\mathbf{G}\Psi_{\mathbf{m}}, \mathbf{G}\mathbf{C}_{\mathbf{m}}\mathbf{G}^T + \mathbf{C}_{\mathbf{e}}). \quad (\text{A-3})$$

where $\mathbf{C}_{\mathbf{e}}$ is the variance matrix for the seismic error, and it can represent the uncertainty of the seismic data. According the statistical theory, the joint distribution of model parameters and seismic data can be expressed as,

$$\begin{bmatrix} \mathbf{d} \\ \mathbf{m} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{G}\Psi_{\mathbf{m}} \\ \Psi_{\mathbf{m}} \end{bmatrix}, \begin{bmatrix} \mathbf{G}\mathbf{C}_{\mathbf{m}}\mathbf{G}^T + \mathbf{C}_{\mathbf{e}} & \mathbf{C}_{\mathbf{m}}\mathbf{G}^T \\ \mathbf{G}\mathbf{C}_{\mathbf{m}} & \mathbf{C}_{\mathbf{m}} \end{bmatrix}\right). \quad (\text{A-4})$$

where $\mathbf{C}_{\mathbf{m}}\mathbf{G}^T$ and $\mathbf{G}\mathbf{C}_{\mathbf{m}}$ are the transposition with each other, and they both present the statistical and spatial correlation between model parameters and seismic data. Using Eq. (A-4), the

conditional probability distribution of $\mathbf{m}$ constrained by $\mathbf{d}$ can be calculated and expressed as follows,

$$\mathbf{m}|\mathbf{d} \sim N(\Psi_{\mathbf{m}|\mathbf{d}}, \mathbf{C}_{\mathbf{m}|\mathbf{d}}), \quad (\text{A-5})$$

and it also includes a posterior mean term and a covariance matrix $\Psi_{\mathbf{m}|\mathbf{d}}$ and $\mathbf{C}_{\mathbf{m}|\mathbf{d}}$, and they can be expressed as

$$\Psi_{\mathbf{m}|\mathbf{d}} = \Psi_{\mathbf{m}} + \mathbf{C}_{\mathbf{m}} \mathbf{G}^T \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{d} - \mathbf{G} \Psi_{\mathbf{m}}), \quad (\text{A-6})$$

$$\mathbf{C}_{\mathbf{m}|\mathbf{d}} = \mathbf{C}_{\mathbf{m}} - \mathbf{C}_{\mathbf{m}} \mathbf{G}^T \Sigma \mathbf{C}_{\mathbf{d}}^{-1} \mathbf{C}_{\mathbf{m}}. \quad (\text{A-7})$$

where the core matrix for BLI is $\mathbf{C}_{\mathbf{d}} = \mathbf{G} \mathbf{C}_{\mathbf{m}} \mathbf{G}^T + \mathbf{C}_{\mathbf{e}}$. The objective of this research is improving the inverse efficiency of $\mathbf{C}_{\mathbf{d}}$.

References

- Aki, K., Richards, P.G., 2002. *Quantitative seismology*.
- Alemie, W., Sacchi, M.D., 2011. High-resolution three-term AVO inversion by means of a Trivariate Cauchy probability distribution. *Geophysics* 76 (3), R43–R55. <https://doi.org/10.1190/1.3554627>.
- Aravkin, A., Friedlander, M.P., Herrmann, F.J., et al., 2012. Robust inversion, dimensionality reduction, and randomized sampling. *Math. Program.* 134, 101–125. <https://doi.org/10.1007/s10107-012-0571-6>.
- Atkinson, S., Zabar, N., 2019. Structured Bayesian Gaussian process latent variable model: applications to data-driven dimensionality reduction and high-dimensional inversion. *J. Comput. Phys.* 383, 166–195. <https://doi.org/10.1016/j.jcp.2018.12.037>.
- Bosch, M., Mukerji, T., Gonzalez, E.F., 2010. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review. *Geophysics* 75 (5), 75A165–75A176. <https://doi.org/10.1190/1.3478209>.
- Buland, A., Omre, H., 2003. Bayesian linearized AVO inversion. *Geophysics* 68 (1), 185–198. <https://doi.org/10.1190/1.1543206>.
- Bunte, K., Schneider, P., Hammer, B., et al., 2012. Limited rank matrix learning, discriminative dimension reduction and visualization. *Neural Netw.* 26, 159–173. <https://doi.org/10.1016/j.neunet.2011.10.001>.
- Cao, Y., Zhou, H., Yu, B., et al., 2024. The estimation of petrophysical parameters based on ensemble smoother with correlation localization. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–12. <https://doi.org/10.1109/TGRS.2024.3403663>.
- Chen, F.B., Zong, Z.Y., Lang, K., et al., 2024a. Geofluid discrimination in stress-induced anisotropic porous reservoirs using seismic AVAZ inversion. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–14. <https://doi.org/10.1109/TGRS.2024.3477943>.
- Chen, F.B., Zong, Z.Y., Yin, X.Y., 2024b. Seismic scattering inversion for multiple parameters of overburden-stressed isotropic media. *Geophysics* 89 (6), T319–T336. <https://doi.org/10.1190/geo2023-0636.1>.
- Chen, S.Y., Wang, N., Shi, Y., et al., 2025. Sparse gabor transform and its application in seismic data analysis. *IEEE Trans. Geosci. Rem. Sens.* 63, 1–10. <https://doi.org/10.1109/TGRS.2025.3560299>.
- Cunningham, J.P., Ghahramani, Z., 2015. Linear dimensionality reduction: survey, insights, and generalizations. *J. Mach. Learn. Res.* 16 (1), 2859–2900. <https://doi.org/10.48550/arXiv.1406.0873>.
- Ding, P.B., Gong, F., Zhang, F., et al., 2021. A physical model study of shale seismic responses and anisotropic inversion. *Pet. Sci.* 18 (4), 1059–1068. <https://doi.org/10.1016/j.petsci.2021.01.001>.
- Downton, J.E., 2005. *Seismic Parameter Estimation from AVO Inversion*. Ph.D. Thesis. University of Calgary.
- Esser, E., Moller, M., Osher, S., et al., 2012. A convex model for nonnegative matrix factorization and dimensionality reduction on physical space. *IEEE Trans. Image Process.* 21 (7), 3239–3252. <https://doi.org/10.1109/TIP.2012.2190081>.
- Gao, Z., Li, C., Liu, N., et al., 2020. Large-dimensional seismic inversion using global optimization with autoencoder-based model dimensionality reduction. *IEEE Trans. Geosci. Rem. Sens.* 59 (2), 1718–1732. <https://doi.org/10.1109/TGRS.2020.2998035>.
- Grana, D., de Figueiredo, L., Azevedo, L., 2019. Uncertainty quantification in Bayesian inverse problems with model and data dimension reduction. *Geophysics* 84 (6), M15–M24. <https://doi.org/10.1190/geo2019-0222.1>.
- Grana, D., Azevedo, L., de Figueiredo, L., et al., 2022. Probabilistic inversion of seismic data for reservoir petrophysical characterization: review and examples. *Geophysics* 87 (5), M199–M216. <https://doi.org/10.1190/geo2021-0776.1>.
- Jumah, B., Herrmann, F.J., 2014. Dimensionality-reduced estimation of primaries by sparse inversion. *Geophys. Prospect.* 62 (5), 972–993. <https://doi.org/10.1190/1.3627931>.
- Liu, M., Grana, D., de Figueiredo, L., 2022. Uncertainty quantification in stochastic inversion with dimensionality reduction using variational autoencoder. *Geophysics* 87 (2), M43–M58. <https://doi.org/10.1190/geo2021-0138.1>.
- Marzouk, Y.M., Najm, H.N., 2009. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *J. Comput. Phys.* 228 (6), 1862–1902. <https://doi.org/10.1016/j.jcp.2008.11.024>.
- Penna, R., Lupinacci, W.M., 2024. Geostatistical seismic inversion and 3D modelling of metric flow units, porosity and permeability in Brazilian presalt reservoir. *Pet. Sci.* 21 (3), 1699–1718. <https://doi.org/10.1016/j.petsci.2024.02.013>.
- Pereira, P., Bordignon, F., Azevedo, L., et al., 2019. Strategies for integrating uncertainty in iterative geostatistical seismic inversion. *Geophysics* 84 (2), R207–R219. <https://doi.org/10.1190/geo2017-0758.1>.
- Pereira, P., Calçôa, I., Azevedo, L., et al., 2020. Iterative geostatistical seismic inversion incorporating local anisotropies. *Comput. Geosci.* 24, 1589–1604. <https://doi.org/10.1007/s10596-020-09966-1>.
- Shi, Y., Yu, B., Zhou, H., et al., 2024. FMG_INV, a fast multi-Gaussian inversion method integrating well-log and seismic data. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–12. <https://doi.org/10.1109/TGRS.2024.3351207>.
- Sun, Y.H., Dong, H.L., Chen, G., et al., 2025. Intelligent seismic AVO inversion method for brittleness index of shale oil reservoirs. *Pet. Sci.* 22 (2), 627–640. <https://doi.org/10.1016/j.petsci.2024.09.024>.
- Thurin, J., Brossier, R., Métivier, L., 2019. Ensemble-based uncertainty estimation in full waveform inversion. *Geophys. J. Int.* 219 (3), 1613–1635. <https://doi.org/10.1093/gji/ggz384>.
- Wang, N., Shi, Y., Ni, J., et al., 2024a. Enhanced seismic attenuation compensation: integrating attention mechanisms with residual learning in neural networks. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–11. <https://doi.org/10.1109/TGRS.2024.3445130>.
- Wang, P., Cui, Y.A., Zhou, L., et al., 2024b. Multi-task learning for seismic elastic parameter inversion with the lateral constraint of angle-gather difference. *Pet. Sci.* 21 (6), 4001–4009. <https://doi.org/10.1016/j.petsci.2024.06.010>.
- Yang, Y.M., Yin, X.Y., Li, K., et al., 2023. Fast pre-stack multi-channel inversion constrained by seismic reflection features. *Pet. Sci.* 20 (4), 2060–2074. <https://doi.org/10.1016/j.petsci.2023.02.009>.
- Yin, Z., Orozco, R., Herrmann, F.J., 2025. WISER: multimodal variational inference for full-waveform inversion without dimensionality reduction. *Geophysics* 90 (2), A1–A7. <https://doi.org/10.1190/geo2024-0483.1>.
- Yu, B., Zhou, H., Wang, L., et al., 2021. Prestack Bayesian linearized inversion with decorrelated prior information. *Math. Geosci.* 53, 437–464. <https://doi.org/10.1007/s11004-020-09899-6>.
- Yu, B., Shi, Y., Zhou, H., et al., 2024. Fast Bayesian linearized inversion with an efficient dimension reduction strategy. *IEEE Trans. Geosci. Rem. Sens.* (62), 1–10. <https://doi.org/10.1109/TGRS.2024.3360031>.
- Zahm, O., Cui, T., Law, K., et al., 2022. Certified dimension reduction in nonlinear Bayesian inverse problems. *Math. Comput.* 91 (336), 1789–1835. <https://doi.org/10.48550/arXiv.1807.03712>.
- Zhang, Y., Oliver, D.S., 2011. Evaluation and error analysis: kalman gain regularization versus covariance regularization. *Comput. Geosci.* 15, 489–508. <https://doi.org/10.1007/s10596-010-9218-y>.
- Zhang, G.Z., Yang, R., Li, L., et al., 2022. Elastic impedance inversion for stress indicator in weakly orthorhombic media. *Pet. Sci.* 19 (5), 2064–2080. <https://doi.org/10.1016/j.petsci.2022.03.020>.