

Article

A Variational Dispersion Mode Decomposition Approach to Extracting Normal-Mode Interference Spectra from Broadband Acoustic Intensity Measurements

Wei Gao * and Guocheng Gao

State Key Laboratory of Acoustics and Marine Information, Institute of Acoustics, Chinese Academy of Sciences, No.21 North 4th Ring West Road, Beijing 100190, China; gaoguocheng@mail.ioa.ac.cn

* Correspondence: gaowei@mail.ioa.ac.cn

Abstract

The primary objective of separating normal-mode interference spectra (NMISs) is to extract more fine-grained coherent structures between different pairs of modes in a shallow water sound field. However, NMISs generally exhibit nonlinear phase frequency and amplitude frequency relationships in one-dimensional broadband sound intensity measurements, which poses a challenge for effective separation. To address the difficulty arising from the nonlinearity of NMISs, this study presents a variational dispersion mode decomposition (VDMD) method. An innovation of VDMD is to incorporate the physical model of NMISs into the framework of Variational Nonlinear Chirp Mode Decomposition (VNCMD). First, it utilizes the waveguide invariant and the interference frequency factor to characterize the phase function of NMISs. Second, this parameterization of the nonlinear phase function provides an analytical expression for the gradient of the objective function in the variational optimization processing of VNCMD, thereby improving the efficiency of its iterative search and enabling accurate extraction of NMISs. Finally, experimental results from a Yellow Sea trial demonstrate that the separated NMISs agree well with the reference modes, validating the effectiveness of our proposed method.

Keywords: variational dispersion mode decomposition; normal-mode interference spectrums; waveguide invariant

1. Introduction

In a shallow water waveguide, the underwater acoustic intensity excited by a point source can be expressed as the superposition of incoherent and coherent components [1–4]. The incoherent component is the sum of the intensities of different normal modes and generally varies slowly with frequency and range. The coherent component characterizes the mutual interference between pairs of normal modes and often exhibits interfering acoustic intensity fringes in a two-dimensional (2D) frequency-range (r - ω) plane. At a fixed range, the one-dimensional (1D) coherent component (as a function of frequency) is characterized by a superposition of quasi-periodic fluctuations produced by each pair of normal modes, especially when the intensities of the modes are comparable [5,6]. In this 1D case, we define the sound intensity fluctuation caused by mutual interference of a pair of normal modes as the normal-mode interference spectrum (NMIS). Previous studies [5–

Academic Editor: Weicheng Cui

Received: 7 June 2026

Revised: 14 July 2026

Accepted: 16 July 2026

Published: 20 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

10] have shown that NMISs generally exhibit nonlinear phase frequency and amplitude frequency relationships. While this nonlinearity carries the waveguide dispersion information, it also makes it more difficult to finely separate and extract the NMIS from the 1Dcoherent component of the acoustic intensity.

To overcome this difficulty, this article attempts to introduce the constraints of the physical model of the NMIS into the framework of Variational Nonlinear Chirp Mode Decomposition (VNCMD) [11] and proposes a variational dispersion mode decomposition (VDMD) method for NMIS separation. VNCMD is a power tool used to estimate the increment of instantaneous frequency using arctangent demodulation technology and regularize the updating process of instantaneous frequency to ensure its smoothness. The greatest advantage of VNCMD is that it establishes a rigorous mathematical model based on variational optimization and the alternating direction method of multipliers (ADMM), transforming signal decomposition into an optimal demodulation problem, thereby outperforming traditional empirical mode decomposition (EMD) algorithms—which lack a clear optimization objective and rely on empirical sifting—in terms of decomposition accuracy, noise robustness, and the ability to mitigate mode mixing. However, in VNCMD, the phase of the signal model lies in a high-dimensional parameter space, which makes it sensitive to initialization conditions and difficult to optimize when dealing with the NMIS separation problem [9]. Fortunately, it is found that the horizontal wavenumber difference in the NMIS phase function depends on the waveguide invariant [12–14]. This prior knowledge from the NMIS physical model provides a physical basis for the initialization and reduces the model dimension of the optimization problem. In addition, the parameterization of the horizontal wavenumber difference provides an analytic form of the gradient of the objective function, which leads to a more definite direction of optimization in the iterations. Simulations and experimental results verify that VDMD has better convergence performance and separation accuracy for NMIS separation than VNCMD. The experimental acoustic data analyzed in this study were acquired using a Type 8106 hydrophone (Brüel & Kjær, Nærum, Denmark), and all numerical simulations and data-processing procedures were performed using MATLAB R2023b (MathWorks, Natick, MA, USA).

This article consists of the following sections. Section 2 introduces the NMIS and some related notions. Section 3 presents the principle and steps of the proposed VDMD method. Sections 4 and 5 present simulation analysis and experimental verification, respectively. Section 6 presents the conclusions.

2. Preliminary

2.1. The Model of NMIS

According to normal-mode theory, the coherent component of the underwater broadband sound intensity excited by a point sound source at depth z_s and angular frequency ω and received at depth z after propagating through a long distance r can be expressed as follows [3]:

$$I_c(\omega) = \sum_{n=1}^N \sum_{m>n}^N I_{mn}(\omega) = \sum_{n=1}^N \sum_{m \neq n}^N A_m A_n \cos\{[k_m(\omega) - k_n(\omega)]r\} \quad (1)$$

$$A_m = \Psi_m(z_s, \omega) \Psi_m(z, \omega) \frac{e^{-\alpha_m(\omega)r}}{\sqrt{k_m(\omega)r}}$$

where N represents the propagation normal-mode number excited by the sound source, $I_{mn}(\omega)$ is the NMIS between the m th and n th normal modes, Ψ_m and k_m represent the

vertical modal function and horizontal wavenumber of the m th normal mode, respectively, and α_m is the attenuation factor of the m th normal mode.

When the frequency of interest is far from the cutoff frequency, the horizontal wavenumber difference between modes m and n can be expressed as follows [2]:

$$k_m(\omega) - k_n(\omega) = \gamma_{mn}\omega^{-1/\beta_{mn}}, \quad (2)$$

where γ_{mn} is a constant determined by the m th and n th modes; β_{mn} , which describes the dispersion characteristic of the ocean waveguide, is the waveguide invariant for the m th and n th modes. It is worth noting that the VDMD proposed in this paper is based on Equation (2), so its applicable frequency band must be kept away from the cutoff frequency. Substituting Equation (2) into Equation (1), $I_c(\omega)$ can be reformulated as follows:

$$I_c(\omega) = \sum_{l=1}^L B_l \cos(\xi_l \omega^{-1/\beta_l}) \quad (3)$$

where symbol l replaces symbols m and n for convenience, $L = \frac{N(N-1)}{2}$ is the total number of NMISs, $\xi_l = \gamma_l r$ is defined as the interference frequency factor, and $B_l = A_m A_n$ is the amplitude of l th NMIS. B_l is mainly affected by the depth of the source and hydrophone and is weakly dependent on the angular frequency.

We define the interference frequency as follows:

$$f_l(\omega) = \frac{1}{2\pi} \left| \frac{d[k_m(\omega) - k_n(\omega)]}{d\omega} r \right| = \left| \frac{\xi_l \omega^{-1/\beta_l-1}}{2\pi\beta_l} \right| \quad (4)$$

Generally, the interference frequency of an NMIS is a smooth function that varies much more slowly than the phase function. Thus, an NMIS can be considered an amplitude-modulated and frequency-modulated broadband non-stationary signal. The phase function of the NMIS is parameterized by the waveguide invariant β_l and the interference frequency factor ξ_l , which provides a physical constraint basis for the NMIS separation method proposed later in this paper.

2.2. VNCMD

VNCMD is a recently introduced method for nonlinear chirp signal decomposition that has aroused notable attention in various fields [11]. We consider the signal $g(t)$ sampled at $t = t_0, \dots, t_{P-1}$ in the following equation:

$$\begin{aligned} g(t) &= \sum_{i=1}^Q a(t) \cos\left(2\pi \int_0^t f_i(s) ds + \phi_i\right) \\ &= \sum_{i=1}^Q u_i(t) \cos\left(2\pi \int_0^t \tilde{f}_i(s) ds\right) + v_i(t) \sin\left(2\pi \int_0^t \tilde{f}_i(s) ds\right) \end{aligned} \quad (5)$$

where Q is the number of nonlinear chirp modes; $u_i(\omega)$ and $v_i(\omega)$ are two demodulated signals:

$$\begin{aligned} u_i(t) &= a(t) \cos\left(2\pi \int_0^t (f_i(s) - \tilde{f}_i(s)) ds + \phi_i\right) \\ v_i(t) &= -a(t) \sin\left(2\pi \int_0^t (f_i(s) - \tilde{f}_i(s)) ds + \phi_i\right) \end{aligned} \quad (6)$$

where $\tilde{f}_i$ represents the instantaneous frequency of the demodulation operators $\cos\left(2\pi \int_0^t \tilde{f}_i(s) ds\right)$ and $\sin\left(2\pi \int_0^t \tilde{f}_i(s) ds\right)$. The two demodulated signals will have the narrowest band corresponding to the bandwidth of the amplitude $a(t)$ if $\tilde{f}_i(s) = f_i(s)$.

The idea of VNCMD is to solve the mode separation problem by minimizing the bandwidths of the demodulated signals:

$$\min_{u_i, v_i, f_i} \left\{ \sum_{i=1}^Q (\|\Omega \mathbf{u}_i\|_2^2 + \|\Omega \mathbf{v}_i\|_2^2) \right\} \quad (7)$$

$$\text{s.t. } \|g - \{\sum_{i=1}^Q A_i u_i + B_i v_i\}\|_2 \leq \varepsilon$$

where $\|\cdot\|_2$ stands for the l_2 norm, $\mathbf{A}_i = \text{diag} [\cos(2\pi \int_0^{t_0} \tilde{f}_i(s) ds), \dots, \cos(2\pi \int_0^{t_{N-1}} \tilde{f}_i(s) ds)]^T$ and $\mathbf{B}_i = \text{diag} [\sin(2\pi \int_0^{t_0} \tilde{f}_i(s) ds), \dots, \sin(2\pi \int_0^{t_{N-1}} \tilde{f}_i(s) ds)]^T$ are discrete forms of two demodulated signals, $\mathbf{u}_i = [u_i(t_0), \dots, u_i(t_{p-1})]^T$ and $\mathbf{v}_i = [v_i(t_0), \dots, v_i(t_{p-1})]^T$ are discrete forms of two demodulation operators, ε is the tolerance upper bound determined by the noise, the bandwidth is evaluated using the square of the l_2 norm of the second-order derivative, and Ω is a second-order difference operator formulated as

$$\Omega = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & -2 & 1 \\ 0 & \dots & 0 & 1 & -1 \end{bmatrix} \quad (8)$$

The constrained optimization problem is solved using the augmented Lagrange multiplier method and the alternating direction method of multipliers. Two demodulated orthogonal signals u_i and v_i are calculated iteratively as follows:

$$u_i^{k+1} = \left(\frac{2}{\alpha} \Omega^T \Omega + A_i^T A_i \right)^{-1} A_i^T \left(g - \sum_{p \neq i} A_p u_p - \sum_p B_p v_p - w - \frac{\lambda}{\alpha} \right), \quad (9)$$

$$v_i^{k+1} = \left(\frac{2}{\alpha} \Omega^T \Omega + B_i^T B_i \right)^{-1} B_i^T \left(g - \sum_{p \neq i} A_p u_p - \sum_p B_p v_p - w - \frac{\lambda}{\alpha} \right),$$

where w is a noise variable, $\lambda \in \mathbb{R}^{W \times 1}$ is the Lagrangian multiplier, and $\alpha > 0$ is a penalty parameter. VNCMD uses the arctangent demodulation technique to obtain the instantaneous frequency increment:

$$\Delta \tilde{f}_i^{k+1}(t) = -\frac{1}{2\pi} \frac{d}{dt} \left(\arctan \left(\frac{v_i^{k+1}(t)}{u_i^{k+1}(t)} \right) \right) \quad (10)$$

$$= \frac{v_i^{k+1}(t) \cdot (u_i^{k+1}(t))' - u_i^{k+1}(t) \cdot (v_i^{k+1}(t))'}{2\pi \left((u_i^{k+1}(t))^2 + (v_i^{k+1}(t))^2 \right)}.$$

VNCMD assumes that the instantaneous frequency is a smooth function, and the final increment can be corrected by a low-pass filter:

$$\Delta f_i^{k+1} = \left(\frac{2}{\mu} \Omega^T \Omega + I \right)^{-1} \Delta \tilde{f}_i^{k+1} \quad (11)$$

It is worth noting that the instantaneous frequency update in VNCMD is a compromise approach, and if the initial instantaneous frequency differs greatly from the real value, VNCMD may fail. The detailed VNCMD algorithm is shown in Ref. [11].

3. VDMD

Building upon the physical model of NMISs established in Section 2.1 and the variational framework of VNCMD reviewed in Section 2.2, we propose a variational dispersion mode decomposition (VDMD) method to overcome the convergence and

accuracy issues of VNCMD for NMIS separation. The basic framework of VDMD is the same as that of VNCMD, which is to achieve modal separation by minimizing the bandwidth of nonlinear modes. The difference is that VDMD introduces the constraint of the NMIS model, which is more suitable for solving the NMIS separation problem. As shown in Equation (4), the nonlinearity of the NMIS interference frequency depends on the waveguide invariant and interference frequency factor. By using this relationship, the analytic form of the gradient of the objective function can be given, and the optimization direction of the interference frequency in the iteration can be determined more precisely. Demodulation, as shown in Equation (7) and Equation (3), can be expressed as follows:

$$\begin{aligned} I_c(\omega) &= \sum_{l=1}^L B_l \cos(\xi_l \omega^{-1/\beta_l}) \\ &= \sum_{l=1}^L u_l \cos(\tilde{\xi}_l \omega^{-1/\tilde{\beta}_l}) + v_l \sin(\tilde{\xi}_l \omega^{-1/\tilde{\beta}_l}) \end{aligned} \quad (12)$$

where

$$\begin{aligned} u_l(\omega) &= B_l \cos(\xi_l \omega^{-1/\beta_l} - \tilde{\xi}_l \omega^{-1/\tilde{\beta}_l}), \\ v_l(\omega) &= -B_l \sin(\xi_l \omega^{-1/\beta_l} - \tilde{\xi}_l \omega^{-1/\tilde{\beta}_l}). \end{aligned} \quad (13)$$

The two scalars $\tilde{\xi}_l$ and $\tilde{\beta}_l$ replace $\tilde{f}_l$, which lies in the high-dimensional parameter space in Equation (9). The optimization problem of VDMD can be expressed as

$$\begin{aligned} \min_{u_l, v_l, \tilde{\xi}_l, \tilde{\beta}_l} & \left\{ \sum_{l=1}^L (\|\Omega u_l\|_2^2 + \|\Omega v_l\|_2^2) \right\} \\ \text{s.t. } & \|I_c - \sum_{l=1}^L (A_l u_l + B_l v_l)\|_2 \leq \varepsilon \end{aligned} \quad (14)$$

where I_c is $I_c(\omega)$ sampled in the band $[\omega_0, \omega_1, \dots, \omega_{W-1}]$:

$$u_l = [u_l(\omega_0), \dots, u_l(\omega_{W-1})]^T, v_l = [v_l(\omega_0), \dots, v_l(\omega_{W-1})]^T,$$

$$A_l = \text{diag} \left[\cos(\tilde{\xi}_l \omega_0^{-1/\tilde{\beta}_l}), \dots, \cos(\tilde{\xi}_l \omega_{W-1}^{-1/\tilde{\beta}_l}) \right]^T,$$

$$B_l = \text{diag} \left[\sin(\tilde{\xi}_l \omega_0^{-1/\tilde{\beta}_l}), \dots, \sin(\tilde{\xi}_l \omega_{W-1}^{-1/\tilde{\beta}_l}) \right]^T.$$

To solve the constrained optimization problem in Equation (16), the corresponding augmented Lagrangian is expressed as follows:

$$\begin{aligned} \mathcal{L}(u_l, v_l, w, \tilde{\xi}_l, \tilde{\beta}_l, \lambda) &= \mathcal{J}_{c_e}(w) + \sum_{l=1}^L (\|\Omega u_l\|_2^2 + \|\Omega v_l\|_2^2) \\ &+ \frac{\alpha}{2} \|w - I_c + \sum_{l=1}^L (A_l u_l + B_l v_l)\|_2^2 + \frac{\lambda}{\alpha} \left(\frac{\lambda}{2} - \frac{\|\lambda\|_2^2}{2\alpha} \right) \end{aligned} \quad (15)$$

where $\mathcal{J}_{c_e}(w)$ is an indicator function. VDMD uses the alternating direction multiplier method (ADMM) to decompose a complex optimization problem into several separately updated subproblems. We will address the issue of updating $\tilde{\xi}_l$ and $\tilde{\beta}_l$, which influence the interference frequency, in Equation (4). The other variable optimization processes are the same as in VNCMD. The gradient of the two sub-optimization problems for $\tilde{\xi}_l$ and $\tilde{\beta}_l$ can be expressed as follows:

$$\begin{aligned}\frac{\partial}{\partial \xi_l} \mathcal{L}(u_l, v_l, \xi_l, \tilde{\beta}_l, w, \lambda) &= \alpha (A_l v_l - B_l u_l)^T H_\xi \left(w - I_c + \sum_{l=1}^L (A_l u_l + B_l v_l) + \frac{\lambda}{\alpha} \right) \\ \frac{\partial}{\partial \tilde{\beta}_l} \mathcal{L}(u_l, v_l, \xi_l, \tilde{\beta}_l, w, \lambda) &= \alpha (A_l v_l - B_l u_l)^T H_\beta \left(w - I_c + \sum_{l=1}^L (A_l u_l + B_l v_l) + \frac{\lambda}{\alpha} \right)\end{aligned}\quad (16)$$

$$\text{where } \begin{cases} H_\xi = \text{diag} \left[\omega_0^{-1/\tilde{\beta}_l}, \dots, \omega_{W-1}^{-1/\tilde{\beta}_l} \right] \\ H_\beta = \text{diag} \left[\frac{\xi_l \omega_0^{-1/\tilde{\beta}_l} \ln \omega_0}{\tilde{\beta}_l^2}, \dots, \frac{\xi_l \omega_{W-1}^{-1/\tilde{\beta}_l} \ln \omega_{W-1}}{\tilde{\beta}_l^2} \right]. \end{cases}$$

Then the Adaptive Moment Estimation (Adam) algorithm [15] is used to update ξ_l and $\tilde{\beta}_l$. Adam is an optimization algorithm that can adaptively adjust the learning rate according to the parameter's gradient history. Because of the adaptive adjustment of the learning rate, the Adam algorithm is insensitive to the initial value of the learning rate, which improves the algorithm's robustness. Take ξ_l as an example:

$$\begin{aligned}d &= \frac{\partial}{\partial \xi_l} \mathcal{L} \\ m_\xi &= \kappa_1 m_\xi^{k-1} + (1 - \kappa_1) d \\ v_\xi &= \kappa_2 v_\xi^{k-1} + (1 - \kappa_2) d^2 \\ \hat{m}_\xi &= \frac{m_\xi}{1 - \kappa_1} \\ \hat{v}_\xi &= \frac{v_\xi}{1 - \kappa_2} \\ \xi_l^{k+1} &= \xi_l^k - \frac{\eta_\xi}{\sqrt{\hat{v}_\xi} + \tau} \hat{m}_\xi\end{aligned}\quad (17)$$

where m_ξ and v_ξ are the gradient estimations of the first and second moments, respectively; $\hat{m}_\xi$ and $\hat{v}_\xi$ are estimations corrected for bias to provide a more accurate update initially. In this article, κ_1 and κ_2 are set to 0.9 and 0.999, respectively, as decay rates controlling the weighted average of the past gradient and the gradient squared, $\tau = 8 \times 10^3$, are used to maintain numerical stability. η_ξ is a hyperparameter that controls the step length of each parameter update.

In VNCMD, ξ_l and $\tilde{\beta}_l$ are replaced by non-parameterized instantaneous frequency functions, and the instantaneous frequency functions are updated using the arctangent demodulation technique and regularization (see Equations (12) and (13)). In other words, VNCMD has a broader application range in nonlinear mode decomposition because it does not require a prescribed phase model, whereas VDMD is a physics-informed specialization designed for NMIS decomposition. The additional physical constraint improves the convergence robustness of VDMD for NMISs, but its applicability depends on the validity of Equation (4). The complete VDMD algorithm is shown in Algorithm 1.

Algorithm 1 VDMD

Initialize: $\mathbf{I}_c; \alpha; \mu; \xi_l^1; \tilde{\beta}_l^1; \mathbf{A}_l^1 = \text{diag} \left[\cos \left(\xi_l^1 \omega^{-1/\tilde{\beta}_l^1} \right) \right], \mathbf{B}_l^1 = \text{diag} \left[\sin \left(\xi_l^1 \omega^{-1/\tilde{\beta}_l^1} \right) \right], k = 0$
 $\mathbf{u}_l^1 = (2/\alpha \mathbf{\Omega}^T \mathbf{\Omega} + (\mathbf{A}_l^1)^T \mathbf{A}_l^1)^{-1} (\mathbf{A}_l^1)^T \mathbf{I}_c, \mathbf{v}_l^1 = (2/\alpha \mathbf{\Omega}^T \mathbf{\Omega} + (\mathbf{B}_l^1)^T \mathbf{B}_l^1)^{-1} (\mathbf{B}_l^1)^T \mathbf{I}_c, \lambda^1 = 0$

- 1: **repeat**
- 2: $k = k + 1$
- 3: $\mathbf{w}^{k+1} = \mathcal{P}_{c_r} (\mathbf{I}_c - \sum_l (\mathbf{A}_l^k \mathbf{u}_l^k + \mathbf{B}_l^k \mathbf{v}_l^k) - \lambda^k / \alpha)$
- 4: **for** $l = 1:L$ **do**
- 5: $\mathbf{u}_l^{k+1} = (2/\alpha \mathbf{\Omega}^T \mathbf{\Omega} + \mathbf{A}_l^T \mathbf{A}_l)^{-1} \mathbf{A}_l^T (\mathbf{I}_c - \sum_{p \neq l} \mathbf{A}_p \mathbf{u}_p - \sum_p \mathbf{B}_p \mathbf{v}_p - \mathbf{w}^{k+1} - \lambda^k / \alpha)$
- 6: $\mathbf{v}_l^{k+1} = (2/\alpha \mathbf{\Omega}^T \mathbf{\Omega} + \mathbf{B}_l^T \mathbf{B}_l)^{-1} \mathbf{B}_l^T (\mathbf{I}_c - \sum_p \mathbf{A}_p \mathbf{u}_p - \sum_{p \neq l} \mathbf{B}_p \mathbf{v}_p - \mathbf{w}^{k+1} - \lambda^k / \alpha)$
- 7: $\xi_l^{k+1} = \xi_l^k - \eta_\xi / (\sqrt{\hat{v}_\xi} + \epsilon) * \hat{m}_\xi$
- 8: $\tilde{\beta}_l^{k+1} = \tilde{\beta}_l^k - \eta_\beta / (\sqrt{\hat{v}_\beta} + \epsilon) * \hat{m}_\beta$
- 9: $\mathbf{A}_l^{k+1} = \text{diag} \left[\cos \left(\xi_l^{k+1} \omega^{-1/\tilde{\beta}_l^{k+1}} \right) \right]$
- 10: $\mathbf{B}_l^{k+1} = \text{diag} \left[\sin \left(\xi_l^{k+1} \omega^{-1/\tilde{\beta}_l^{k+1}} \right) \right]$
- 11: **end for**

```

12: if  $\|\mathbf{w}^{k+1} + \sum_l (\mathbf{A}_l^{k+1} \mathbf{u}_l^{k+1} + \mathbf{B}_l^{k+1} \mathbf{v}_l^{k+1}) - \mathbf{I}_c\|_2^2 > \|\mathbf{I}_c\|_2^2$  then
13:    $\lambda^{k+1} = 0$ 
14:   for  $l = 1:L$  do
15:      $\mathbf{u}_l^{k+1} = (2/\alpha \mathbf{\Omega}^T \mathbf{\Omega} + (\mathbf{A}_l^{k+1})^T \mathbf{A}_l^{k+1})^+ (\mathbf{A}_l^{k+1})^T \mathbf{I}_c$ 
16:      $\mathbf{v}_l^{k+1} = (2/\alpha \mathbf{\Omega}^T \mathbf{\Omega} + (\mathbf{B}_l^{k+1})^T \mathbf{B}_l^{k+1})^+ (\mathbf{B}_l^{k+1})^T \mathbf{I}_c$ 
17:   end for
18: end if
19:  $\lambda^{k+1} = \lambda^k + \alpha (\mathbf{w}^{k+1} + \sum_l \mathbf{I}_l^{k+1} - \mathbf{I}_c)$ 
20: until  $\|\mathbf{I}_c^{k+1} - \mathbf{I}_c^k\|_2^2 / \|\mathbf{I}_c^k\|_2^2 \leq \delta$ 

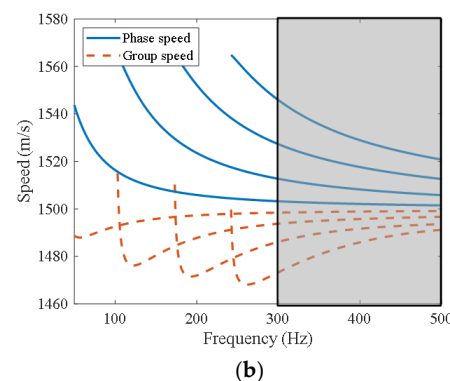
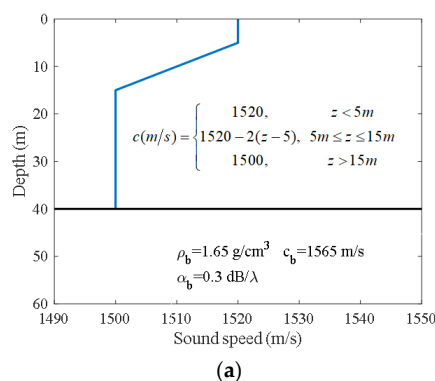
```

4. Simulation

4.1. VDMD for NMIS Separation

Specific numerical simulations are used to demonstrate the performance of VDMD in NMIS separation. Figure 1a shows a simulated waveguide, which has a negative thermocline at depths of 5–15 m. The source and hydrophone are both located 30 m below the sea surface. The horizontal distance between the source and receiver is 10 km. The synthesized acoustic data in the frequency band of interest, 300 Hz to 500 Hz, is calculated by KRAKEN [16]. The sound field is mainly dominated by normal modes 1–4, and their group velocity and phase velocity are shown in Figure 1b. It is worth noting that the frequency band of interest is far from the cutoff frequency, which satisfies the assumption of Equation (2). Figure 1c plots the vertical modal function of the first to fourth normal modes in 400 Hz, and the black dashed line represents the depth of the source and hydrophone. Gaussian white noise is added to the data and the signal-to-noise ratio (SNR) is 20 dB. Figure 1d shows the coherent component of broadband intensity with noise, the reconstruction waveform and the modes separated by VDMD. The separated modes respectively correspond to 1–2, 1–3, 1–4, and 2–4 pair NMISs. The good consistency validates the effectiveness of VDMD. The 2–3 and 3–4 pair NMISs are submerged in noise because the amplitude of the third normal mode is minimal, as shown in Figure 1c.

It should be noted that the key advantage of VDMD over the existing VNCMD algorithm lies in its use of the physical characteristic constraints of the NMIS to improve the iterative search efficiency of the objective function, thereby effectively avoiding the problems of sensitivity to initial values and susceptibility to local minima when separating NMISs. Therefore, in the simulation example of Figure 1, we do not present a direct comparison between VDMD and VNCMD; instead, we provide a statistical comparison in the subsequent analysis of issues such as initial values, convergence, and signal-to-noise ratio.



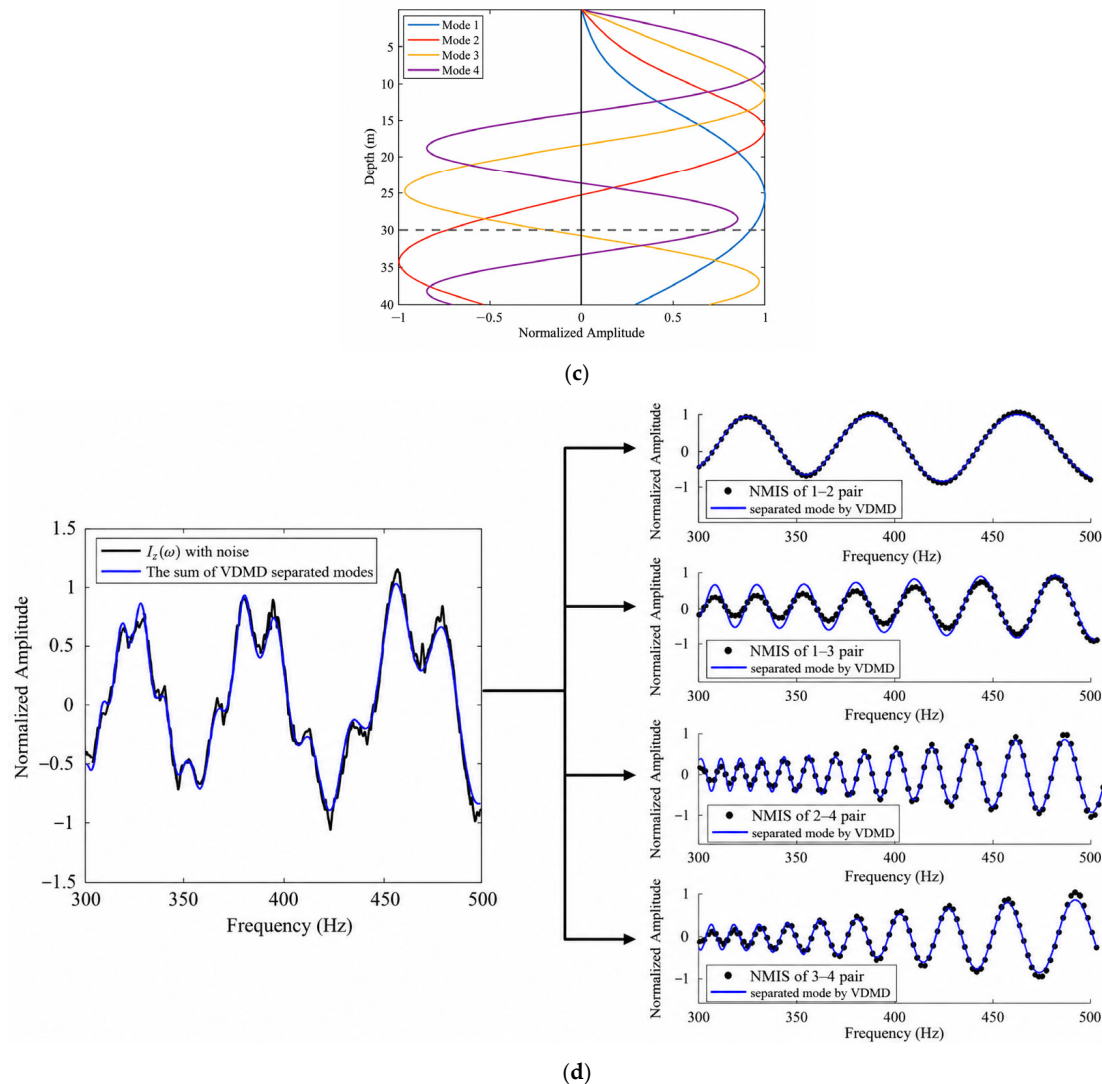


Figure 1. (a) Simulation environment schematic diagram. (b) The group velocity and phase velocity of normal modes 1–4. The shaded area is the frequency band of interest. (c) The vertical modal functions of normal modes 1–4 from 300 Hz to 500 Hz. (d) The coherent component of broadband intensity with noise, the reconstruction waveform and the modes separated by VDMD.

4.2. Initialization of β_l and ξ_l

The theoretical and experimental values of β_l are approximately equal to one in the ideal Pekeris waveguide [1], so we suggest setting $\tilde{\beta}_l^1 = 1$ in the initialization. The initialization of $\tilde{\xi}_l^1$ can be estimated by the interference frequency peaks using the Fourier transform. Assuming the frequency corresponding to the l th peak is f_l^{peak} , and the center angular frequency of the interested frequency band is ω_c , it can be obtained according to Equation (4):

$$\tilde{\xi}_l^1 = \frac{2\pi f_l^{\text{peak}} \tilde{\beta}_l^1}{\omega_c^{-1/\tilde{\beta}_l^1 - 1}} = 2\pi f_l^{\text{peak}} \omega_c^2 \quad (18)$$

Here, we use the interference frequency at the center angular frequency ω_c to approximate the peak frequency to obtain the initialization of $\tilde{\xi}_l^1$. It is worth noting that there is a coupling relationship between β_l and ξ_l in representing the interference frequency. The relative error (RE) is defined as follows:

$$RE(\tilde{\xi}, \tilde{\beta}) = \frac{\|f(\tilde{\xi}, \tilde{\beta}) - f(\xi, \beta)\|_2}{\|f(\xi, \beta)\|_2} = \frac{\left\| \frac{\tilde{\xi} \omega^{-1/\tilde{\beta}-1}}{\tilde{\beta}} - \frac{\xi \omega^{-1/\beta-1}}{\beta} \right\|_2}{\left\| \frac{\xi \omega^{-1/\beta-1}}{\beta} \right\|_2}, \quad (19)$$

where $\tilde{\xi}$ and $\tilde{\beta}$ are the estimated values; β_l and ξ_l denote the real values. Considering a typical 1–3 NMIS with nonlinear phase frequency and amplitude frequency relationships in Figure 1d and setting a search interval for $\tilde{\xi}$ and $\tilde{\beta}$, the REs calculated over the interval are shown in Figure 2a. In Figure 2a, the real β_l and ξ_l of the NMIS are located in the position of the two white dashed lines. There is a curved band with a low RE from the upper left to the lower right of the true positions, which indicates that the interference frequency can be represented by many combinations of $\tilde{\xi}$ and $\tilde{\beta}$, and there is a coupling relationship between $\tilde{\xi}$ and $\tilde{\beta}$. Although the estimates of β_l and ξ_l may be different from their true values, this does not affect the representation of the interference frequency. Figure 2b and Figure 2c, respectively, draw the RE curves with the fixed real β_l and ξ_l , corresponding to the RE on the two white dashed lines in Figure 2a. It can be seen that the RE is more sensitive to the change in $\tilde{\beta}$. Therefore, we suggest that a small update step of $\tilde{\beta}$ should be selected, which is conducive to the algorithm's stability.

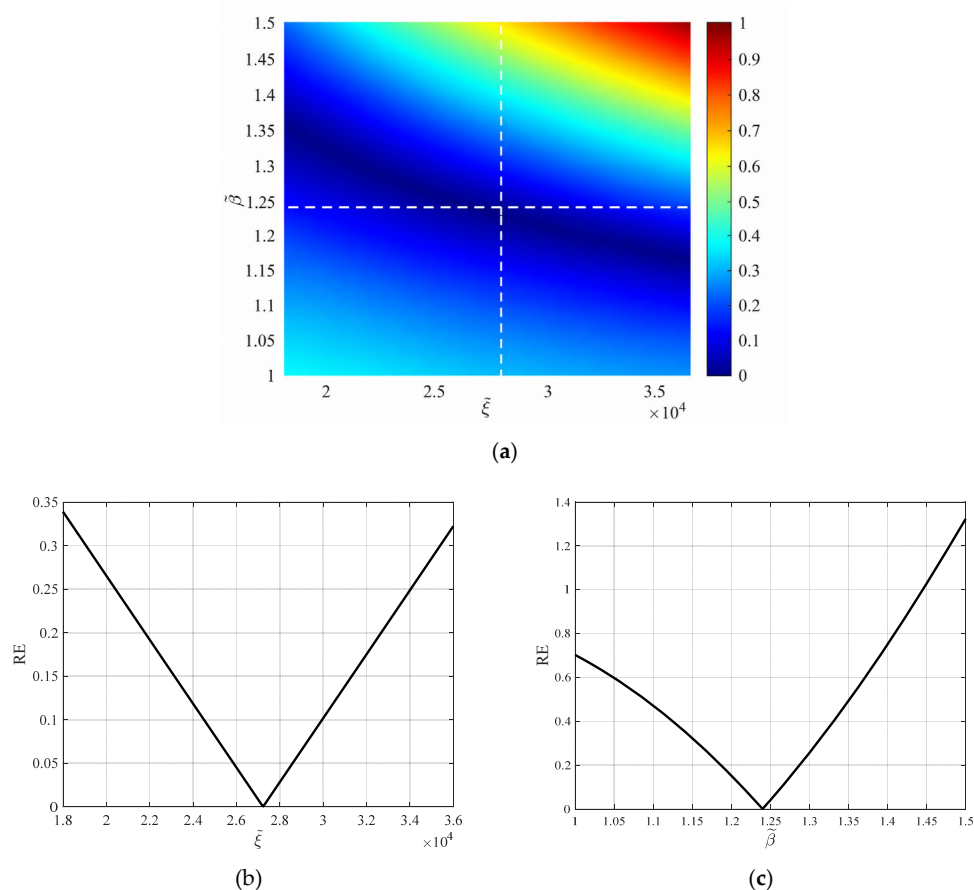


Figure 2. (a) The RE of interference frequency in the prior search interval; the intersection of the two white dashed lines is the true position of β_l and ξ_l . (b) RE along the line with the real ξ_l . (c) RE along the line with the real β_l .

4.3. Convergence of the VDMD

To measure the accuracy of modal reconstruction, we define the convergence degree as follows:

$$C_{\text{deg}} = 20 \log_{10} \left(\frac{\|I_c(\omega)\|_2}{\|I_c(\omega) - \tilde{I}_c(\omega)\|_2} \right), \quad (20)$$

where $I_c(\omega)$ and $\tilde{I}_c(\omega)$ represent the real and estimated NMISs. Firstly, the convergence of the VDMD is verified by the single NMIS in Figure 2a. We generate 100 randomly generated initial ξ^1 values:

$$\xi^1 = \xi \cdot (1 + 0.5e_r), \quad (21)$$

where e_r belongs to a normal distribution function with a mean of 0 and a variance of one. In the absence of noise, the convergence performance of VDMD and VNCMD is tested using the above 100 random initial ξ^1 values. As is shown in Section 4.2, the initial value of waveguide variance $\tilde{\beta}^1$ is set to one. Here, the initial ξ^1 and $\tilde{\beta}^1$ have been converted to the same initial frequency function in VNCMD. The penalty factors for both VDMD and VNCMD are set to 0.01, and the maximum number of iterations is set to 200. Figure 3a plots the final convergence degree of VDMD and VNCMD. It can be seen from Figure 3a that VDMD has a higher convergence degree than VNCMD in the trials. Two typical convergence curves of VDMD and VNCMD are shown in Figure 3b. It can be seen from Figure 3b that, for the representative trial shown here, VDMD converges faster and reaches the prescribed termination tolerance, whereas VNCMD converges to a stationary solution with a relatively large reconstruction error. This behavior can be attributed to the different optimization strategies of the two methods. In VNCMD, the instantaneous frequency curve is updated as a non-parametric function in a relatively high-dimensional optimization space. Consequently, when the initial instantaneous frequency curve deviates substantially from the true curve, the iteration may converge to a local stationary solution that minimizes the demodulated bandwidth but does not represent the physical NMIS phase. In contrast, VDMD restricts the interference frequency curve to the physically motivated form in Equation (4), thereby reducing the dimension of the search space and providing an analytical gradient for updating the phase-related parameters.

Nevertheless, the comparison in Figure 3 should not be interpreted as indicating that VNCMD necessarily converges to a local minimum in every nonlinear mode decomposition problem. VNCMD can converge accurately when the initial instantaneous frequency curves are sufficiently close to the true ones, the nonlinear modes are well separated, and the SNR is sufficiently high. In particular, when the NMIS phase nonlinearity is weak over a relatively narrow frequency band, VNCMD and VDMD may achieve comparable reconstruction accuracy. However, inaccurate initialization, close interference frequency curves, weak component amplitudes, and noise can increase the susceptibility of the non-parametric instantaneous frequency update to a local stationary solution. VDMD reduces this susceptibility by restricting the interference frequency curve to a physically motivated parametric form. Thus, the advantage observed here is specific to NMIS decomposition under the tested conditions and does not imply that VDMD universally outperforms VNCMD for all nonlinear chirp signals.

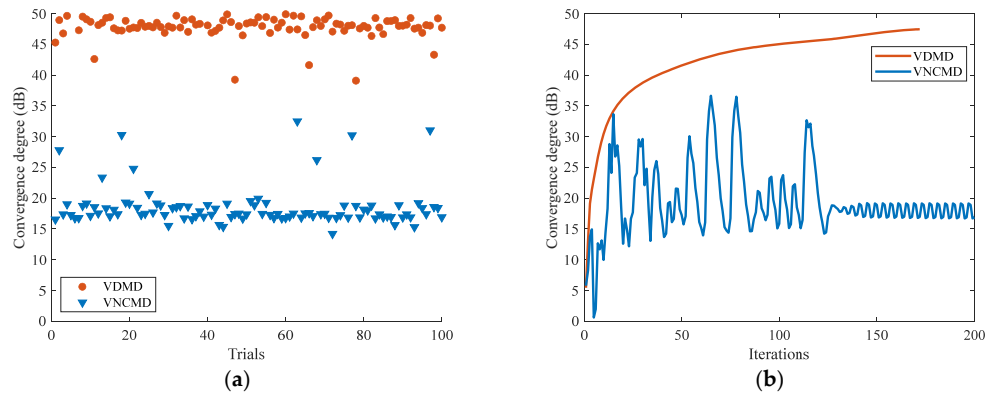


Figure 3. (a) Final convergence degrees of VDMD and VNCMD for 100 randomly generated initial interference frequency curves. (b) Representative convergence curves of VDMD and VNCMD under the same initialization condition.

4.4. Denoising Performance

Next, we will test the performance of the VDMD under different signal-to-noise ratios (SNRs). We set the maximum number of iterations to 500 and the convergence tolerance to $1e-8$. In this example, the errors of the initial interference frequency are 38% and 85%. In the case of the above initialization, we test the algorithm 100 times with the single NMIS in Figure 2a at different SNRs, and the success rate is shown in Figure 4. Here, we consider the test successful only if the RE of the estimated interference frequency is less than 5%. Figure 4 shows that VDMD maintains a relatively high success rate at low SNRs (e.g., SNR = 0 dB) when the initial interference frequency RE is 38%. Under the two initialization error conditions considered here, the overall success rate of VNCMD is lower than that of VDMD, particularly at low SNRs and for a larger initialization error. This result indicates that the physical parameterization adopted by VDMD improves its robustness to the combined effects of noise and initialization uncertainty in the present NMIS decomposition problem. It should be noted that the performance of VNCMD also depends strongly on the accuracy of the initial instantaneous frequency curve. With sufficiently accurate initialization and well-separated components, VNCMD may achieve a success rate comparable to that of VDMD. Therefore, Figure 4 does not imply that VNCMD is intrinsically less noise-resistant in all nonlinear mode decomposition applications.

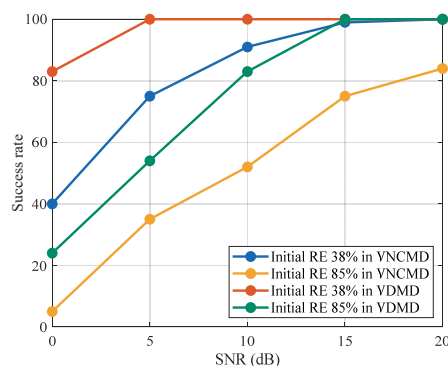


Figure 4. Success rates of VDMD and VNCMD at different SNRs and initial interference frequency REs. Each result was calculated from 100 independent noise realizations, and a trial was regarded as successful when the RE of the estimated interference frequency was less than 5%.

Finally, we provide another example of NMIS separation for multimode superposition in a Pekeris waveguide. The water depth, source depth, and receiver depth are 30 m, 15 m, and 20 m, respectively. The distance between the source and receiver is 10 km. The sound speed in water is 1510 m/s. The sound speed, density, and attenuation of the bottom are 1700 m/s, 1.7 g/cm³, and 0.1 dB/λ, respectively. Figure 5a shows the interference frequency curves of three NMISs resulting from interference among the first three normal modes. Gaussian white noise with an SNR of 10 dB was added to the signal, after which both VNCMD and VDMD were applied to decompose the three NMISs. As shown in Figure 5b, for the separation results above 300 Hz corresponding to the NMIS between two and three normal modes, the VNCMD results exhibit a significant deviation from the theoretical true values. In this case, VDMD achieves better mode separation performance than VNCMD. For this example, the VNCMD iteration converges to a stationary solution that deviates from the theoretical interference frequency curve above 300 Hz, indicating local convergence under the adopted initialization and parameter settings. In contrast, VDMD incorporates the physical dispersion model of the NMIS and therefore obtains a more accurate separation result. This example illustrates the greater convergence robustness of VDMD when the NMIS components have relatively close interference frequencies, but it does not imply that VNCMD fails under all initialization conditions.

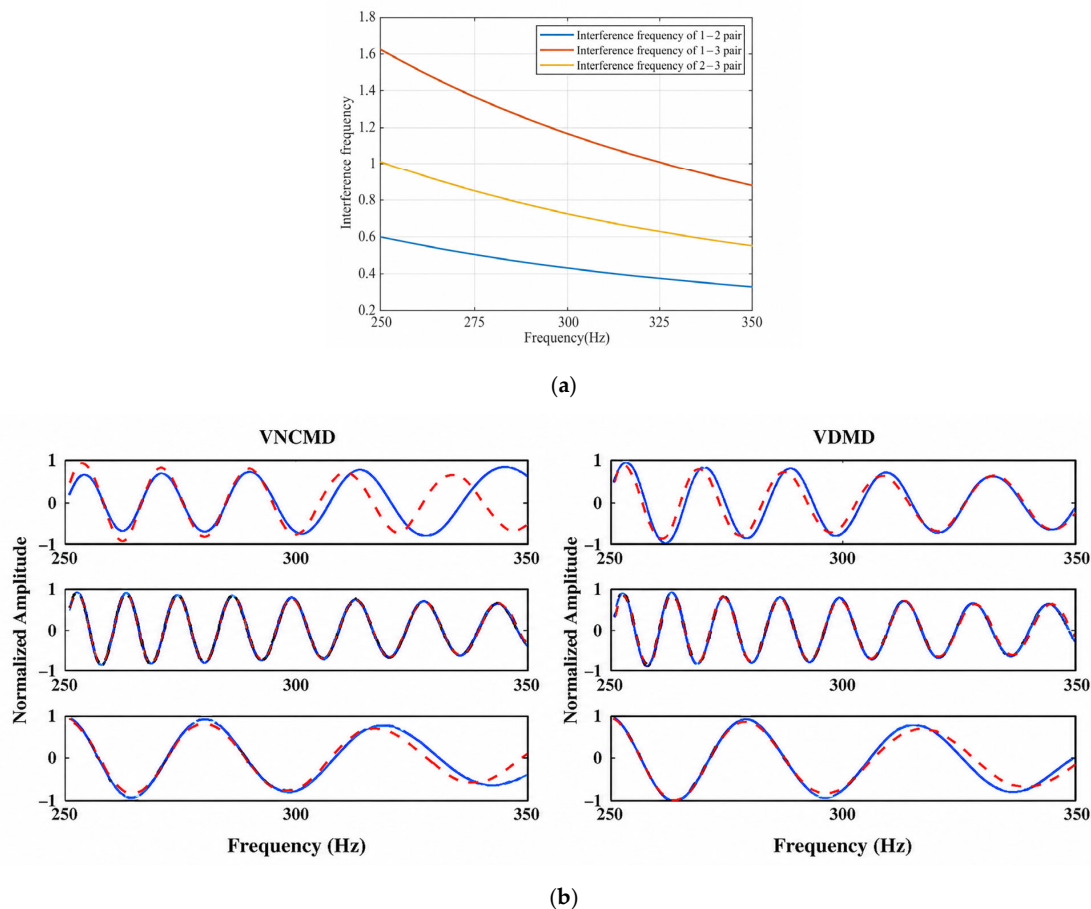


Figure 5. (a) The interference frequency curves of three NMISs in a Pekeris waveguide. (b) The separated NMISs with VNCMD and VDMD.

5. Experimental Results

In December 2011, an underwater acoustic propagation experiment was carried out in an area of the Yellow Sea. The seafloor in the test area is flat and the depth of the sea is about 60 m. The SSP is shown in Figure 6a, and the sound speed is roughly uniform. The hydrophone was laid on the seabed. The sound source circulates a linear frequency modulation signal with a center frequency of 380 Hz and a bandwidth of 100 Hz at a depth of 5 m. One can obtain the distance between the source and the receiver through the data recorded by GPS. Figure 6b shows a 2.56 s signal received by a single hydrophone at a horizontal distance of 1.92 km from the source. The one-dimensional broadband sound intensity with the trend term removed is shown in Figure 6c, and obvious fluctuations can be observed. To achieve sufficient frequency resolution, we extract a 10 s data segment that includes the signal and the ambient ocean noise both preceding and following it, and then compute the power spectrum. The NMISs are separated using VDMD, and the sum of the separated NMISs is shown in Figure 6c. Compared with the simulated NMISs using the test sound speed profile and geoacoustic inversion model [17], the separated NMISs respectively correspond to 5–8, 2–7, 3–7 and 2–5 pair normal modes. It can be seen from Figure 6c that the separated NMISs show good agreement with the reference results, with correlation coefficients for all NMISs exceeding 0.95. This result validates the effectiveness of using VDMD to separate NMISs. For comparison, VNCMD was also applied to the same experimental data, as shown in Figure 6c. Although the sum of the VNCMD components approximately reproduces the main large-scale fluctuations of the measured coherent component, its individual components corresponding to the 5–8, 2–7, and 3–7 mode pairs exhibit evident deviations from the reference NMISs. The 2–5 component is reasonably recovered by VNCMD, but its agreement with the reference NMIS is still less accurate than that obtained by VDMD. In contrast, all four components separated by VDMD closely follow the corresponding reference NMISs, with correlation coefficients exceeding 0.95. This comparison further demonstrates the advantage of incorporating the physical NMIS dispersion constraint into VDMD for component-level separation of the experimental data. It should be noted that the seabed affects the experimental mode structure through its reflection and absorption properties. In the present analysis, these effects were included in the KRAKEN normal-mode calculation using the measured sound–speed profile and the geoacoustic model obtained from the inversion results in Ref. [17]. The corresponding bottom parameters influence the modal wavenumbers, modal attenuation, and the reference NMISs shown in Figure 6c. The water depth was approximately 60 m, while the acoustic wavelength in the analyzed 330–430 Hz band was approximately 3.5–4.5 m, giving a depth-to-wavelength ratio of approximately 13–17. This ratio supports multiple propagating waterborne modes but does not by itself determine the fraction of energy coupled into the seabed. Quantitative acoustic-to-seismic conversion also depends on the elastic and stratified properties of the bottom. Because shear-wave parameters were not measured during this experiment, explicit elastic seabed coupling is beyond the scope of the present NMIS separation study.

One possible application of NMIS separation is typical guide-source ranging [18]. This paper presents preliminary results of sea trial data processing. A detailed analysis will be reported elsewhere. Suppose there is a broadband $[\omega_0, \omega_{W-1}]$ guide source with the same depth as the target source and a distance r_g from the hydrophone. The broadband fluctuation number T_l^g of the l th NMIS from the broadband guide source can be expressed as

$$T_l^g = \left| \int_{\omega_0}^{\omega_{W-1}} f_l^g(\omega) d\omega \right| \quad (22)$$

where $f_l^g(\omega)$ is the l th interference frequency of NMIS from the guide source. According to Equation (4), T_l^g is proportional to the horizontal distance, and the distance r_s between the source and the hydrophone can be estimated through the following:

$$r_s = \chi^* r_g$$

$$\chi^* = \arg \min_{\chi} \sum_{l=1}^L \|T_l^g - \chi T_l^s\|_2^2 \quad (23)$$

where $\chi \in \mathbb{R}^+$ and T_l^s is the broadband fluctuation number of the l th NMIS from the source. We assume that the guide source is 1.92 km away from the hydrophone, and the signal in Figure 6b is excited by a guided source. The distance of the other four sites can be estimated by the above method, where T_l^s is estimated using VDMD. For the four experimental ranges of 1, 3, 4, and 5 km, the preliminary range estimates obtained using this guide-source approach have a mean relative error of 7.6%. This result is provided only as an illustrative application of the separated NMISs and should not be interpreted as the general ranging accuracy of VDMD. In this work, we assumed that the normal-mode numbers corresponding to the NMIS are known. When multiple NMISs are observable, how to obtain their corresponding mode numbers remains an issue worthy of future attention, for example, using the mode-matching method.

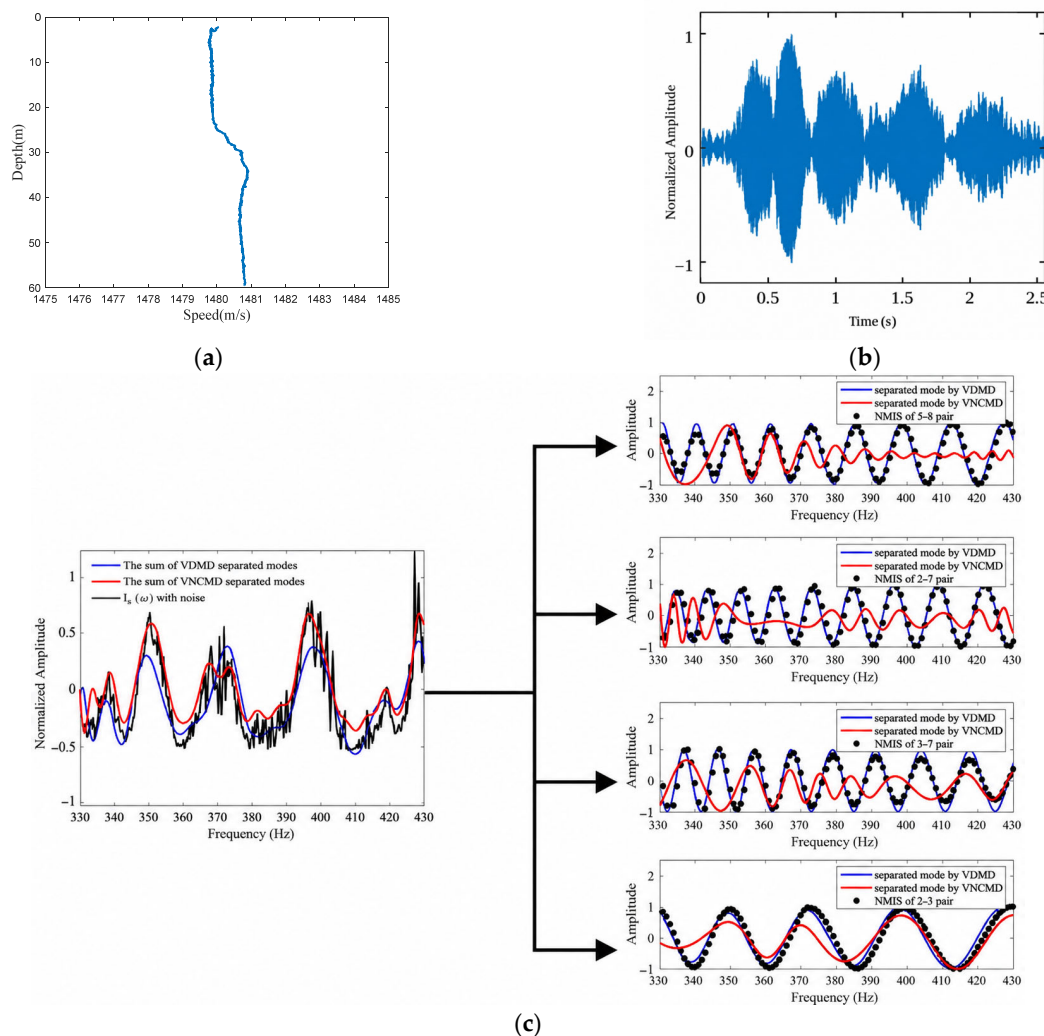


Figure 6. (a) The sound–speed profile in the experiment area. (b) A 2.56 s signal received by a single hydrophone at a horizontal distance of 1.92 km from the source. (c) The measured coherent

component and the NMIS separation results obtained by VDMD, with VNCMD included as a comparison. The black dots denote the simulated reference NMISs for the 5–8, 2–7, 3–7, and 2–5 mode pairs.

6. Conclusions

In this paper, a VDMD method for NMIS separation based on VNCMD and waveguide invariant theory is presented. By incorporating the physical model and dispersion characteristics of NMISs, VDMD improves both separation accuracy and convergence performance compared to VNCMD. Under the initialization and noise conditions considered in the numerical examples, the statistical results indicate that VDMD has a faster convergence rate, a higher successful-decomposition rate, and better reconstruction accuracy than VNCMD for NMIS decomposition. These improvements result from the physical parameterization of the interference frequency curve, which reduces the dimension of the optimization space and alleviates the sensitivity to initialization. Experimental results from the 2011 Yellow Sea campaign verified the effectiveness of our proposed VDMD. Moreover, a preliminary guide-source ranging example demonstrates that the separated NMISs retain physically meaningful range-dependent interference information. This example is intended only to illustrate a potential application of NMIS separation, while the development and evaluation of a dedicated source-ranging method are beyond the scope of the present study. Future work will focus on more experimental validation and on applying NMIS separation results to passive source ranging or waveguide invariant estimating.

Author Contributions: W.G.: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Visualization, Resources, Investigation, and Writing—Review and Editing. G.G.: Resources, Investigation, and Data Curation. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw reanalysis datasets analyzed in this study are publicly available from their respective official repositories. The in situ observational data and the custom code developed for data extraction and analysis are available from the corresponding author on request.

Acknowledgments: We express our sincere gratitude to all personnel involved in instrument deployment, data collection, and preliminary data processing for their support. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chuprov, S.D.; Brekhovskikh, L.M. Interference Structure of a Sound Field in a Layered Ocean. In *Ocean Acoustics, Current State*; Nauka: Moscow, Russia, 1982; pp. 71–91.
2. D’Spain, G.L.; Kuperman, W.A. Application of waveguide invariants to analysis of spectrograms from shallow water environments that vary in range and azimuth. *J. Acoust. Soc. Am.* **1999**, *106*, 2454–2468.
3. Jensen, F.B.; Kuperman, W.A.; Porter, M.B. *Computational Ocean Acoustics*, 2nd ed.; Springer: New York, NY, USA, 2011.
4. Godin, O.A.; Katsnelson, B.G.; Tan, T.W. Normal Mode Dispersion and Time Warping in the Coastal Ocean. *J. Acoust. Soc. Am.* **2019**, *146*, EL205–EL211.
5. Gao, W.; Li, X.; Wang, H. Multi-resolution estimation of the interference spectrum per pair of modes in the frequency domain. *J. Acoust. Soc. Am.* **2000**, *148*, EL340–EL346.
6. Zhen-Dong, Z.; Ning, W.; Da-Zhi, G.; Hao-Zhong, W. Broadband Source Ranging in Shallow Water Using the Ω -Interference Spectrum. *Chin. Phys. Lett.* **2010**, *27*, 064301.

7. Qi, Y.B.; Zhou, S.H.; Zhang, R.H.; Ren, Y. A Waveguide-Invariant-Based Warping Operator and Its Application to Passive Source Range Estimation. *J. Comput. Acoust.* **2015**, *23*, 1550003.
8. Song, W.; Gao, D.; Li, X.; Kang, D.; Li, Y. Waveguide Invariant in a Gradual Range-and Azimuth-Varying Waveguide. *JASA Express Lett.* **2022**, *2*, 056002.
9. Zhang, S.; Piao, S. Broadband Sound Intensity Interference Frequency Periodicity and Pulse Source Localization. *J. Mar. Sci. Eng.* **2021**, *9*, 200.
10. Niu, H.; Zhang, R.; Li, Z. Theoretical Analysis of Warping Operators for Non-Ideal Shallow Water Waveguides. *J. Acoust. Soc. Am.* **2014**, *136*, 53–65.
11. Chen, S.; Dong, X.; Peng, Z.; Zhang, W.; Meng, G. Nonlinear Chirp Mode Decomposition: A Variational Method. *IEEE Trans. Signal Process.* **2017**, *65*, 6024–6037.
12. Byun, G.; Song, H. The Waveguide Invariant for a Pekeris Waveguide. *J. Acoust. Soc. Am.* **2022**, *151*, 846–851.
13. Tao, H.; Krolik, J.L. Waveguide Invariant Focusing for Broadband Beamforming in an Oceanic Waveguide. *J. Acoust. Soc. Am.* **2008**, *123*, 1338–1346.
14. Bonnel, J.; Thode, A.; Wright, D.; Chapman, R. Nonlinear Time-Warping Made Simple: A Step-by-Step Tutorial on Underwater Acoustic Modal Separation with a Single Hydrophone. *J. Acoust. Soc. Am.* **2020**, *147*, 1897–1926.
15. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
16. Porter, M.B. *The KRAKEN Normal Mode Program*; Naval Research Laboratory: Washington, DC, USA, 1992.
17. Gao, W.; Wang, N.; Wang, H.Z. Statistical Geoacoustic Inversion from Vertical Correlation of Shallow Water Reverberation. *Acta Acoust.* **2008**, *33*, 109–115.
18. Turgut, A.; Orr, M.; Rouseff, D. Broadband Source Localization Using Horizontal-Beam Acoustic Intensity Striations. *J. Acoust. Soc. Am.* **2010**, *127*, 73–83.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.