



HAL
open science

Automated approach for source location in shallow waters

Angèle Niclas, Josselin Garnier

► **To cite this version:**

Angèle Niclas, Josselin Garnier. Automated approach for source location in shallow waters. 2023. hal-04173073

HAL Id: hal-04173073

<https://hal.science/hal-04173073v1>

Preprint submitted on 28 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automated approach for source location in shallow waters

Angèle Niclas^a and Josselin Garnier^b

CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, Palaiseau, 91120, France

(Dated: 28 July 2023)

This paper proposes a fully automated method for recovering the location of a source and medium parameters in shallow waters. The scenario involves an unknown source emitting low-frequency sound waves in a shallow water environment, and a single hydrophone recording the signal. Firstly, theoretical tools are introduced to understand the robustness of the warping method and to propose and analyze an automated way to separate the modal components of the recorded signal. Secondly, using the spectrogram of each modal component, the paper investigates the best way to recover the modal travel times and provides stability estimates. Finally, a penalized minimization algorithm is presented to recover estimates of the source location and medium parameters. The proposed method is tested on experimental data of right whale gunshot and combustive sound sources, demonstrating its effectiveness in real-world scenarios.

[<https://doi.org/DOI number>]

[XYZ]

Pages: 1–12

I. INTRODUCTION

This paper presents a fully automated method for recovering the location of a source and medium parameters in shallow waters. Specifically, we consider a scenario where an unknown source emits low-frequency sound waves (typically less than 500 Hz) in a coastal environment, and a single hydrophone records the signal at a distance of more than 1 km from the source. The ability to accurately recover the source location from the recorded signal has important military applications for localizing quiet sources (Ainslie, 2010), as well as for monitoring fish (Stanton *et al.*, 2018) and whale populations (Mellinger *et al.*, 2007).

To study this inverse problem and propose a fully automated resolution method, we model the source as a stationary point source and the shallow water environment as a semi-infinite Pekeris waveguide (Pekeris, 1948) with a finite layer of water and an infinite layer of sediments. Under these ideal circumstances the Fourier transform \hat{u} of the recorded signal u can be decomposed into modal components,

$$\hat{u}(\omega) = \sum_{n \in \mathbb{N}} A_n(\omega) e^{i\Phi_n(\omega)}, \quad (1)$$

where n is the mode index, A_n is the complex amplitude of mode n , $\Phi_n = -k_n r$ with k_n the horizontal wavenumber of mode n and r the distance from the source to hydrophone (Jensen *et al.*, 2011). However, due to uncertainties in the propagation environment and

the source characteristics, the modal decomposition (1) of the recorded signal \hat{u} is not exact and is difficult to extract and process.

Different techniques to extract information about the source and environment have been proposed. One of the most commonly used techniques is called Time-of-Arrival (TOA) estimation (see for instance (Aubauer *et al.*, 2000; Bonnel, 2010)). This method involves measuring the modal travel times $t_n(\omega) = -\Phi'_n(\omega)$ (where the prime stands for the frequency derivative) at the hydrophone and using this information to estimate the distance to the source. Methods have been developed to retrieve directly the phases Φ_n from the signal \hat{u} (Bonnel *et al.*, 2012), but there are not very effective when data are noisy. Most existing methods take advantage of the modal dispersion curves in the time-frequency domain, and consists in warping the signal to extract each modal component, recovering the dispersion curves $\omega \mapsto t_n(\omega)$ associated with each mode using time-frequency analysis, and matching estimated dispersion curves with simulated replicas.

The warping method is a widely used technique that aims at increasing the distance between modes in the time-frequency space, making it easier to separate them (see a full description in (Bonnel *et al.*, 2020)). Its main advantage is its robustness to environmental uncertainties. However, as pointed out in (Bonnel *et al.*, 2020), "Warping is not a universal solution and requires expertise and judgment to be used effectively, as there are currently no automated methods for bulk warping." This comment also applies to the recovery of dispersion curves using time-frequency analysis, where there is no known automated method for recovering dispersion curves with stability estimates. Most experimental data are thus processed manually, depending on the situation (Bonnel *et al.*, 2018; Thode *et al.*, 2017). The inverse problem,

^aangele.niclas@polytechnique.edu

^bjosselin.garnier@polytechnique.edu

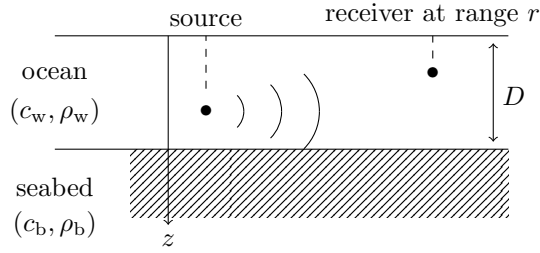


FIG. 1. Scheme of the Pekeris waveguide.

usually presented as a minimization problem between the estimated dispersion curves and simulated replicas, also lacks of an automated procedure. In practice, the minimization problem is solved numerically by computing the function's value everywhere and taking its minimum value, which can be computationally expensive.

This paper proposes a solution to these limitations by presenting a completely automated method to recover source location in shallow waters. The design of this automated method is based on the warping method (Bonnell *et al.*, 2018) and on several additional theoretical arguments. Specifically, we provide theoretical tools to understand the robustness of the warping method and use a watershed algorithm to implement it automatically. We then present two methods to recover the dispersion curves $\omega \mapsto t_n(\omega)$ and quantify their stability with noisy data. Finally, we implement a penalized minimization algorithm that can use prior estimates on some medium parameters (if available) to estimate the source location.

The paper is organized as follows: Section II describes the general framework of the Pekeris waveguide and the associated modal decomposition. In Section III, we focus on the warping technique and provide an automated way to separate modal components in the signal. In Section IV, we introduce and compare the maximum and mean methods for recovering dispersion curves and provide stability estimates in both cases. Finally, in Section V, we present a penalized minimization algorithm that makes it possible to estimate the source location and medium parameters.

II. GENERAL FRAMEWORK

A. Acoustic waves propagation

Let us model the shallow water environment by a Pekeris waveguide $\mathbb{R}^2 \times (0, +\infty)$ represented in Figure 1 and formed with a thin layer of ocean for depths between 0 and D and an infinite layer of sediment for depths greater than D (Pekeris, 1948). The celerity c and the density ρ are assumed to be piecewise constant, and an impulsive source transmits a signal received by a receiver (hydrophone) located at range r . The case of non-impulsive signals can be handled similarly using deconvolution techniques or dispersion curve differences described in (Bonnell *et al.*, 2020). The acoustic wave u emitted by the impulsive source satisfies the acoustic

wave equation

$$\rho \nabla \cdot \left(\frac{1}{\rho} \nabla u \right) - \frac{1}{c^2} \partial_{tt} u = 0, \quad (2)$$

with the free surface (Dirichlet) boundary condition $u|_{z=0} = 0$. Here ρ is the density and c is the wave speed. Following the analysis presented in (Jensen *et al.*, 2011, §2.4.5), the Fourier transform $\hat{u}(\omega) = \int_{\mathbb{R}} u(t) \exp(-i\omega t) dt$ of the signal $u(t)$ recorded by the receiver can be decomposed as a sum of modal components

$$\hat{u}(\omega) = \sum_{n \in \mathbb{N}} \hat{u}_n(\omega) = \sum_{n \in \mathbb{N}} A_n(\omega) e^{i\Phi_n(\omega)}, \quad (3)$$

where the complex amplitude A_n is slowly varying in ω and depends on the shape of the emitted signal and the depths and ranges of the source and receiver (Jensen *et al.*, 2011) and the phase is of the form $\Phi_n(\omega) = -k_n(\omega)r$, where r is the distance from the source to the receiver (range) and k_n is the n -th wavenumber given by the n -th solution of the dispersion relation

$$\tan \left(D \sqrt{\omega^2/c_w^2} \right) = -\frac{\rho_b}{\rho_w} \sqrt{\frac{\omega^2/c_w^2 - k_n^2}{k_n^2 - \omega^2/c_b^2}}. \quad (4)$$

Given the little information that we have, the general idea is to use $\Phi_n(\omega)$ and (4) to recover information about r and the parameters of the Pekeris waveguide. However, phrase retrieval is difficult with noisy data (see however (Bonnell *et al.*, 2012)), and it is easier to retrieve the modal travel time $t_n(\omega)$ defined by

$$t_n(\omega) = -\Phi'_n(\omega) = r k'_n(\omega). \quad (5)$$

In Figure 2, we plot an example of the dispersion curves $\omega \mapsto t_n(\omega)$. The time-frequency representation (Boashash, 2016) of signals is known to concentrate energy around these dispersion curves, which enables to recover r and the parameters of the Pekeris waveguide. Following the investigation led in (Bonnell, 2010) on the most appropriate time-frequency representation, we choose to use the short-time Fourier transform (STFT) of the signal and the spectrogram S defined by

$$S(t, \omega) = |STFT(t, \omega)|^2, \quad (6)$$

$$STFT(t, \omega) = \int_{\mathbb{R}} u(\tau) h(\tau - t) e^{-i\omega\tau} d\tau \quad (7)$$

$$= \frac{e^{i\omega t}}{2\pi} \int_{\mathbb{R}} \hat{u}(\xi + \omega) \hat{h}(\xi) e^{i\xi t} d\xi. \quad (8)$$

Similarly, we define the modal component S_n of the spectrogram S by

$$S_n(t, \omega) = \left| \int_{\mathbb{R}} u_n(\tau) h(\tau - t) e^{-i\omega\tau} d\tau \right|^2 \quad (9)$$

$$= \frac{1}{4\pi^2} \left| \int_{\mathbb{R}} \hat{u}_n(\xi + \omega) \hat{h}(\xi) e^{i\xi t} d\xi \right|^2. \quad (10)$$

In this definition, h is a centered window function. Different choices of h are possible (rectangular windows,

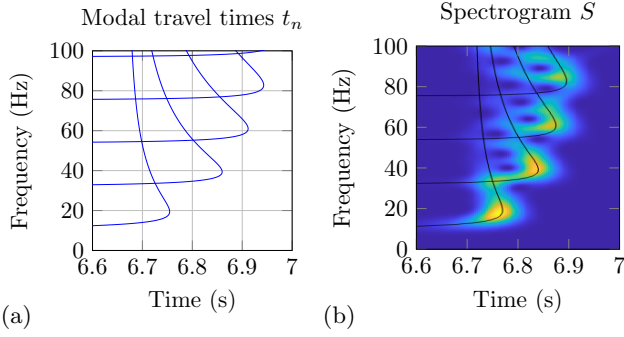


FIG. 2. (a): Example of dispersion curves $\omega \mapsto t_n(\omega)$ in the Pekeris model with parameters given in Table I-line 1. (b): Spectrogram S computed with a Gaussian window defined in (11) with $\sigma = 20$ Hz, superposed with the dispersion curves.

sine windows, Hamming windows...), and in the following, we choose to work with a Gaussian window of standard deviation σ to simplify the computations. The function h is then defined by

$$h(t) = \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\sigma^2 t^2}{2}\right), \quad \hat{h}(\omega) = \exp\left(-\frac{\omega^2}{2\sigma^2}\right). \quad (11)$$

An example of a spectrogram with $\sigma = 20$ Hz is presented in Figure 1. As explained in (Bonnell *et al.*, 2020), if σ is small enough, the support of \hat{h} is small and we can hope to simplify the expression of S . The general idea, detailed in the following sections, first uses the fact that modes can be separated and that one can use S to recover the modal spectrograms S_n that are of the form

$$S_n(t, \omega) = \frac{1}{4\pi^2} \left| \int_{\mathbb{R}} A_n(\xi + \omega) e^{i\Phi_n(\xi + \omega)} \hat{h}(\xi) e^{i\xi t} d\xi \right|^2. \quad (12)$$

Then, using the fact that the support of \hat{h} is small,

$$A_n(\xi + \omega) \approx A_n(\omega), \quad (13)$$

$$\Phi_n(\xi + \omega) \approx \Phi_n(\omega) + \xi \Phi'_n(\omega) = \Phi_n(\omega) - \xi t_n(\omega), \quad (14)$$

and we can find a direct link between $S_n(t, \omega)$ and $t_n(\omega)$:

$$S_n(t, \omega) \approx \frac{1}{4\pi^2} |A_n(\omega)|^2 \left| \int_{\mathbb{R}} \hat{h}(\xi) e^{i\xi(t - t_n(\omega))} d\xi \right|^2 \quad (15)$$

$$= |A_n(\omega)|^2 h(t - t_n(\omega))^2. \quad (16)$$

We will detail these approximations in the following, as well as the use of spectrograms to recover the modal travel times and the environment parameters. Throughout the paper, we will illustrate our method on synthetically generated data and on experimental data presented in the following subsection.

B. Synthetical and experimental data

In this paper, all numerical simulations on synthetically generated data are presented for a Pekeris model

whose parameters are given in the first line of Table I. These data are generated on Matlab using the code developed by J. Bonnel (Bonnell *et al.*, 2020) where the Fourier transform \hat{u} is computed using the modal expansion presented in (Jensen *et al.*, 2011), and then moved into the time domain using an inverse fast Fourier transform. Spectrograms are calculated using the time-frequency toolbox (F. Auger and Flandrin, 2008), and enable to recover four different propagative modes in the modal decomposition.

We also implement our algorithm on experimental data from two different experiments. The first one is the recording of a right whale gunshot presented in (Thode *et al.*, 2017), where four propagative modes are recorded with a maximal frequency $f_{\max} = 200$ Hz. In this environment, we have some prior estimates of parameters given in the second line of Table I. The second one is the recording of a combustive sound source (CSS) presented in (Bonnell *et al.*, 2018), where 11 modes are clearly visible and 7 additional modes are hardly visible. In this environment, we have some prior estimates of parameters given in the third line of Table I.

III. EXTRACTION OF THE MODAL COMPONENTS OF THE SPECTROGRAM

This section aims to extract the modal components S_n of the spectrogram from the full spectrogram S . We use the warping method described in (Bonnell *et al.*, 2020). The general idea of this method is to apply a change of variable in the time axis to improve the separation of the modal components in the spectrogram. We first provide some theoretical tools to assess the mode separability. Then, we describe a way to implement this method numerically in a completely automated way.

A. Warping method

Let us denote $\omega \mapsto t_n(\omega)$ the modal travel time and $t \mapsto \omega_n(t)$ its reciprocal function. Given the fact that most of the energy of the Gaussian window h (resp. \hat{h}) is contained in the interval $(-\sigma^{-1}, \sigma^{-1})$ (resp. $(-\sigma, \sigma)$), the energy of each modal component of the spectrogram is concentrated in the rectangle

$$(\min(t_n) - \sigma^{-1}, \max(t_n) + \sigma^{-1}) \times (\min(\omega_n) - \sigma, \max(\omega_n) + \sigma). \quad (17)$$

If the intersection of these rectangles for different modes is empty, we say that the modes are well separated and we can easily extract the different modal components of the spectrogram. However, these rectangles often intersect, making this separation impossible. Hence the use of warping methods.

The general idea of the warping method is to notice that in a perfectly reflecting waveguide with a boundary condition $\partial_z u|_{z=D} = 0$ at depth $z = D$, dispersion curves

	r (km)	D (m)	c_w (m/s)	c_b (m/s)	ρ_w (kg/m ³)	ρ_b (kg/m ³)	f_{\max} (Hz)
Pekeris model	10	100	1500	1600	1000	1500	100
Right whale (Thode <i>et al.</i> , 2017)	8.8	51	1450	?	1000	1600	204
CSS (Bonnel <i>et al.</i> , 2018)	4.8	69.5	1464.5	?	1000	1600	488

TABLE I. Numerical values for parameters in the Pekeris waveguide. Values are exact for the Pekeris model (line 1) and estimated for the right whale gunshot (line 2) and the combustive sound source (line 3).

are explicit and given by

$$\omega_n(t) = \frac{tc_w\pi(2n-1)}{2D\sqrt{t^2-r^2/c_w^2}} \quad (18)$$

when the emission of the pulse is at time 0. To improve the separability of modal components, one can change the time variable with a change of variable $t \mapsto \psi(t)$ and consider the spectrogram of the signal

$$\tilde{u}(t) = \sqrt{|\psi(t)|} u \circ \psi(t). \quad (19)$$

The dispersion curves of the transformed signal are linked to the ones of the original signal through the relation $\tilde{\omega}_n(t) = \psi'(t)\omega_n(\psi(t))$, which enables to compute the new dispersion curves:

$$\tilde{\omega}_n(t) = \psi'(t) \frac{\psi(t)c_w\pi(2n-1)}{2D\sqrt{\psi(t)^2-r^2/c_w^2}}. \quad (20)$$

To simplify these new curves, choosing $\psi(t) = \sqrt{t^2+r^2/c_w^2}$ is natural (assuming that the range r is known or can be estimated). This change of variable leads to constant dispersion curves

$$\tilde{\omega}_n(t) = \frac{c_w\pi(2n-1)}{2D} \quad (21)$$

which are separable using a small value of σ .

In the Pekeris waveguide, the warping is more complicated. Given the similarities between Pekeris waveguides and perfectly reflecting waveguides, we choose to use the same type of warping and we define $\psi(t) = \sqrt{t^2+t_0^2}$ for a certain value $t_0 > 0$. The new dispersion curves are then illustrated in Figure 3 and given by

$$\tilde{\omega}_n(t) = \frac{t}{\sqrt{t^2+t_0^2}} \omega_n \left(\sqrt{t^2+t_0^2} \right). \quad (22)$$

Since dispersion curves intersect, it is impossible to find a change of variable that completely separates the curves as in the perfectly reflecting waveguide. Hence, we now only aim at separating parts of dispersion curves that do not overlap (see an illustration in Figure 3) and contain most of the modal energy. These parts are defined as follows.

Definition 1. For each curve $\omega \mapsto t_n(\omega)$, we define $T_n^{(2)} = \max_{\omega}(t_n(\omega))$ and $\Omega_n^{(2)} = \operatorname{argmax}_{\omega}(t_n(\omega))$ and we consider portions of the dispersion curves contained in an interval $(\Omega_n^{(1)}, \Omega_n^{(3)})$ such that $\Omega_n^{(1)} < \Omega_n^{(2)} < \Omega_n^{(3)}$. We

denote $T_n^{(j)} = t_n(\Omega_n^{(j)})$ for $j = 1, 2, 3$. The energy rectangle associated to the mode number n is then defined by

$$E_n(\sigma) := (\min(T_n^{(1)}, T_n^{(3)}) - \sigma^{-1}, T_n^{(2)} + \sigma^{-1}) \times (\Omega_n^{(1)} - \sigma, \Omega_n^{(3)} + \sigma), \quad (23)$$

and we consider that two modal components n and m are separated if there exists $\sigma > 0$ such that

$$E_n(\sigma) \cap E_m(\sigma) = \emptyset. \quad (24)$$

An illustration is provided in Figure 3. Figures 3b show that, no matter the choice of σ , modal components cannot be separated, making warping necessary. Figures 3c-d-e show the spectrogram and the dispersion curves of the warped signal.

The following proposition quantifies the separability of modal components:

Proposition 1. Given two portions of dispersion curves associated to modes $n < m$ that do not overlap, there exists a maximum value Ω_{\max} such that if $\Omega_m^{(3)} < \Omega_{\max}$, then there exists a warping and parameters $t_0 > 0, \sigma > 0$ that separate the two modal components n and m .

The proof is provided in supplementary. This proposition shows that if $\Omega_m^{(3)}$ is not too large, then portions of modal dispersion curves can be separated with an appropriate choice of t_0 and σ . As explained in the proof, the parameter σ needs to be chosen as large as possible in order to separate modal components in the vertical direction (see an illustration in Figure 3-e). Regarding the choice of t_0 , the proof does not provide a concrete way to compute it. As illustrated in Figures 3d-e, a small change in t_0 may produce very different dispersion curves. We give in the next section a method to choose t_0 in an automated and appropriate way.

Let us also point out that in all the numerical and experimental examples shown below, Ω_{\max} turns out to be much larger than the maximum accessible frequencies, ensuring that $\Omega_m^{(3)} < \Omega_{\max}$. Furthermore, one can always ensure that parts of curves do not overlap by considering smaller portions of the dispersion curves and keeping only the high energy levels in the spectrogram.

B. Numerical implementation

To numerically implement the previous warping method, one must find a good value of t_0 and separate

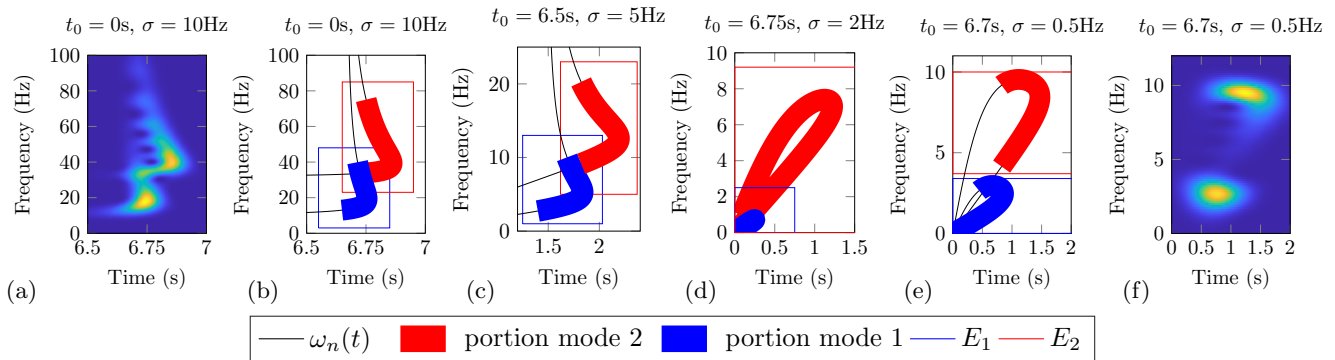


FIG. 3. Illustration of the warping of mode 1-2 in the Pekeris model for different choices of t_0 . (a) initial spectrogram computed with $\sigma = 10$ Hz. (b) simplified diagram with a portion of mode 1 and 2. (c) new diagram with the warping $\psi(t) = \sqrt{t^2 + t_0^2}$ with $t_0 = 6.5$ s. (d) new diagram with $t_0 = 6.75$ s. (e) new diagram with $t_0 = 6.7$ s. (f) new spectrogram of the warped signal at $t_0 = 6.7$ s, computed with $\sigma = 0.5$ Hz.

modes in the spectrogram. In this subsection we develop an algorithm that iteratively separates each mode, using the smallest possible value for σ and starting with the largest propagative mode number N . Assuming that t_0 is well chosen, each modal component is separated from the others in the spectrogram. To separate them automatically, we use a watershed transform which returns a set of drainage basins associated with local maximum points of the spectrogram in the time-frequency space (t, ω) . The watershed transform is an algorithm used for image processing, segmentation, and analysis. It is based on the concept of a topographic map, where the grey level values of an image represent elevations, and the image is viewed as a surface. The basic idea behind the watershed transform is to use the topographic map of an image to identify regions or objects within the image. This is done by “flooding” the topographic map from its minima, which correspond to local minima in the image. As the flooding proceeds, the water from different minima starts to meet and form ridges or boundaries, which separate different regions in the image. More details about this method can be found for instance in (Meyer, 1994).

Using the shape of dispersion curves in Pekeris waveguides (see for instance Figure 3), we know that most of the energy of each mode is concentrated near its high frequencies. Therefore, we choose to apply the following algorithm to sort drainage basins:

- We sort local maximum points (t_i, ω_i) associated with basins B_i by descending frequencies.
- We associate B_1 with the mode number N . For $i \geq 1$, if $S(t_{i+1}, \omega_{i+1}) < S(t_i, \omega_i)$, then we keep the same mode number, otherwise, we associate B_{i+1} to a new mode number (the mode number of B_i minus one).

An example is shown in Figure 4.

We still need to choose a good t_0 to separate modes. To do so, we maximize a quality factor defined to favor warpings where each mode is separated from the next one, both in amplitude and frequency distance.

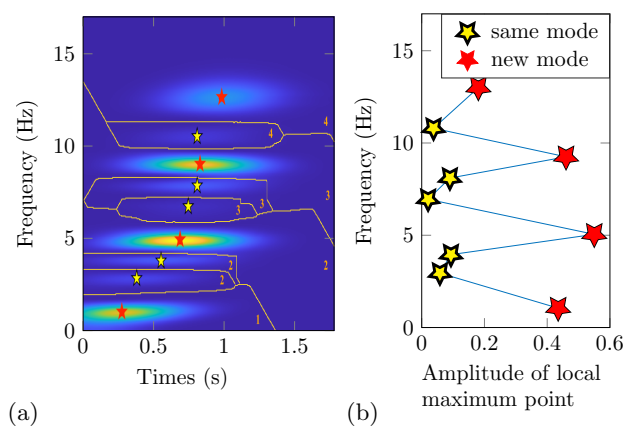


FIG. 4. (a): A warped spectrogram superposed with associated drainage basins and local maximum points (yellow and red stars). (b): Amplitude of each local maximum point sorted by descending frequencies; the top red star is associated with mode number 4, and so is the yellow star below it. The second red star from the top is associated to the mode number 3, as well as the two following yellow stars, and so on.

We privilege situations where each mode contains different drainage basins with an amplitude gap with the next mode. For a mode number n associated to basins $i = I, \dots, J$, we define

$$Q^{(n)}(t_0) = \begin{cases} 0 & \text{if } I = J, \\ S(t_{J+1}, \omega_{J+1}) - S(t_J, \omega_J) & \text{else.} \end{cases} \quad (25)$$

This quality factor is relatively simple and it encapsulates the warping quality very well as shown in Figure 5. Its maximization gives the best t_0 to separate modes. If $Q^{(n)}$ vanishes everywhere, we remove the first condition and only measure the amplitude gaps between the basins.

After choosing the best t_0 to maximize the quality factor $Q^{(N)}$, we remove all the drainage basins associated with the mode N , and we do the same process for the

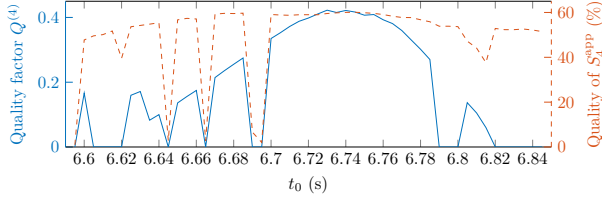


FIG. 5. Comparison between the quality factor $Q^{(4)}$ and the quality of the reconstruction of S_4 , with respect to t_0 . The quality of the precision is computed using the formula $1 - \|S_4 - S_4^{\text{app}}\|_2 / \|S_4\|_2$, where S_4 is the true spectrogram of the mode 4, S_4^{app} is the one obtained after modal separation, and $\|\cdot\|_2$ is the L^2 -norm.

next mode $N - 1$ with the quality factor $Q^{(N-1)}$ until the spectrogram is empty. The fully automated algorithm is summarized in Algorithm 1.

ALGORITHM 1: Filtering of modal components in the received signal

Input: Signal u measured on the regular grid
 $\tau = (\tau_1, \dots, \tau_T)$
Number of propagative modes N

Output: Modal components u_1, u_2, \dots, u_N

for $n = N, \dots, 2$ **do**
 for $\ell = 1, \dots, T$ **do**
 $t_0 = \tau_\ell$
 $\tilde{u} \leftarrow$ warping of u associated to $\psi(t) = \sqrt{t^2 + t_0^2}$
 (see (19))
 $\tilde{S} \leftarrow$ spectrogram of \tilde{u} computed with the smallest possible value of σ (see (9) and (11))
 $B_i, (t_i, \omega_i) \leftarrow$ drainage basins of \tilde{S} , local maxima of \tilde{S} , sorted by descending frequencies
 $Q^{(n)}(\ell) \leftarrow$ quality factor at t_0 computed using (B_i, t_i, ω_i) (see (25))
 end for
 $t_0 \leftarrow \text{argmax}(Q^{(n)})$
 $B_i, (t_i, \omega_i) \leftarrow$ drainage basins of \tilde{S} at t_0 , local maxima of \tilde{S} at t_0
 mask $\leftarrow 0 \times \tilde{S}$
 while $\tilde{S}(t_{i+1}, \omega_{i+1}) < \tilde{S}(t_i, \omega_i)$ **do**
 mask = mask + $\mathbf{1}_{B_i}$ (indicator function of B_i)
 end while
 $u_n \leftarrow$ inverse warping of the inverse spectrogram of $\tilde{S} \times \text{mask}$
 $u \leftarrow u - u_n$
end for
 $u_1 \leftarrow u$

A modal separation done with this algorithm on synthetic data in the Pekeris model is presented in Figure 6. Then, this algorithm is applied to experimental data with a right whale gunshot (Figure 7) and a combustive sound source (Figure 8) and produces great reconstruction comparable to the ones proposed in (Bonnell *et al.*, 2018, 2020; Thode *et al.*, 2017) but in a completely automated way.

From now on, we assume that we can separate modal components and get good approximations of each modal spectrogram $S_n(t, \omega)$. In the following section, we will discuss the best action to use these spectrograms to get approximations of dispersion curves $\omega \mapsto t_n(\omega)$.

IV. RECONSTRUCTION OF THE MODAL TRAVEL TIMES

In this section, for a fixed n , we aim at approaching the dispersion curves $\omega \mapsto t_n(\omega)$ given the modal spectrogram $S_n(t, \omega)$, and we consider a signal of the form

$$u(t) = u_n(t) + W(t), \quad u_n(t) = \frac{1}{2\pi} \int \hat{u}_n(\omega) e^{i\omega t} d\omega, \quad (26)$$

where $W(t)$ is a Gaussian additive noise satisfying $\mathbb{E}(W(t)) = 0$ and

$$\text{Cov}(W(t), W(t')) = \delta^2 \exp\left(-\frac{(t-t')^2}{2T_\delta^2}\right). \quad (27)$$

This noise model seems to be a good approximation of the measurement noise and the environment noise (Aparicio *et al.*, 2015), and can also account for the possible small errors in the previous separation method. As mentioned in Section II, we have a link between the spectrogram $S(t, \omega)$ and the dispersion curves $\omega \mapsto t_n(\omega)$ since

$$S(t, \omega) \approx S^{\text{app}}(t, \omega) := |A_n(\omega)|^2 h(t - t_n(\omega))^2. \quad (28)$$

Two competing methods would then allow to recover $t_n(\omega)$ using $S(t, \omega)$:

- Maximum method: we notice that

$$\begin{aligned} \partial_t \sqrt{S^{\text{app}}(t, \omega)} = 0 &\Leftrightarrow \partial_t h(t - t_n(\omega)) = 0 \\ &\Leftrightarrow t = t_n(\omega). \end{aligned} \quad (29)$$

To recover an approximated value of $t_n(\omega)$, we look for the maximum value of $t \mapsto S^{\text{app}}(t, \omega)$ by solving the equation $\partial_t \sqrt{S^{\text{app}}(t, \omega)} = 0$.

- Mean method: since h is a centered window function,

$$\begin{aligned} \frac{\int_{\mathbb{R}} t S^{\text{app}}(t, \omega) dt}{\int_{\mathbb{R}} S^{\text{app}}(t, \omega) dt} &= \frac{\int_{\mathbb{R}} t h(t)^2 dt + t_n(\omega) \int_{\mathbb{R}} h(t)^2 dt}{\int_{\mathbb{R}} h(t)^2 dt} \\ &= t_n(\omega). \end{aligned} \quad (30)$$

To recover an approximated value of $t_n(\omega)$, we can compute the quantity

$$\frac{\int_{\mathbb{R}} t \phi(t) S^{\text{app}}(t, \omega) dt}{\int_{\mathbb{R}} \phi(t) S^{\text{app}}(t, \omega) dt}, \quad \phi(t) = \exp\left(-\frac{t^2}{2T_w^2}\right), \quad (31)$$

where ϕ is a weight function to account for the fact that we do not have access to measurements for every $t \in \mathbb{R}$.

To discriminate these two methods, we study the stability of each method and the reconstruction error given the noise level. In each case, we present the reconstruction error for a fixed window width σ , and then we optimize the width σ to minimize the reconstruction error.

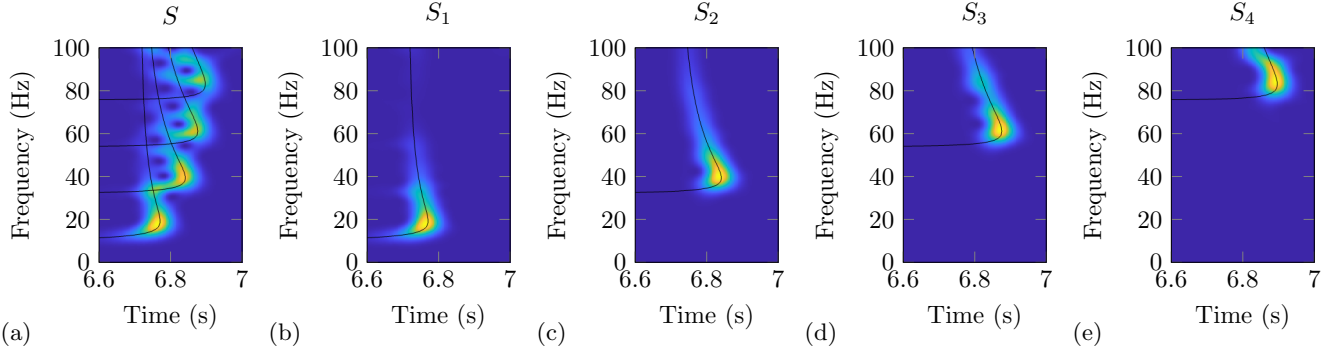


FIG. 6. Initial spectrogram and separate modal components computed with $\sigma = 20$ Hz and obtained with the warping algorithm on synthetically generated data in the Pekeris model given in Table I-line 1. (a) initial spectrogram superposed with the dispersion curves. (b) spectrogram of mode 1. (c) spectrogram of mode 2. (d) spectrogram of mode 3. (e) spectrogram of mode 4.

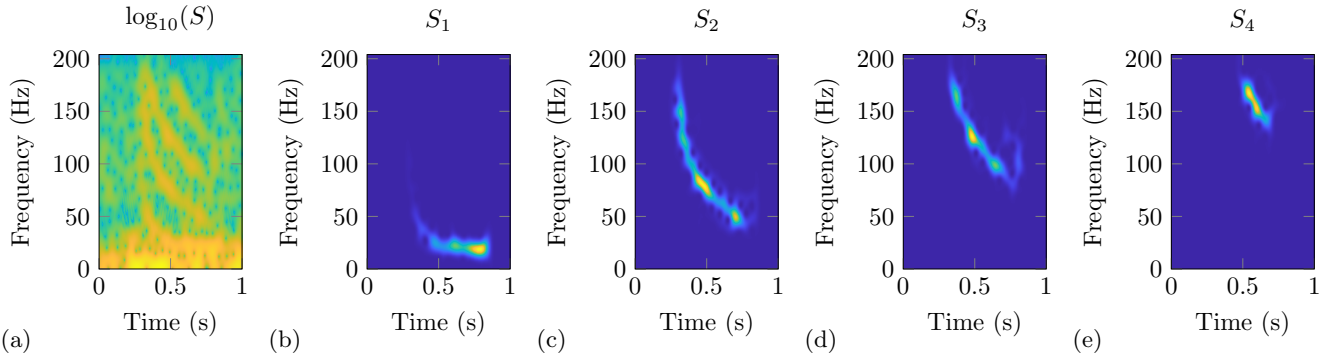


FIG. 7. Initial spectrogram and separate modal components computed with $\sigma = 41$ Hz and obtained with the warping algorithm on data of a right whale gunshot (Thode *et al.*, 2017). (a) initial spectrogram (in log scale). (b) spectrogram of mode 1. (c) spectrogram of mode 2. (d) spectrogram of mode 3. (e) spectrogram of mode 4.

A. Maximum method

First, we investigate the stability of the maximum method. In the following, for each function ζ_n depending on the mode n , we denote

$$\|\zeta_n\|_\infty = \sup_{\tilde{\omega} \in (\omega_{c,n}, +\infty)} |\zeta_n(\tilde{\omega})|, \quad (32)$$

where $\omega_{c,n}$ is the cut-off frequency of the mode n (Jensen *et al.*, 2011, §2.4.5.1).

Proposition 2. Let $t_n^{\text{app}}(\omega)$ be an approximation of $t_n(\omega)$ defined as the unique solution of the equation $\partial_t \sqrt{S_n(t, \omega)} = 0$. Defining $\sigma_{\text{lim}}(\omega) = \omega_{c,n}/3 - \omega/4$ (which do not depend on δ , T_δ , σ), we have

$$\begin{aligned} & \mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|) \\ & \leq \frac{\sup_{\tilde{\omega} \in (4\omega_{c,n}/3, +\infty)} |A'_n(\tilde{\omega})|}{|A_n(\omega)|} + \frac{\|A'_n\|_\infty}{|A_n(\omega)|} \mathbf{1}_{\sigma > \sigma_{\text{lim}}(\omega)} \\ & \quad + C_2 \sigma \frac{\|A'_n\|_\infty \|\Phi'_n\|_\infty}{|A_n(\omega)|} + C_3 \frac{\delta T_\delta^{1/2}}{|A_n(\omega)| \sigma^{3/2}}. \end{aligned} \quad (33)$$

This quantity is minimal when $\sigma = \sigma_{\text{opt}}$, where

$$\sigma_{\text{opt}}(\omega) \in \left\{ \sigma_{\text{lim}}(\omega), \left(\frac{3\delta T_\delta^{1/2} 2^{-11/4}}{\|A_n\|_\infty \|\Phi'_n\|_\infty} \right)^{2/5} \right\}. \quad (34)$$

Here, all dimensionless quantities C_i are explicit constants detailed in supplementary.

Due to the term $\mathbf{1}_{\sigma > \sigma_{\text{lim}}(\omega)}$ in the control of the reconstruction error, the optimal value of σ cannot be computed easily and one needs to test the two values in (34) and determine each time which one provides the minimal error. However, in the Pekeris model, the most common situation seems to be the one where $\sigma_{\text{opt}} = \sigma_{\text{lim}}$, as illustrated in Figure 9 where we compute the optimal choice of σ for different values of $\delta T_\delta^{1/2}$ and ω . We also compute the associated reconstruction error $\mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|)$.

B. Mean method

We now investigate the stability of the mean method.

Proposition 3. Let $t_n^{\text{app}}(\omega)$ be an approximation of $t_n(\omega)$ defined by

$$t_n^{\text{app}}(\omega) = \frac{\int_{\mathbb{R}} t \phi(t) S(t, \omega) dt}{\int_{\mathbb{R}} \phi(t) S(t, \omega) dt}, \quad \phi(t) = \exp\left(-\frac{t^2}{2T_w^2}\right). \quad (35)$$

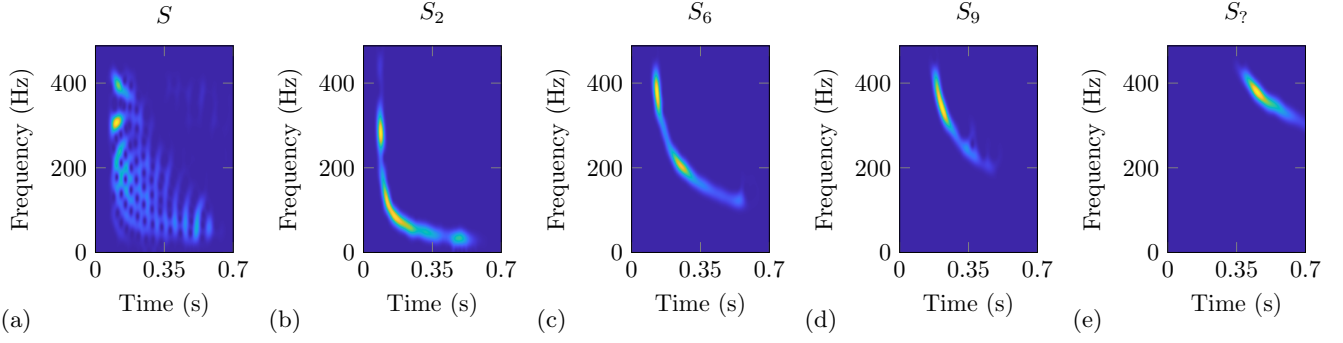


FIG. 8. Initial spectrogram and separate modal components computed with $\sigma = 98$ Hz and obtained with the warping algorithm on data of a combustive sound source (Bonnell *et al.*, 2018). As in (Bonnell *et al.*, 2018), we recover the 11 first modes, but due to a gap in the data, we cannot be sure of the mode numbers after that. We recover a total of 14 modal components using our warping algorithm. (a) initial spectrogram. (b) spectrogram of mode 2. (c) spectrogram of mode 6. (d) spectrogram of mode 9. (e) spectrogram of an unidentified mode.

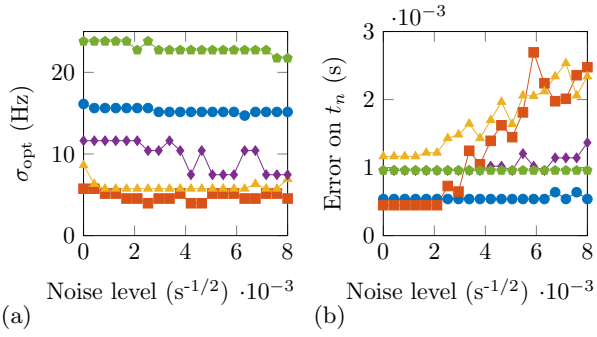


FIG. 9. (a) Optimal choice of σ to minimize the error of reconstruction of t_n for increasing values of the noise level and different frequencies f . The noise level is defined as $\delta T_\delta^{1/2}/|A_n(\omega)|$. (b) Mean reconstruction error $\mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|)$ for the optimal choice $\sigma = \sigma_{\text{opt}}$ with respect to the noise level $\delta T_\delta^{1/2}/|A_n(\omega)|$. The mean is computed using 50 simulations, and $T_\delta = 0.01s$.

If $\sigma t_n(\omega) > 1$ and $\delta T_\delta^{1/2}/(t_n(\omega)^{1/2}|A_n(\omega)|) \ll 1$ then

$$\mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|) \leq \frac{\|A'_n\|_\infty \|A_n\|_\infty}{4|A_n(\omega)|^2} + D_1 \frac{t_n(\omega)^2}{T_w} + D_2 \frac{\|A'_n\|_\infty^2 t_n(\omega)}{|A_n(\omega)|^2} \sigma^2 + D_3 t_n(\omega) \frac{\delta T_\delta^{1/2}}{\sqrt{\sigma}|A_n(\omega)|}. \quad (36)$$

This quantity is minimal for

$$\sigma_{\text{opt}}(\omega) = D_4 \frac{|A_n(\omega)|^{4/5}}{\|A'_n\|_\infty^{4/5}} \left(\frac{\delta T_\delta^{1/2}}{|A_n(\omega)|} \right)^{2/5}, \quad (37)$$

and provides an error

$$\mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|) \leq \frac{\|A'_n\|_\infty \|A_n\|_\infty}{4|A_n(\omega)|^2} + D_1 \frac{t_n(\omega)^2}{T_w} + D_5 \frac{t_n(\omega) \|A'_n\|_\infty^{2/5}}{|A_n(\omega)|^{2/5}} \left(\frac{\delta T_\delta^{1/2}}{|A_n(\omega)|} \right)^{4/5}. \quad (38)$$

Here, all dimensionless quantities D_i are explicit and depend on ω , T_w , A_n and Φ_n . Their expressions are given in supplementary.

We illustrate these results in Figure 10 where we plot $\sigma_{\text{opt}}(\omega)$ with respect to $\delta T_\delta^{1/2}/|A_n(\omega)|$. Numerically, we notice that for each plotted frequency, we have

$$\sigma_{\text{opt}}(\omega) = \mathcal{O} \left(\left(\frac{\delta T_\delta^{1/2}}{|A_n(\omega)|} \right)^{0.38} \right), \quad (39)$$

$$\mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|) = \mathcal{O} \left(\left(\frac{\delta T_\delta^{1/2}}{|A_n(\omega)|} \right)^{0.73} \right), \quad (40)$$

which corroborates the theoretical results of Proposition 3.

C. Comparison between both methods

We can now compare the maximum and the mean methods using Propositions 2 and 3. Let us denote by e_{max} the error of reconstruction with the maximum method and by e_{mean} the error with the mean method. First, we notice that, when there is no noise in the data (i.e., when $\delta T_\delta^{1/2} = 0$), and assuming that T_w is large enough, we have

$$e_{\text{max}} \leq \frac{\|A'_n\|_\infty}{|A_n(\omega)|}, \quad e_{\text{mean}} \leq \frac{\|A'_n\|_\infty \|A_n\|_\infty}{4|A_n(\omega)|^2}. \quad (41)$$

When $|A_n(\omega)| \approx \|A_n\|_\infty$ and without noise, the mean method turns out to be more precise.

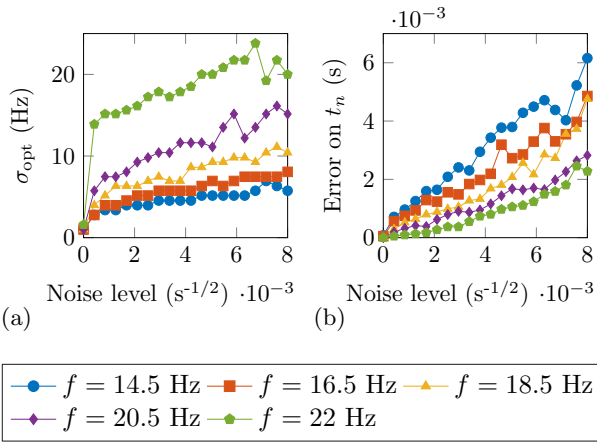


FIG. 10. (a) Optimal choice of σ to minimize the error of reconstruction of t_n for increasing values of the noise level and different frequencies f . The noise level is defined as $\delta T_\delta^{1/2}/|A_n(\omega)|$. (b) Mean reconstruction error $\mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|)$ for the optimal choice $\sigma = \sigma_{\text{opt}}$ with respect to the noise level $\delta T_\delta^{1/2}/|A_n(\omega)|$. The mean is computed using 50 simulations, and $T_\delta = 0.01s$.

Assuming that $\sigma_{\text{opt}} = \sigma_{\text{lim}}$ for the maximum method when data are noisy, we notice that e_{max} grows as $\delta T_\delta^{1/2}/|A_n(\omega)|$ while e_{mean} increases as $(\delta T_\delta^{1/2}/|A_n(\omega)|)^{4/5}$, meaning that the maximum method becomes more precise as the noise level grows. Numerical tests for different frequencies are presented in Figure 11 and confirm this comparison. We give the relative reconstruction error at the optimal choice $\sigma = \sigma_{\text{opt}}$ with the maximum and the mean methods for different frequencies in each case.

We then investigate the behavior of the global error

$$\max_{\omega \in \Omega_n} \mathbb{E}(|t_n^{\text{app}}(\omega) - t_n(\omega)|), \quad (42)$$

where Ω_n is the set of frequencies

$$\Omega_n = \{\omega > 0 \text{ s.t. } S_n(t_n^{\text{app}}(\omega), \omega) > p \max(S)\}, \quad (43)$$

for a fixed threshold p (its choice is discussed in the next section). This definition is very similar to the one presented in (Bonnell *et al.*, 2020), where one must select the significant parts of the dispersion curves $\omega \mapsto t_n^{\text{app}}(\omega)$ manually. We plot in Figure 12 the optimal value σ_{opt} to minimize the global error and the resulting global error for both methods. Again, the mean method works better for a small noise level and the maximum method for a higher noise level. We also notice that the optimal value σ_{opt} in the maximum method is very stable with respect to $\delta T_\delta^{1/2}/|A_n(\omega)|$, which makes it easy to calibrate a priori.

Given the theoretical and numerical results presented in this section, we choose to work with the optimized maximum method to recover an approximated value of the dispersion curves. This choice differs from the one

used in paper (Bonnell *et al.*, 2020) where the value $\sigma = 20$ Hz is always chosen, and the mean method is used to compute the approximation of t_n . These non-optimized choices are represented in Figure 12 for comparison purposes.

Using the maximum method with an optimized choice $\sigma_{\text{opt}} = 5f_{\text{max}}/100$, we present in Figure 13 the approximated reconstructions $t_n^{\text{app}}(\omega)$ obtained using Algorithm 2 and the separation of modes presented in Algorithm 1.

ALGORITHM 2: Reconstruction of dispersion curves

Input: Modal component u_n sampled at frequency f_{max}
Set of frequencies F used to compute the spectrogram
Threshold $P = p \max(S)$
Output: Dispersion curves $F \mapsto t_n^{\text{app}}(F)$
 $S_n \leftarrow$ spectrogram of u_n computed with $\sigma = 5f_{\text{max}}/100$ (see (9) and (11))
for $f \in F$ **do**
 $t_n^{\text{app}}(f) = \arg\max S_n(f, \cdot)$
if $S_n(t_n^{\text{app}}(f), f) < P$ **then**
 $t_n^{\text{app}}(f) = \text{NaN}$ {One could also remove the values from the list}
end if
end for

V. INVERSE PROBLEM

We now work with approximated values of the dispersion curves $\omega \mapsto t_n^{\text{app}}(\omega)$ obtained using the method described in the previous section, and aim at recovering the range r and the parameters of the Pekeris waveguide. Dispersion curves $\omega \mapsto t_n(\omega)$ do not have an explicit expression, but it is possible to compute excellent approximations by looking at the zeros of the relation (4). As described in (Bonnell *et al.*, 2020), the most naive method to recover all the parameters would be to minimize the functional

$$J(r, c_w, c_b, \rho_w, \rho_b, D, dt) = \|t_n(r, c_w, c_b, \rho_w, \rho_b, D, \cdot) - dt - t_n^{\text{app}}(\cdot)\|_{\ell^2}^2, \quad (44)$$

where

$$\|t_n\|_{\ell^2}^2 = \sum_{n=1}^N \int_{\Omega_n} t_n(\omega)^2 d\omega, \quad (45)$$

and Ω_n is the set of frequencies associated to the mode n defined in (43).

However, different difficulties arise from this approach:

- Firstly, all parts of curves $\omega \mapsto t_n^{\text{app}}(\omega)$ may not be significant and one must choose an appropriate threshold p to define Ω_n in (43). Especially we prove in the proofs of Propositions 2 and 3 that the reconstruction error increases as $1/|A_n(\omega)|$, which leads to large errors for frequencies with low energy.

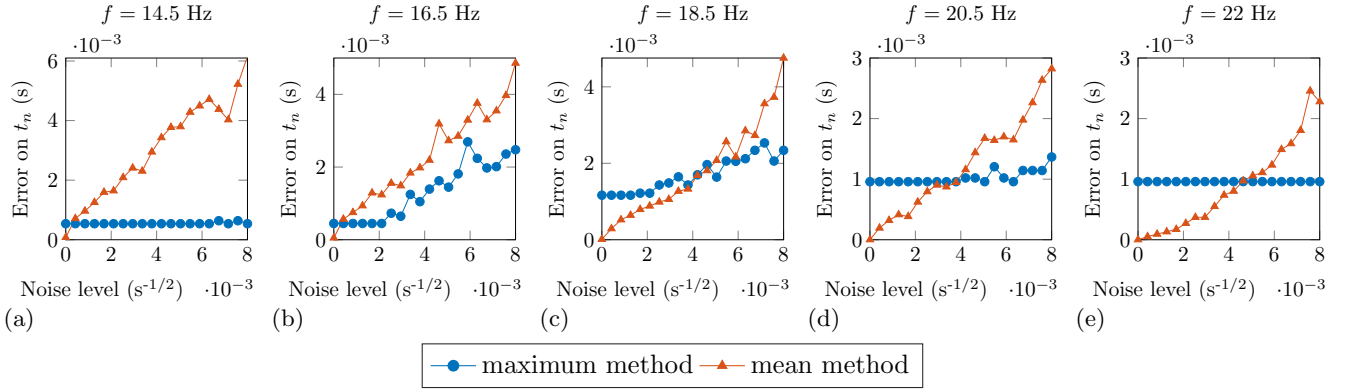


FIG. 11. Reconstruction error $\mathbb{E}(|t_n(\omega) - t_n^{\text{app}}(\omega)|)(\sigma_{\text{opt}})$ with respect to the noise level $\delta T_\delta^{1/2}/|A_n(\omega)|$ for different frequencies. (a) $f = 14.5$ Hz (b) $f = 16.5$ Hz (c) $f = 18.5$ Hz (d) $f = 20.5$ Hz (e) $f = 22$ Hz. The mean is computed with 50 simulations, and $T_\delta = 0.01$ s

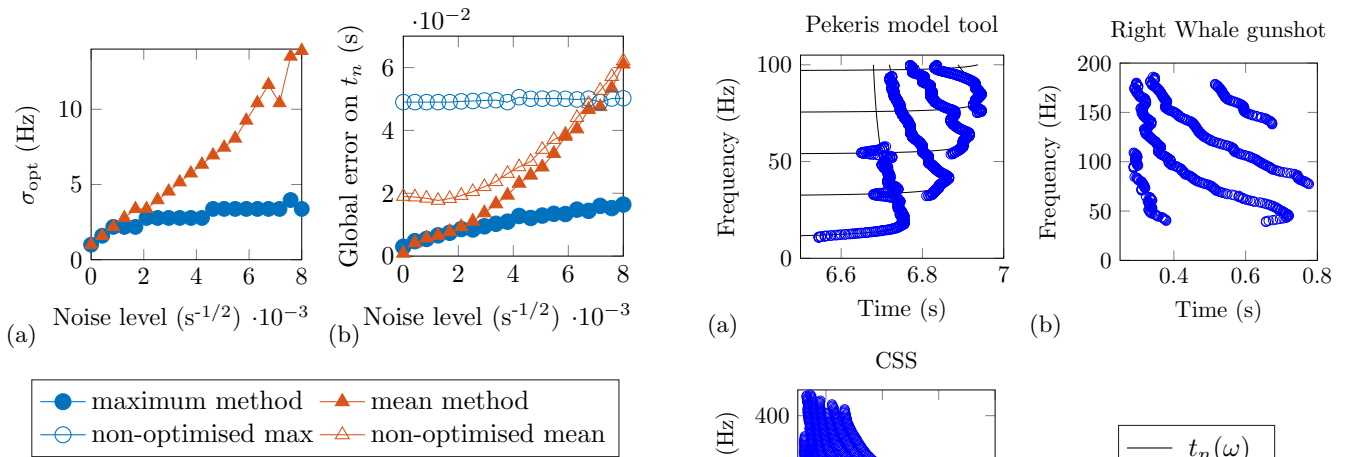


FIG. 12. (a) Optimal choice of σ to minimize the global error in the reconstruction of t_n for increasing values of the noise level $\delta T_\delta^{1/2}/|A_n(\omega)|$. (b) Global reconstruction error for the optimal choice $\sigma = \sigma_{\text{opt}}$, and comparisons with non-optimized errors with $\sigma = 20$ Hz (choice in (Bonnel *et al.*, 2020)). The choice of threshold is $p = 0.5$.

- Secondly, the functional J is not convex and its optimization can be costly.
- Thirdly, the minimum in (44) may be reached at non-physical values, especially for c_w and ρ_w which actually do not vary much in shallow waters.

To solve the first issue, we propose to determine empirically the optimal threshold value, denoted as p , by executing the algorithm on the Pekeris model. This will allow us to identify the ideal value of p that minimizes the relative reconstruction error for all the parameters.

To solve the second and third issues, the paper (Bonnel *et al.*, 2020) chooses to reduce the number of variables and to fix the values of c_w, ρ_w, ρ_b, D . The optimization is done with the unknown r, c_b, dt and turns out to be much simpler. We develop a more powerful approach and adapt

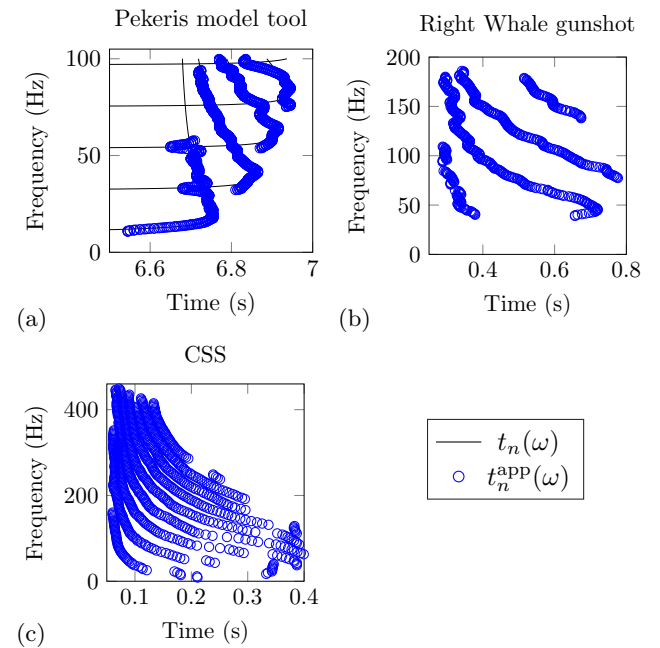


FIG. 13. Approximated dispersion curves $\omega \mapsto t_n^{\text{app}}(\omega)$ obtained using the maximum method with $\sigma = 5f_{\text{max}}/100$. We only present here the representative portions of curves such that $S_n(t_n^{\text{app}}(\omega), \omega) > 0.05 \max(S)$. (a) Pekeris model defined in Table I-line 1 whose modal component have been extracted in Figure 6. (b) impulsive right whale gunshot whose modal component have been extracted in Figure 7. (c) combustive sound source whose modal component have been extracted in Figure 8.

a Bayesian approach with prior knowledge of the values of some parameters. This approach helps us account for the fact that the Pekeris model is just an approximation of the actual propagation in shallow water. To do so, we add a penalization term to the functional J depending on hyper-parameters accounting for our confidence in the a

priori values $c_w^0, \rho_w^0, \rho_b^0, D^0$ of the parameters c_w, ρ_w, ρ_b and D .

These hyper-parameters are usually determined by training on big sets of data (Mohammad-Djafari, 1993; Pascal *et al.*, 2021). However, we only have two experiences in very different contexts. Despite this, we know that the estimates on the water layer are usually very reliable, while those on ρ_b and D are more uncertain. For this reason, we propose the following choice of hyper-parameters for the optimization and define a new functional

$$\begin{aligned} \tilde{J}(r, c_w, c_b, \rho_w, \rho_b, D, dt) &= J(r, c_w, c_b, \rho_w, \rho_b, D, dt) \\ &+ \alpha \left(\left| \frac{D - D^0}{D^0} \right|^2 + \left| \frac{\rho_b - \rho_b^0}{\rho_b^0} \right|^2 \right. \\ &\left. + 10 \left| \frac{\rho_w - \rho_w^0}{\rho_w^0} \right|^2 + 10 \left| \frac{c_w - c_w^0}{c_w^0} \right|^2 \right). \end{aligned} \quad (46)$$

Here, α is chosen so that the penalization part is of the same order as J . We will show in the following section that this choice of hyper-parameters yields good results for both experimental setups and the Pekeris model. By minimizing the penalized functional \tilde{J} , we obtain the *a posteriori* maximum of the corresponding Bayesian approach, which incorporates *a priori* information on $c_w^0, \rho_w^0, \rho_b^0$, and D^0 . For a more detailed explanation, see (Gribonval, 2011).

First, we investigate the optimal choice of threshold p on the Pekeris model. Then, we present the reconstruction obtained by our algorithm on experimental data.

A. Choice of the threshold p

The choice of the threshold p is important: if p is too close to 1, then the optimization is done with very few measurements and can become very sensitive to measurement errors. On the other hand, if p is too close to 0, then the reconstruction t_n^{app} can be very far from the actual values t_n which can tamper the reconstruction of the parameters. The good choice results from a trade-off to keep enough measurement points to prevent errors and remove non-significative measurements. Using the dispersion curves presented in Figure 13 for the Pekeris model, we minimize the functional \tilde{J} for different choices of p and present the relative reconstruction errors in Table II. Given these data, we choose from now on to work with the choice of threshold $p = 0.4$ and provide the reconstruction Algorithm 3.

B. Experimental data

We now minimize the functional \tilde{J} with a threshold $p = 0.4$ on the approximated dispersion curves obtained in Figure 13 for the experimental data. As explained in (Bonnell *et al.*, 2018, 2020), the environment of the combustive sound source experiment is only comparable to a Pekeris waveguide when frequencies are between 100 Hz and 300 Hz, and we only consider these data for the inversion. In Figure 14, we present the dispersion curves t_n associated with the parameters minimizing the functional

p	r	c_w	c_b	ρ_w	ρ_b	D	dt
0.2	2 %	2 %	2 %	<1%	2%	3%	1%
0.4	<1 %	<1 %	<1 %	<1%	2%	<1%	<1%
0.6	<1 %	<1 %	<1 %	<1%	3%	2%	1%
0.8	<1 %	<1 %	<1 %	<1%	9%	3%	<1%

TABLE II. Relative reconstruction errors of $r, c_w, c_b, D, \rho_w, \rho_b$ in the Pekeris model with respect to the threshold p .

ALGORITHM 3: Recovery of parameters

Input: Measured signal u at times $\tau = \tau_1, \dots, \tau_T$
sampled at frequency f_{\max}
Set of frequencies F to compute the spectrogram
Number of modes N
Estimated values of $c_w^0, \rho_w^0, \rho_b^0$, and D^0

Output: Recovered parameters $r, c_w, c_b, D, \rho_w, \rho_b$
 $S \leftarrow$ spectrogram of u computed with $\sigma = 5f_{\max}/100$ (see (9) and (11))
 $u_1, \dots, u_n \leftarrow$ filtering of u using Algorithm 1
 $P \leftarrow 0.4 \max(S)$
 $t_1^{\text{app}}, \dots, t_n^{\text{app}} \leftarrow$ result of Algorithm 2 on u_1, \dots, u_n with the threshold P
 $r^0, c_b^0, dt^0 \leftarrow$ arbitrary values
 $\alpha \leftarrow J(r^0, c_w^0, c_b^0, \rho_w^0, \rho_b^0, D^0, dt^0)$
 $(r, c_w, c_b, \rho_w, \rho_b, D, dt) \leftarrow$ minimization of \tilde{J} defined in (46) and starting at $(r^0, c_w^0, c_b^0, \rho_w^0, \rho_b^0, D^0, dt^0)$

\tilde{J} . For the impulsive right whale gunshot, we obtain the following parameters: $r = 8.8\text{km}$, $c_b = 1727\text{m/s}$, $dt = 5.9\text{s}$, $c_w = 1449\text{m/s}$, $\rho_w = 1007\text{kg/m}^3$, $\rho_b = 1481\text{kg/m}^3$, $D = 51.2\text{m}$. For the combustive sound source, we obtain $r = 4.8\text{km}$, $c_b = 1649\text{m/s}$, $dt = 3.2\text{s}$, $c_w = 1476\text{m/s}$, $\rho_w = 996\text{kg/m}^3$, $\rho_b = 1282\text{kg/m}^3$, $D = 73.1\text{m}$. These results are consistent with the ground truth and those exposed in Refs. (Bonnell, 2010; Bonnell *et al.*, 2020; Thode *et al.*, 2017) and in Table I. We even find a better localization result in the combustive sound source case, thanks to the relaxation of the parameters c_w, ρ_w, ρ_b , and D .

VI. CONCLUSION

In this paper, we have presented theoretical tools and numerical methods for automating the process of source localization and medium parameter estimation in shallow waters. Our approach involves 1) using a watershed algorithm and optimized warping to effectively separate modal components of the signal and 2) developing a stability-optimized maximum method to obtain accurate approximations of dispersion curves. The method also makes it possible to incorporate prior estimates on certain medium parameters. The application of this automated method on two experimental data sets show that it accurately recovers source location and additional medium parameters.

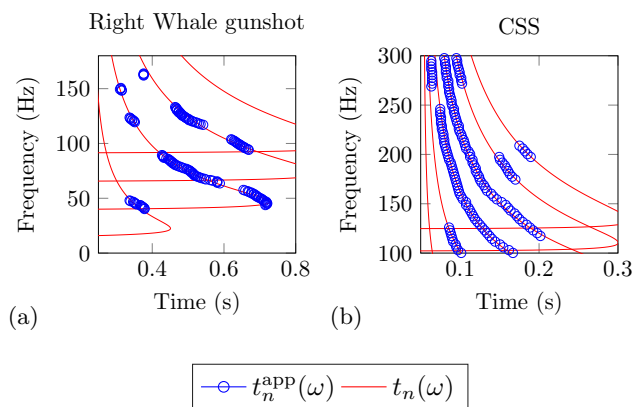


FIG. 14. Dispersion curves $\omega \mapsto t_n(\omega)$ obtained by minimizing the functional \tilde{J} on data $t_n^{\text{app}}(\omega)$ presented in Figure 13-(b-c). (a) Right whale gunshot data. (b) Combustive sound source data.

The definition of the quality factor of the warping in (25) and the choices of hyper-parameters in (46) follow from theoretical considerations. We have shown that these choices provide reliable estimates of the parameters of interest and that the automated method demonstrates stable performance with respect to the chosen values of the hyper-parameters. They could, however, be updated in the presence of additional information about the experimental setup. Indeed, with further experimental data, it is possible to refine the choices to improve the accuracy of our estimates even further. To summarize, our current method demonstrates significant potential for completely automating the process of source localization and medium parameter estimation in shallow waters.¹

¹See supplementary materials attached to find the mathematical developments of the paper, and at https://github.com/niclas-angele/source_localization for all the Matlab codes associated to the article.

Ainslie, M. (2010). *Principles of Sonar Performance Modelling* (Springer Berlin Heidelberg).

Aparicio, J., Jiménez, A., Ureña, J., and Álvarez, F. J. (2015). “Realistic modeling of underwater ambient noise and its influence on spread-spectrum signals,” in *OCEANS 2015 - Genova*, pp. 1–6, doi: [10.1109/OCEANS-Genova.2015.7271473](https://doi.org/10.1109/OCEANS-Genova.2015.7271473).

Aubauer, R., Lammers, M. O., and Au, W. W. L. (2000). “One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water,” *The Journal of the Acoustical Society of America* **107**(5), 2744–2749, doi: [10.1121/1.428660](https://doi.org/10.1121/1.428660).

Boashash, B. (2016). *Time-Frequency Signal Analysis and Processing (Second Edition)* (Academic Press, Oxford).

Bonnell, J. (2010). “Analyse de la dispersion acoustique UBF (0-150 Hz) pour la surveillance et la caractérisation du milieu marin,” Theses, Institut National Polytechnique de Grenoble, <https://theses.hal.science/tel-00522789>.

Bonnell, J., Gervaise, C., Nicolas, B., and Mars, J. I. (2012). “Single-receiver geoacoustic inversion using modal reversal,” *The Journal of the Acoustical Society of America* **131**(1), 119–128, doi: [10.1121/1.3664083](https://doi.org/10.1121/1.3664083).

Bonnell, J., Lin, Y.-T., Eleftherakis, D., Goff, J. A., Dosso, S., Chapman, R., Miller, J. H., and Potty, G. R. (2018). “Geoacoustic inversion on the new england mud patch using warping and dispersion curves of high-order modes,” *The Journal*

of the Acoustical Society of America **143**(5), 405–411, doi: [10.1121/1.5039769](https://doi.org/10.1121/1.5039769).

Bonnell, J., Thode, A., Wright, D., and Chapman, R. (2020). “Non-linear time-warping made simple: A step-by-step tutorial on underwater acoustic modal separation with a single hydrophone,” *The Journal of the Acoustical Society of America* **147**(3), 1897–1926, doi: [10.1121/10.0000937](https://doi.org/10.1121/10.0000937).

F. Auger, O. Lemoine, P. G., and Flandrin, P. (2008). “Time-frequency toolbox” <https://tftb.nongnu.org>, accessed: 2022.03.21.

Gribonval, R. (2011). “Should penalized least squares regression be interpreted as maximum a posteriori estimation?,” *IEEE Transactions on Signal Processing* **59**(5), 2405–2410, doi: [10.1109/TSP.2011.2107908](https://doi.org/10.1109/TSP.2011.2107908).

Jensen, F. B., Kuperman, W. A., Porter, M. B., Schmidt, H., and Tolstoy, A. (2011). *Computational ocean acoustics*, **794** (Springer).

Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (2007). “An overview of fixed passive acoustic observation methods for cetaceans,” *Oceanography* **20**(4), 36–45.

Meyer, F. (1994). “Topographic distance and watershed lines,” *Signal Processing* **38**(1), 113–125, doi: [https://doi.org/10.1016/0165-1684\(94\)90060-4](https://doi.org/10.1016/0165-1684(94)90060-4) mathematical Morphology and its Applications to Signal Processing.

Mohammad-Djafari, A. (1993). “On the estimation of hyperparameters in bayesian approach of solving inverse problems,” in *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 495–498, doi: [10.1109/ICASSP.1993.319857](https://doi.org/10.1109/ICASSP.1993.319857).

Pascal, B., Vaiter, S., Pustelnik, N., and Abry, P. (2021). “Automated data-driven selection of the hyperparameters for total-variation-based texture segmentation,” *Journal of Mathematical Imaging and Vision* **63**(7), 923–952, doi: [10.1007/s10851-021-01035-1](https://doi.org/10.1007/s10851-021-01035-1).

Pekeris, C. L. (1948). *Theory of propagation of explosive sound in shallow water* (Geological Society of America).

Stanton, T. K., Lee, W.-J., and Baik, K. (2018). “Echo statistics associated with discrete scatterers: A tutorial on physics-based methods,” *The Journal of the Acoustical Society of America* **144**(6), 3124–3171, doi: [10.1121/1.5052255](https://doi.org/10.1121/1.5052255).

Thode, A., Bonnell, J., Thieury, M., Fagan, A., Verlinden, C., Wright, D., Berchok, C., and Crance, J. (2017). “Using non-linear time warping to estimate north pacific right whale calling depths in the bering sea,” *The Journal of the Acoustical Society of America* **141**(5), 3059–3069, doi: [10.1121/1.4982200](https://doi.org/10.1121/1.4982200).