# Passive source depth discrimination in deep-water

Remi Emmetiere, Julien Bonnel, Xavier Cristol, Marie Géhant, Thierry Chonavel

# Passive source depth discrimination in deep-water

Rémi Emmetière, Julien Bonnel, *Member, IEEE,* Xavier Cristol, Marie Géhant and Thierry Chonavel, *Member, IEEE.*

**Abstract**

This paper addresses the problem of passive source depth discrimination in ocean acoustics using a horizontal line array (HLA). The scope is restricted to low-frequency sources (frequency $f < 500$ Hz), broadband signals (bandwidth $B$ of a few Hz), deep-water environment (water depth $D > 1000$ m), and distant sources (range $r$ greater than several km) at the endfire position. In this context, the environment acts as a dispersive waveguide, and one should not use classical source localization methods based on plane waves or any other simplistic wave model. Instead, a method based on the modal behavior driving the propagation is proposed. It notably uses the concept of the waveguide invariant, a scalar that summarizes the waveguide dispersion. In deep water, the waveguide invariant largely depends on source depth, and thus is an interesting input for source depth discrimination. An algorithm is proposed to compute energy ratio in groups of modal interferences. The input data for the algorithm is a range-frequency intensity, as measured on a HLA. The modal interference groups are defined based on their respective waveguide invariant values which in turns depend on source depth. This idea is formalized to propose a source depth discrimination method, which is performed as a binary classification problem. As long as the sound speed profile features a surface thermocline, the algorithm does not require detailed knowledge about the environment and it allows the classification of sources under two hypotheses, above or under a user-chosen threshold depth.

**Index Terms**

Source localization, underwater acoustics, waveguide invariant, interference pattern, binary classification,

## I. Introduction

When considering low frequency acoustic waves, the oceanic soundscape arises from various sources (biological, anthropological or geological). In passive acoustics, it is of primary interest to localize and/or classify sources contributing to this soundscape. For instance, human experts or algorithms are able to recognize sources because they feature specific sound characteristics such as known time-frequency signatures. In some cases, assessing the depth also allows classification of the source. In particular, such idea is very attractive in an anti-submarine warfare context. Indeed, depth estimation algorithms can passively distinguish noise radiated by a surface ship from noise radiated by a submerged submarine.

We warmly acknowledge D. Fattaccioli, Direction Générale de l'armement (DGA), for his scientific support.

In this paper we address the problem of passive source depth discrimination using a Horizontal Line Array (HLA). The objective of source depth discrimination is to classify a sound source based on its depth, above or below a chosen threshold. Such discrimination methods are useful in place of traditional source depth estimation methods, in context that are too complex for classical methods to perform properly. This is notably the case of the underwater environment, which is highly fluctuating, supports a multipath/multimodal propagation, and is usually poorly known. The scope of the paper is restricted to low-frequency sources (frequency $f < 500$ Hz), broadband signals (bandwidth $B$ of a few Hz), deep-water environment (water depth $D > 1000$ m), and distant sources (range $r$ greater than at least several km) at the endfire position. In this context, the acoustic propagation is largely impacted by the oceanic environment (the geo-acoustic properties of the propagation media). Thus, classical plane wave or other simple wave models cannot be used, and beamforming-like localization methods are not performing well. In fact, the acoustic propagation is more conveniently described by the normal modes theory [1]. The oceanic environment acts as a dispersive waveguide, where several modes are propagating, each mode having its own dispersive (i.e. non-linear) group delay. The acoustic field is then described in terms of depth functions (vertically standing wave) and range propagating dispersive waves. It gave birth to several physics-based methods to either estimate the source depth or perform depth discrimination.

For instance, matched-field processing (MFP) has been proposed to include better environmental models, but it is known to perform poorly as soon as the environment is not perfectly known [2], [3]. Within a classification framework, MFP-like processing will be tested here using the generalized likelihood ratio test [4], [5], [6]. In real-life, the ocean environment is a dynamic system with properties varying over space and time so that it is usually not possible to model it accurately. Thus, MFP is often not applicable. Specific signal processing methods have been developed to detect this environmental mismatch [7] or to mitigate it [8], [9]. Another way to circumvent this issue is to design localization schemes that extract specific features of the acoustic signal which are less sensitive to environmental mismatch than the pressure field itself. Most of these methods, such as matched mode processing (MMP) [10], [11] or modal subspace analysis (MSA) [12], [13], [14], use propagation features that are extracted from the sound pressure. In particular, these algorithms are based on mode filtering, and thus require the modes to be resolvable. Unfortunately, this is usually not doable in a deep water. Indeed, the modal density increases with water depth, and even advanced modal filtering methods (e.g. [15], [16], [17]) usually fail in deep-water. As a result, MMP and MSA are mostly adapted to shallow water environments. However, a possible extension of the MSA to scenarios where modes are not resolvable is suggested by Conan et al. [18], who propose to use a simple wavenumber spectrum analysis instead of mode filters. The original method is proposed for shallow-water environment and monochromatic sources. In this paper, the method will be applied to deep-water scenario. This -as well as the MFP likelihood ratio test mentioned earlier- will be used as references to benchmark the method proposed in this paper.

On the other hand, other localization methods are based on the acoustic intensity signal (square modulus of the pressure), such as localization schemes based on the modal scintillation [19], [20] or the mode (or ray) interference pattern [21], [22], [23], [24]. In deep water, interference properties of the intensity have been used to localize sources at relatively short ranges (at max several tens of kilometers). In particular, one can use Lloyd mirror properties and

a deep vertical array [22], the interference pattern of four eigenrays in the first shadow zone [24], or interference of highly refracted waves as measured on a deep sensor [25]. All these methods are based on the specific physics driving the propagation in a particular area of the oceanic waveguide, as illustrated in Fig. 1. As such, they perform very well in their own area of interest, but none of them are suitable for our context: a distant source observed with a HLA whose maximum immersion depth is about a few hundred meters.

This paper proposes a depth discrimination method based on the underlying physics driving the propagation in the context described above, and thus is complementary with the literature. It notably uses the mode interference pattern (intensity signal) as an input data and a quantity called *waveguide invariant* [26]. Indeed, the waveguide invariant is a scalar, usually noted $\beta$, that summarizes the waveguide dispersion. Its relative invariance makes it a robust input for source range localization in shallow waters [27]. However, it is also well known that the invariant is not strictly constant for all the mode pairs [28]. In particular, the variation of $\beta$ in shallow water can be related to the source depth [29], [21]. A previous study [30] has shown that the $\beta$-values are also largely driven by the source depth in deep waters, and is thus interesting for our problem. This idea is further explored in this paper, and extended to propose a new depth discrimination method.

To do so, we use the concept of mode trapping and some general features of the wide-band interference pattern involving the waveguide invariant theory. The depth discrimination method is designed to work in environments where the water sound speed profile (SSP) has one unique minimum far from the surface, which is a common situation in the Mediterranean sea. Also, the source/array range has to be known or preliminary estimated [27], [31], although further analysis will show that the poposed method is robust to a $10\%$ ranging error. However, the method does not require detailed knowledge about the environment, and is particularly robust to mismatch. In particular, neither the water SSP nor the seabed properties need to be known accurately. Because of the physical hypotheses used to derive the method, it is designed to perform well for sufficiently long range, which corresponds to the distance after which the propagating modes have performed at least a half interference cycle [30]. This range is about 10 km in a classical Mediterranean sea context. Because of the underlying physics, the proposed method is restricted to source/array depth that are less than a few hundred meters. This is not a strong constraint because sources of interest are rarely deeper than that.

The paper is organized as follows. In Sec. II, the physical derivation of the acoustic field under the normal mode theory is reminded. Then introducing the concept of mode trapping, the general idea of the depth discrimination concept is presented. In Sec. III, the depth discrimination problem is stated as a binary classification problem. The statistical classifier based on the generalized likelihood ratio test along with the physic-based approaches involving wavenumber and interference spectral analyses are exposed in details. The main contribution of this paper relies on a new interference spectral analyses which notably involves a ratio of energy estimated on the 2-D Fast Fourier Transform (2-D FFT) of the interference pattern. Finally, in Sec. IV, the classifiers are validated against synthetic data in ideal deep-water experimental conditions and their performances are compared. The classifiers are also tested against synthetic data in degraded experimental conditions, by introducing realistic distortion of the array geometry or of the environmental model. The limitations of the methods are discussed in Sec. V, and the article ends with a brief conclusion.

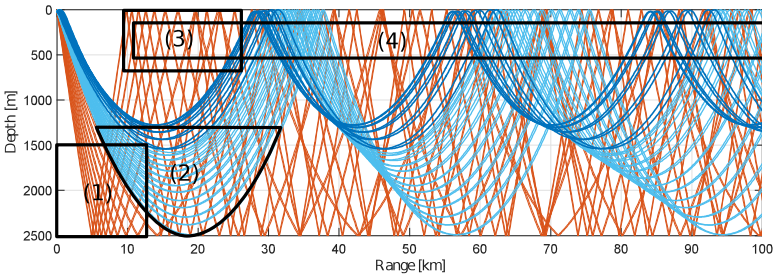


Fig. 1. A typical rays trace corresponding to a shallow source in a deep-water Mediterranean environment. One finds three types of rays: water born rays (dark blue), surface reflected rays (light blue) and bottom reflected surface reflected rays (orange). The superimposed black lines represent the possible positions of the receiver(s) where the McCargar's depth estimation method [22], indexed (1), the Weng's depth estimation method [24], indexed (2), the Duan's depth estimation method [25], indexed (3) and the proposed depth discrimination, indexed (4), are expected to work well.

## II. Normal mode propagation

### A. The acoustic field

At low-frequencies, underwater sound propagation is conveniently described by a set of normal modes [1]. We denote $s$ the complex pressure received at range $r$ and depth $z_r$ which originates from a point source radiating a broadband signal with spectrum $S(f)$ at depth $z_s$. From the normal mode theory, $s$ is a sum of $M$ propagating waves called modes. Up to a multiplicative constant, the pressure field is

$$s(z_s, z_r, r, f) = S(f)p(z_s, z_r, r, f) = S(f)\sum_{m=1}^{M} a_m(z_s, z_r, r, f)e^{i\phi_m(r,f)}, \tag{1}$$

with $p(z_s, z_r, r, f) = \sum_{m=1}^{M} a_m(z_s, z_r, r, f)e^{i\phi_m(r,f)}$ the impulse response of the waveguide. The phase of mode $m$ is given by

$$\phi_m(r, f) = k_m(f)r, \tag{2}$$

and its amplitude by

$$a_m(z_s, z_r, r, f) = \frac{\psi_m(z_s, f)\psi_m(z_r, f)}{\sqrt{k_m(f)r}}, \tag{3}$$

where $k_m$ and $\psi_m$ are the wavenumber and the depth function of the mode $m$, respectively.

In the following, the source spectrum is considered to be flat over the frequency bandwidth under study and is removed from the equations for notation convenience. Also, when considering relatively long range propagation, the modal amplitude $a_m$ is a slowly varying function of $r$ and $f$, compared to the exponential term in Eq. (1). As a result, the range/frequency dependence of $a_m$ can be ignored [1]. This approximation is referred to as the long range approximation in the paper. The validity of this approximation dictates the minimal source/receiver range to be considered with our method, as illustrated on Fig. 1.

In our propagation context, up to a constant term, the intensity of the signal is given by

$$I(z_s, z_r, r, f) \equiv |p(z_s, z_r, r, f)|^2 = \sum_{m,n,m\neq n} a_m(z_s, z_r)a_n(z_s, z_r)cos(r\Delta k_{mn}(f)), \tag{4}$$

with $\Delta k_{mn} = k_m - k_n$. The intensity shows interferences between modes that lead to striations in the $r - f$ plane [26] as illustrated by Fig. 2. Thus, $I(r, f)$ is called the interference pattern. In particular, the interference between modes $m$ and $n$ is defined by

$$I_{mn}(z_s, z_r, r, f) = a_m(z_s, z_r)a_n(z_s, z_r)cos(r\Delta k_{mn}(f)), \tag{5}$$

and it will be called interference striation in the following.

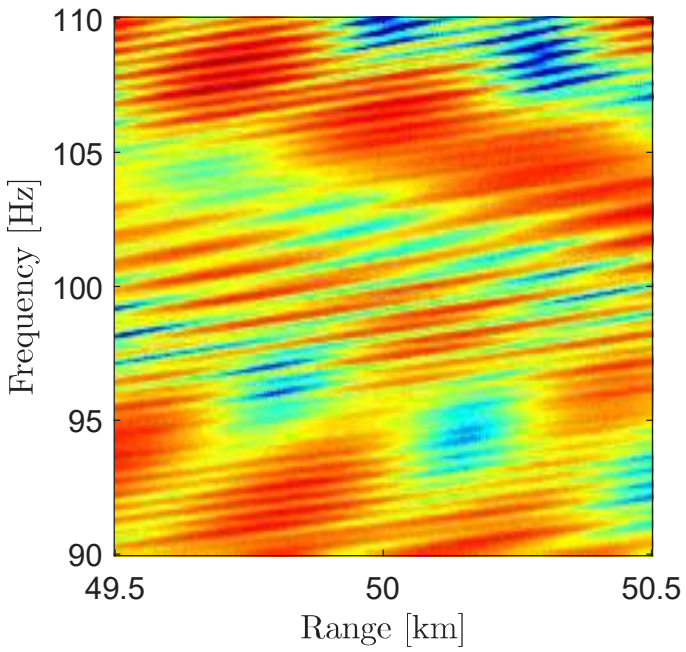


Fig. 2. The interference pattern $I(r, f)$ of a flat broadband spectrum source at $z_s = 150$ m recorded by a 1 km long HLA submerged at $z_r = 200$ m. The source-to-array range is $r = 50$ km. The intensity is computed using the normal mode KRAKEN code [32] with the deep-water waveguide environment defined by the TABLE I.

One already notes that, under the long range approximation, the source depth information is carried only by the mode amplitudes in both the pressure and the intensity fields.

### *B. Mode trapping*

From the normal modes theory, the modal depth functions oscillate between an upper and a lower turning point (at depths $z_m^+$ and $z_m^-$, respectively) and their amplitudes are exponentially decaying beyond these points [1]. For a given sound speed profile (SSP), noted $c(z)$, these turning points verify

$$c(z_m^{\pm}) = V_{p,m}, \tag{6}$$

with

$$V_{p,m} = 2\pi \frac{f}{k_m} \tag{7}$$

the phase velocity of the mode indexed by $m$.

The modes can be thus grouped in different types depending on where the depth functions oscillate. In deep-water the SSP is composed of at least one barocline. Indeed, increasing static pressure with depth produces a positive sound speed gradient in the water column. Especially in summer, the ocean surface is heated over the first few hundred meters and a negative thermocline (negative gradient of temperature) appears close to the surface. Thus, the SSP at the surface can be modeled by a negative gradient. An example of a typical Mediterranean sea environment

TABLE I
TYPICAL SUMMER DEEP WATER MEDITERRANEAN ENVIRONMENT

| depth [m] | sound speed [m/s] | density [kg/m$^3$] | attenuation [dB/$\lambda$] |
|---|---|---|---|
| 0 | $c_{surf} = 1530$ | 1030 | 0 |
| 100 | $c_{min} = 1500$ | 1030 | 0 |
| 2500 | $c_{max} = 1550$ | 1030 | 0 |
| 2500 | $c_{seabed} = 1700$ | 1700 | 0.6 |
| $\infty$ | $c_{seabed} = 1700$ | 1700 | 0.6 |

is presented in TABLE I and illustrated on Fig. $3(b)$. Considering this type of environment with four specific values of sound speed, it leads to three types of modes defined as follows:

- Trapped Modes (TM) for the modes verifying $c_{min} \leq V_{p,m} < c_{surf}$,
- Surface Interacting Modes (SIM) for the modes verifying $c_{surf} \leq V_{p,m} < c_{max}$ and
- Surface Interacting-Bottom Interacting Modes (SIBIM) for the modes verifying $c_{max} \leq V_{p,m} < c_{seabed}$.

An example of one depth function of each type is displayed on Fig. $3(a)$. The colored zones of Fig. $3(b)$ illustrate the depth areas where the mode functions are non-zero as a function of their phase velocity.

### C. The source depth discrimination concept

The source depth discrimination problem can be formalized as a binary classification problem. One wants to know if the source is above or under a threshold depth. One way to do so arises from the mode trapping phenomena. In fact, when considering a SSP with a surface thermocline, TMs have low amplitude near the surface. As a result, the mode amplitude $a_m$ [see Eq. (2)] is very small for a surface source if $m$ is a TM, while it is likely higher if the source is submerged. Thus, the depth discrimination problem can be recasted as a mode amplitude estimation problem. It can also be rephrased as follows: does the source excites TMs ?

According to Eq. (2) and Eq. (4), the acoustic field is a sum of complex exponential terms, and thus, they are conveniently represented in the Fourier domain. Using a HLA, the mode amplitudes can be assessed either in the spatial Fourier domain of the pressure field (wavenumbers) or in the spatial Fourier domain of the intensity field (interferences). In this paper, the first case is referred to as the *wavenumber spectral analysis* and the second as the *interference spectral analysis*. Please note that, the wavenumber spectral analysis is used in [12], [13], [14] to perform depth discrimination in shallow-water but has never been used in deep water. On the other hand, Turgut et

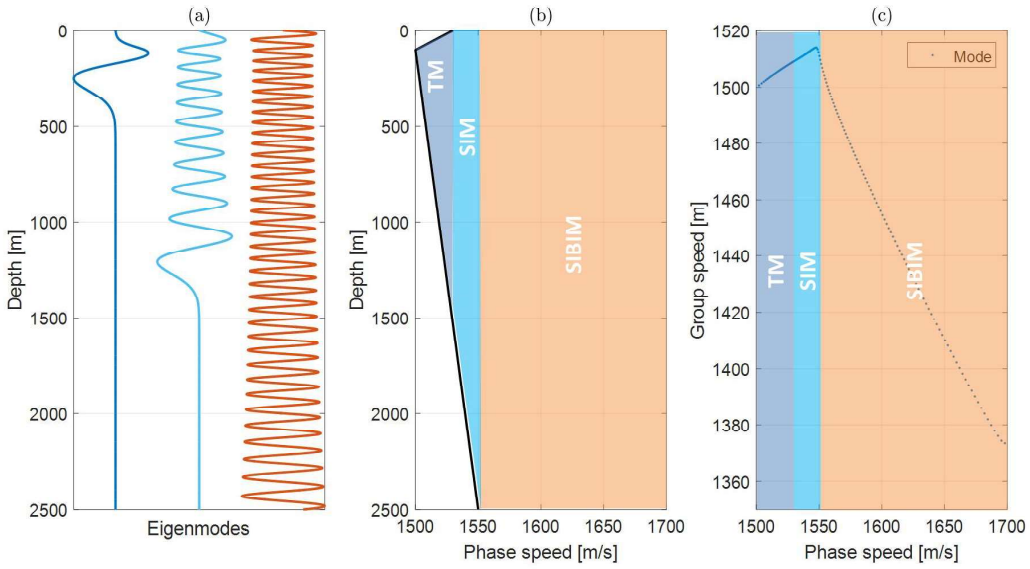


Fig. 3. $(a)$ The depth functions of one arbitrary mode of each type: TM (dark blue), SIM (light blue) and SIBIM (orange). $(b)$ The colored zones illustrate the depth areas where the mode functions are oscillating as a function of the phase velocity considering the SSP (drawn in black) with values listed in TABLE I. $(c)$ The modes are displayed as a function of their phase speed and group speed. The modes are calculated using KRAKEN at frequency $f = 100$Hz.

al. [21] has demonstrated that the interference pattern can be used to perform depth discrimination in shallow-water. However, none of these methods has ever been extended to deep-water. In this paper, building on Turguts idea, we will formalize a depth discrimination method based on the interference pattern. Because of the deep-water context, more modes are propagating, and the interference patterns are more complex than in shallow water. We thus propose a discrimination method that is explicitly based on the modal groups presented earlier: TM, SIM and SIBIM. As stated earlier, the method proposed by Conan et al [14] will also be directly transferred to a deep-water scenario. The deep-water extension of Conans method, as well as a generalized likelihood ratio test based on MFP, will serve as references to benchmark the proposed method. The next section presents these two existing methods, and then details the method proposed in this paper.

## III. Depth discrimination

### A. Tested hypotheses

As explained before, the source depth classification is formalized as a binary classification problem. The observation is classified into two classes: above or under a discrimination depth $z_{lim}$, within a set of possible depth $Z \in [0\ D]$. The classifier must choose between the two hypotheses

$$\begin{aligned} H_0 &: 0 < z_s \leq z_{lim} \\ H_1 &: z_{lim} < z_s < D, \end{aligned} \tag{8}$$

where $H_0$ and $H_1$ are the tested hypotheses corresponding to a shallow source and a submerged source, respectively.

For the specific application of the classification of radiated noise originating from surface ships or from submerged submarines, the discrimination depth is chosen to be a few meters below the surface. Indeed, when recorded at large range, a surface ship is known to be roughly equivalent to a point source with source depth from 1 m to 20 m. Thus, a discrimination depth of $z_{lim} = 20$ m is an appropriate candidate and will be used in the following. Nonetheless, the classifiers may features better performance for other discrimination depths where no obvious application arises.

Within this binary classification framework, the classifier also relies on a decision metric noted $\tau$. Then $\tau$ is compared to a decision threshold $\nu$, enabling a classification decision. Considering the two hypotheses $H_0$ and $H_1$ the problem is defined as follows:

$$\begin{aligned} \tau \leq \nu &\rightarrow H_0 \\ \tau > \nu &\rightarrow H_1. \end{aligned} \tag{9}$$

In the next section, three different approaches will be presented to design $\tau$, based on the signal model given by Eq. (10) . In particular, we introduce a statistical test based on the well known generalized likelihood ratio test [4], [5], [6]. Two other methods are detailed based on more heuristic arguments originating from the wavenumber and interference spectral analysis. The last one is the original contribution presented in this paper. Because it carefully takes into account the underlying physics, it will be shown to be more robust to model mismatch than the state-of-art.

### *B. Generalized likelihood ratio test (LRT)*

As stated above in Eq. (8), the alternative hypothesis $H_1$ is such that the parameter $z_s$ is in the complement of $H_0$. Traditionally, to solve this kind of problem one uses a statistical test, namely the generalized likelihood ratio test (LRT) where unknown parameters for each hypothesis are replaced by their maximum likelihood estimates (MLE). For notation convenience, we consider a version of Eq. (2) spatially sampled over a HLA. Then, the noisy signal sampled at positions given by vector $\mathbf{r}$ yields the multivariate observation:

$$\mathbf{s}(\mathbf{r}, f; z_s) = S(f)\mathbf{p}(\mathbf{r}, f; z_s) + \eta(f), \tag{10}$$

where $\eta(f)$ is spatially white noise following $\eta(f) \sim \mathcal{N}(0, \sigma^2)$ and $\mathbf{p}$ is a vector collecting the theoretical impulse response of the waveguide at ranges $\mathbf{r}$, as given by Eq. (1).

In our case, unknown parameters are $S(f)$, $\sigma^2$ and $z_s$. Considering the model (10) at a fixed frequency $f$, up to a constant factor and a power transform, the likelihood function is given by

$$\mathcal{L}(\mathbf{r}, f; z_s) = \frac{1}{\|\mathbf{s}(\mathbf{r}, f; z_s) - \mathbf{p}(\mathbf{r}, f; z_s) S_{\text{MLE}}\|} \tag{11}$$

where

$$S_{\text{MLE}} = \frac{\mathbf{p}^\dagger(\mathbf{r}, f; z_s)\mathbf{s}(\mathbf{r}, f; z_s)}{\|\mathbf{p}(\mathbf{r}, f; z_s)\|^2} \tag{12}$$

is the maximum likelihood estimation of the source spectrum at frequency $f$ and $^\dagger$ denotes the transpose conjugate operator. This function measures the likeness of the noisy signal to replica $\mathbf{p}$ generated for different values of the parameter $z_s$.

Then, the decision metric associated to the LRT is defined as follows [6]

$$\tau^{(\mathrm{LRT})}(f;z_s) = \frac{\max\{\mathcal{L}(\mathbf{r},f;z_s) : H_1\}}{\max\{\mathcal{L}(\mathbf{r},f;z_s) : H_0 \cup H_1\}}. \tag{13}$$

The LRT takes values between 0 and 1. Defined as Eq. (13), a high value of the test means the source is likely to be under hypothesis $H_1$ whereas a low value means it is more likely represented by the alternative $H_0$.

### C. Wavenumber spectral analysis (WSA)

In this section, we summarize the method developed in [14] for shallow water. It will later be applied to our deep water context. First, we introduce the wavenumber spectrum of the signal given by Eq. (1) by performing the spatial Fourier transform (FT) as

$$\tilde{s}(k,f;z_s) = \int_{r_0-\frac{L}{2}}^{r_0+\frac{L}{2}} s(r,f;z_s)e^{-ikr}\mathrm{d}r, \tag{14}$$

where $L$ is the array aperture and $r_0$ the mean source-to-array range.

Using Eqs. (6) and (7), the resulting wavenumber spectrum $\tilde{s}$ is split into two subspaces. The first subspace is associated to the TMs

$$k_{lim} < k < k_{min}. \tag{15}$$

The second subspace gathers all the propagating modes together

$$k_{seabed} < k < k_{min}, \tag{16}$$

with $k_{min} = \frac{2\pi f}{c_{min}} + \frac{2\pi}{L}$, $k_{lim} = \frac{2\pi f}{c(z_{lim})}$, and $k_{seabed} = \frac{2\pi f}{c_{seabed}} - \frac{2\pi}{L}$. When needed, the bounds have been stretched out by $\pm\frac{2\pi}{L}$ in order to include the main lobes due to the finite aperture $L$ of the HLA. An example of such subspaces are displayed in Fig. 4 for a surface source (red) and a submerged source (black).

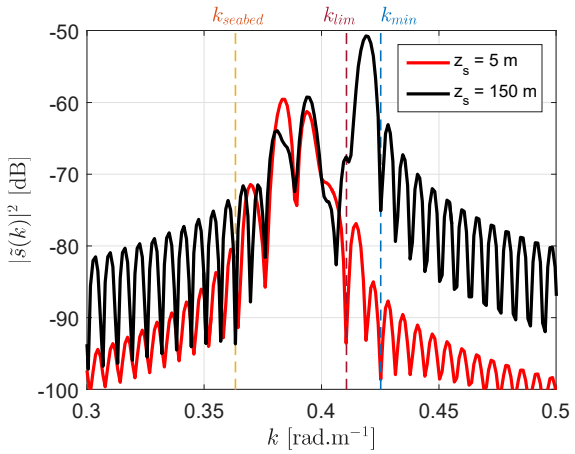


Fig. 4. Wavenumber spectrum calculated at $f = 100$ Hz for a 1 km long HLA with a 2 m spatial sampling and an infinite SNR. The pressure signal has been generated using the KRAKEN code with the same experimental conditions as the ones in Fig. 2.

In this example, there are 145 propagating modes in the deep-water waveguide, including 27 TMs. Fig. 4 clearly illustrates that the modes cannot be resolved and that modes amplitudes cannot be extracted from the wavenumber

spectrum. As suggested by Conan et al. [14], the decision metric for the classifier is computed as a simple integral over the two subspaces of wavenumber defined by Eq. (15) and Eq. (16)

$$\tau^{(\mathrm{WSA})}(f;z_s)=\frac{\int_{k_{lim}}^{k_{min}}\tilde{s}(k,f;z_s)\mathrm{d}k}{\int_{k_{seabed}}^{k_{min}}\tilde{s}(k,f;z_s)\mathrm{d}k}. \tag{17}$$

As a reminder, the decision metric is expected to be close to 0 for surface sources and close to 1 for submerged sources which excite TMs.

### *D. Interference spectral analysis (ISA)*

The main contribution of this paper is to conduct a similar analysis in the intensity domain. The goal is to define two subspaces, one gathering all the TMs together and the other all the propagating modes. Since it is not straightforward, the subsection is divided into successive steps for the sake of clarity. First, we define the 2-D Fourier transform of the interference pattern. Then, using a specific feature of the broadband structure of the acoustic field (the waveguide invariant), we show how to identify the interferences which involve TMs. This leads to the definition of two specific subspaces of the interference pattern. Finally, the decision metric for the classifier is presented.

*1) Interference spectrum:* The 2-D Fourier transform (2-D FT) of the interference pattern of bandwidth $B$ and central frequency $f_0$ observed by a HLA is defined by

$$\tilde{I}(\kappa,t;z_s)=\int_{f_0-\frac{B}{2}}^{f_0+\frac{B}{2}}\int_{r_0-\frac{L}{2}}^{r_0+\frac{L}{2}}I(r,f;z_s)e^{-i2\pi(\kappa r+tf)}\mathrm{d}r\mathrm{d}f, \tag{18}$$

with $\kappa$ (in $\mathrm{m}^{-1}$) and $t$ (in s) the Fourier transform variables conjugate to range and frequency, respectively. Please note that $\kappa$ is a wavenumber up to a $2\pi$ multiplicative term. Inserting Eq. (4) into Eq. (18), the 2-D FT of a given striation pattern is

$$\begin{aligned}\tilde{I}(\kappa,t;z_s)=\sum_{m,n,m\neq n}&a_m(z_s,z_r)a_n(z_s,z_r)sinc(\pi\kappa L)sinc(\pi tB)*\\&\Big[\delta(\kappa-\kappa_{mn},t-t_{mn})+\delta(\kappa+\kappa_{mn},t+t_{mn})\Big],\end{aligned} \tag{19}$$

where $*$ denotes the convolution, $\delta$ is the Dirac delta function, and the pair $(\kappa_{mn},t_{mn})$ is the Fourier coordinates of the particular interference striation $I_{mn}$. In terms of image processing, $\kappa_{mn}$ is the spatial frequency of $I_{mn}$ along the range axis and $t_{mn}$ is the spatial frequency of $I_{mn}$ along the frequency axis. According to Eq. (2) one finds

$$\kappa_{mn}(f)=\frac{\Delta k_{mn}(f)}{2\pi} \tag{20}$$

and

$$t_{mn}(r,f)=r\frac{\partial\Delta k_{mn}(f)}{2\pi\partial f}. \tag{21}$$

Recognizing the phase velocity $V_{p,m}=2\pi\frac{f}{k_m}$ and the group velocity $V_{g,m}=2\pi\frac{\partial f}{\partial k_m}$ (respectively the phase slowness $S_{p,m}=\frac{1}{V_{p,m}}$ and the group slowness $S_{g,m}=\frac{1}{V_{g,m}}$), one obtains

$$\kappa_{mn}(f)=f\Delta S_{p,mn}(f) \tag{22}$$

and

$$t_{mn}(r,f)=r\Delta S_{g,mn}(f), \tag{23}$$

with $\Delta S_{p,mn}(f) = S_{p,m}(f) - S_{p,n}(f)$ and $\Delta S_{g,mn}(f) = S_{g,m}(f) - S_{g,n}(f)$. In physical terms, $t_{mn}$ is the difference of travel times between the two interfering modes $m$ and $n$. Furthermore, $\kappa_{mn}$ refers to a wavenumber difference which is similar to a difference of arrival angles.

An example of such a transform is displayed by Fig. 5. This is the 2-D FT of the intensity shown in Fig. 2. Focusing on the interference identified by the dark dot, one recognizes the 2-D main lobe and the secondary lobes along the $\kappa$ and the $t$ axis. Especially, this interference correspond to striations in Fig. 2 with spatial frequency along the frequency axis of $t = -0.95$ Hz$^{-1}$ and along the range axis of $\kappa = 0.0038$ m$^{-1}$. In other terms this interference involves modes $m$ and $n$ with travel time difference $t_{mn} = -0.95$ s and wavenumber difference $\kappa_{mn} = 0.0038$ m$^{-1}$. These specific striations are visible in figure 2 over the 1 km aperture between 95 to 105 Hz.

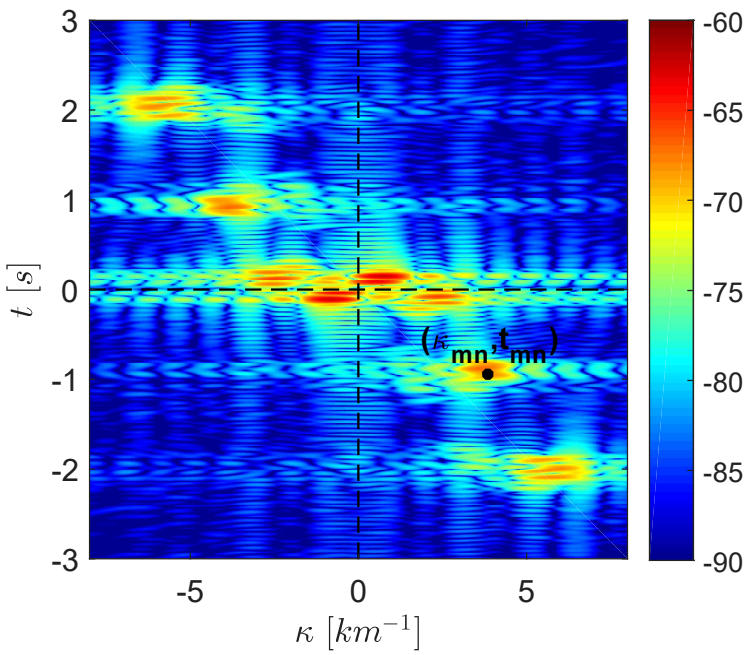


Fig. 5. 2-D FT $\tilde{I}(\boldsymbol{\kappa}, \mathbf{t})$ of the interference pattern presented in figure 2

Furthermore, from Eq. (19), the intensity is real-valued. Thus, the interference spectrum is symmetrical along the horizontal and the vertical axis, because the Fourier transform is Hermitian. In the following we chose to consider the right half of the interference spectrum ($\kappa \geq 0$) which contains the entire signal information.

*2) Interferences involving trapped energy:* In this subsection we use the broadband properties of the modes to identify the interferences that involve TMs. In particular, a quantity called *waveguide invariant*, usually noted $\beta$, has been introduced to quantify the slope of the interference striations in the log $r-$log$f$ plane [26]. This quantity is not derived in this manuscript, but the following result can be demonstrated by setting to zero the first-order Taylor series expansion of $I(\log r, \log f)$. For more details see [1]. Considering the particular interference between

modes $m$ and $n$ the waveguide invariant is defined by:

$$\beta_{mn}(f) = -\frac{\Delta S_{p,mn}(f)}{\Delta S_{g,mn}(f)}. \tag{24}$$

One notes from Eq. (24) that the waveguide invariant describes the slope of the interference striation with a scalar that does not depend on the source-receiver configuration and only depends on the interfering modes $m$ and $n$. This is a key property for our problem. It allows the characterization of interferences with a quantity that does not depend on the source/receiver configuration. In the Fourier domain, inserting Eqs. (22) and (23) in Eq. (24), the waveguide invariant is given by

$$\beta_{mn}(f) = -\frac{r}{f}\frac{\kappa_{mn}(f)}{t_{mn}(r,f)}. \tag{25}$$

This means that the position of one interference $(\kappa_{mn}, t_{mn})$ in the interference spectrum can be associated to the waveguide invariant.

Furthermore, usually the waveguide invariant is roughly constant for modes having the same type of propagation (i.e. TMs, SIMs, SIBIMs). This can be visualized by plotting the group velocity as a function of the phase velocity, as illustrated by Fig. 3(c). According to Eq. (24), the waveguide invariant is closely related to this representation since it relies on these two quantities (or more precisely the inverse of these two quantities). Considering the SSP previously presented, it is well known that $\beta_{mn} \approx 1$ if $m$ and $n$ are both SIBIMs and $\beta_{mn} \approx -3$ if $m$ and $n$ are TMs or SIMs [1] .

Finally, Fig. 6 displays an example of $I(\kappa, t)$ for $(a)$ a shallow source and $(b)$ a submerged source. Lines have been superimposed corresponding to waveguide invariants $\beta = -3$ and $\beta = 1$, according to Eq. (25). One can see that different interferences with different values of $\beta$ are located into different areas of $I(\kappa, t)$. One can thus define subspaces of $I(\kappa, t)$ that are associated to a given type of interference, estimate the modal energy in these subspaces, and use it as an input for depth discrimination. This will be detailed in the following subsection, but one can already note that a high energy content along the black line $\beta = -3$ in Fig. 6 refers to a submerged source whereas a low energy content is associated to a surface source.

*3) Subspaces of the interference spectrum:* As stated before, there are three types of modes : TMs, SIMs, SIBIMs. There are thus six types of possible interferences: TM-TM,TM-SIM, TM-SIBIM, SIM-SIM, SIM-SIBIM and SIBIM-SIBIM. Among them, we are interested in the subspace of the interference spectrum that contains at least one TM (i.e. interferences TM-TM, TM-SIM and TM-SIBIM), because the energy in this subspace highly depends on source depth, as explained before. In the following, we will use physical arguments to localize various interference subspaces in the $\tilde{I}(\kappa, t)$ domain. To do so, we define a subset of modes $\mathbf{N}$ driven by a common waveguide invariant noted $\beta_{\mathbf{N}}$. In particular, with the SSP defined in TABLE I

- $\beta_{\mathbf{N}} \approx -3$ if $\mathbf{N}$ is a subset of TMs,
- $\beta_{\mathbf{N}} \approx -3$ if $\mathbf{N}$ is a subset of SIMs,
- $\beta_{\mathbf{N}} \approx 1$ if $\mathbf{N}$ is a subset of SRIRMs.

Next, lets consider a mode $n \in \mathbf{N}$. For any other mode $m$, the Fourier coordinates of the interference between modes $m$ and $n$ is given by Eq. (25). However, $\beta_{mn}$ is unknown if $m \notin \mathbf{N}$ , in this case $\beta_{mn}$ may take a wide

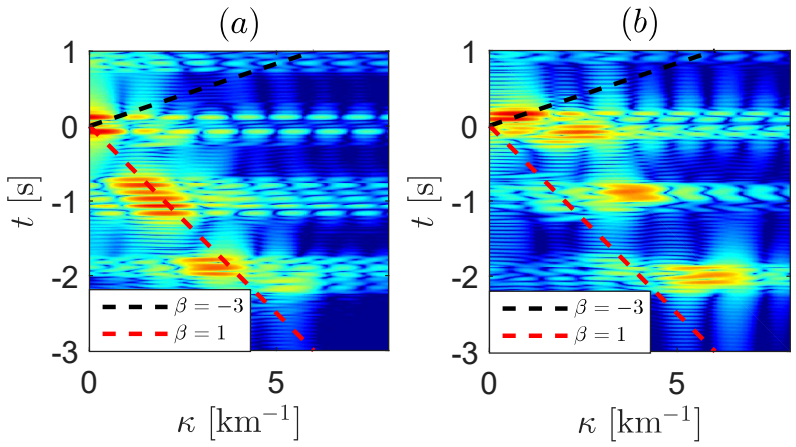


Fig. 6. The 2-D FFT $\tilde{I}(\kappa, t)$ of a 20 Hz bandwidth acoustic intensity simulated over a 1 km long array with the KRAKEN code, $(a)$ for a shallow source $z_s = 5$m and $(b)$ for a deep source$z_s = 150$m. The red lines shows the direction of interference with $\beta = 1$ and the black line shows the direction of interference with $\beta = -3$.

range of values from almost $-\infty$ to $+\infty$. Nevertheless, because Eq. (23) is linear, one can always decompose Eq. (25) by picking an arbitrary mode $l$ as

$$t_{mn} = t_{ml} + t_{ln} = -\frac{r}{f}\left(\frac{1}{\beta_{ml}}\kappa_{ml} + \frac{1}{\beta_{ln}}\kappa_{ln}\right). \tag{26}$$

To simplify Eq. (26), we restrain the choice of $l$ so that $l \in \mathbf{N}$ and $\beta_{ml} = +\infty$ hence

$$t_{mn} = -\frac{r}{f\beta_{\mathbf{N}}}\kappa_{ln}. \tag{27}$$

We will see later the impacts of such a choice. For now, using the linearity of Eq. (22), Eq. (27) can be reformulated as follows

$$t_{mn} = -\frac{r}{f\beta_{\mathbf{N}}}(\kappa_{mn} - \kappa_{ml}). \tag{28}$$

Finally, remembering that $\kappa_{mn}$ is the Fourier coordinate conjugate to the range variable, the interferences between a mode $m$ and all the modes of the subset $\mathbf{N}$ are aligned in the interference spectrum along a straight line of equation

$$t(\kappa) = -\frac{r}{f\beta_{\mathbf{N}}}(\kappa - \kappa_{ml}), \;\; \kappa \in [f\Delta S_{p,min} \;\; , \;\; f\Delta S_{p,max}], \tag{29}$$

where $\Delta S_{p,min}$ and $\Delta S_{p,max}$ are defined as the minimum and the maximum phase slowness difference between the mode $m$ and the subset $\mathbf{N}$.

Equation (29) is particularly important. It gives the position of the modal interferences in the $\tilde{I}(\kappa, t)$ domain. However, if the environment is unknown, the quantities $\kappa_{ml}$, $\Delta S_{p,min}$ and $\Delta S_{p,max}$ are also unknown. Nonetheless,

they can be roughly estimated based on the mode type(s) of $m$ and $\mathbf{N}$. For instance one can show, based on Fig. 3(c), that

- if $m$ and $\mathbf{N}$ are both of the same type then $\beta_{ml} = +\infty$ only if $\Delta S_{g,ml} = 0$. As a result, $m = l$ and thus $\kappa_{ml} = 0$.
- if $m$ is a SIM and $\mathbf{N}$ the subset of SIBIMs then $\kappa_{ml}$ can be roughly bounded by $0 < \kappa_{ml} < f(\frac{1}{c_{surf}} - \frac{1}{c_{max}})$.
- if $m$ is a TM and $\mathbf{N}$ the subset of SIBIMs then $\kappa_{ml}$ can be roughly bounded by $f(\frac{1}{c_{surf}} - \frac{1}{c_{max}}) < \kappa_{ml} < f(\frac{1}{c_{min}} - \frac{1}{c_{max}})$
- if $m$ is a TM and $\mathbf{N}$ the subset of SIMs then $\Delta S_{g,ml} = 0$ only if $m = l$ hence $\kappa_{ml} = 0$.

The previous list, giving the values of $\beta_{\mathbf{N}}$, $\kappa_{ml}$, $\Delta S_{p,min}$ and $\Delta S_{p,max}$ depending on mode type, is summarized by the TABLE II.

TABLE II

INTERFERENCE PARAMETERS SORTED BY TYPE OF INTERFERENCES

| $m$ / $\mathbf{N}$ | TMs | SIMs | SIBIMs |
|---|---|---|---|
| TM | $\beta_{\mathbf{N}} = -3$<br>$\Delta S_{p,min} = 0$<br>$\Delta S_{p,max} = \frac{1}{c_{min}} - \frac{1}{c_{surf}}$<br>$\kappa_{ml} = 0$ | $\beta_{\mathbf{N}} = -3$<br>$\Delta S_{p,min} = 0$<br>$\Delta S_{p,max} = \frac{1}{c_{min}} - \frac{1}{c_{max}}$<br>$\kappa_{ml} = 0$ | $\beta_{\mathbf{N}} = 1$<br>$\Delta S_{p,min} = \frac{1}{c_{surf}} - \frac{1}{c_{max}}$<br>$\Delta S_{p,max} = \frac{1}{c_{min}} - \frac{1}{c_{seabed}}$<br>$f\Delta S_{p,min} < \kappa_{ml} < f(\frac{1}{c_{min}} - \frac{1}{c_{max}})$ |
| SIM | | $\beta_{\mathbf{N}} = -3$<br>$\Delta S_{p,min} = 0$<br>$\Delta S_{p,max} = \frac{1}{c_{surf}} - \frac{1}{c_{max}}$<br>$\kappa_{ml} = 0$ | $\beta_{\mathbf{N}} = 1$<br>$\Delta S_{p,min} = 0$<br>$\Delta S_{p,max} = \frac{1}{c_{surf}} - \frac{1}{c_{seabed}}$<br>$\Delta S_{p,min} < \kappa_{ml} < f(\frac{1}{c_{surf}} - \frac{1}{c_{max}})$ |
| SIBIM | | | $\beta_{\mathbf{N}} = 1$<br>$\Delta S_{p,min} = 0$<br>$\Delta S_{p,max} = \frac{1}{c_{max}} - \frac{1}{c_{seabed}}$<br>$\kappa_{ml} = 0$ |

Using the model defined by Eq. (29) and the bounds given in TABLE II, we build two subspaces of the interference spectrum noted $D_0$ and $D_1$. Subspace $D_0$ gathers all the possible interferences and $D_1$ gathers only interferences involving at least one TM. In addition, the interference pattern is windowed (finite bandwidth over a finite range aperture) and thus the 2-D FT is a sum of delta function convoluted by sinc functions. In order to include in the subspaces the main lobes, some bounds have been stretched out by $\frac{1}{L}$. Moreover, one knows that exotic $\beta$-values can arise from the micro-structures in the SSP [26] or from spatial variations of the environment [33]. To be less restrictive, the upper borders of the $D_0$ and $D_1$ are defined with $\beta = -1$ (instead of $-3$) and the left border with $\beta = 0.75$ (instead of $1$ previously). The shape of these subspaces are schematically illustrated by Fig. 7.

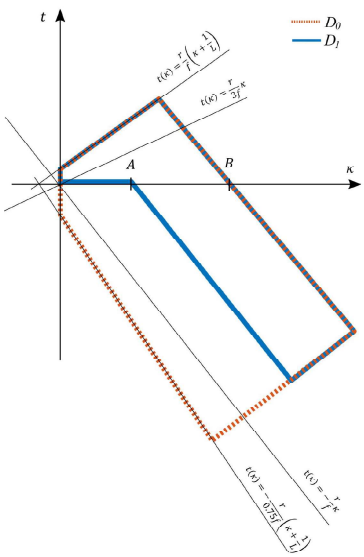


Fig. 7. From TABLE II, the interferences involving TMs are within $D_1$ and the all the possible interferences are within $D_0$. The specific points $A = (f(\frac{1}{c_{max}} - \frac{1}{c_{surf}}), 0)$ and $B = (f(\frac{1}{c_{min}} - \frac{1}{c_{seabed}}) + \frac{1}{L}, 0)$.

Now that the subspaces are clearly defined, the decision metric used by the classifier is computed as follow

$$\tau^{(\mathrm{ISA})}(f_0; z_s) = \frac{\int_{D_1} \tilde{I}(\kappa, t; z_s) \mathrm{d}\kappa \mathrm{d}f}{\int_{D_0} \tilde{I}(\kappa, t; z_s) \mathrm{d}\kappa \mathrm{d}f}. \tag{30}$$

An example of the subspaces are superimposed on the interference spectrum of Fig. 8 for $(a)$ a surface source and $(b)$ a submerged source. The decision metric is expected to be close from 1 for submerged sources and close from 0 for surface source.

The proposed methods has several differences with the state of the art [12], [13], [14], where modal subspaces are built in the wavenumber domain. Because we consider intensity, a quadratic quantity, one has to deal with cross terms, and thus interferences between mode types. This complicates the definition of the subspaces that are used to build the discrimination metrics. On the other hand, by using broadband information and the waveguide invariant concept, we showed that the interference subspaces can be defined with minimal environment information: only crude estimates of $\beta$ are required. Nonetheless, the price to pay is that the method requires a priori information about the source/array range $r$, which can be a strong limitation in real-life. However, it will be shown later that $r$ and $\beta$ do not need to be known accurately. In the next section, the three approaches (LRT, WSA and ISA) will be compared for different scenarios that include realistic model errors.

## IV. Performance evaluation with Monte Carlos method

### A. General methodology

The performance of a binary problem can be evaluated using the detection probability $P_D$ and the false alarm probability $P_{FA}$ of a given classifier. We chose to call detection the observations attributed to the hypothesis $H_1$.

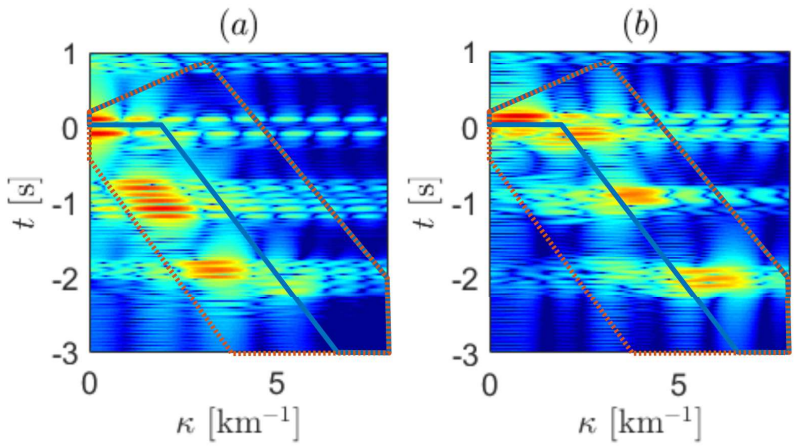


Fig. 8. Interference spectrum calculated for the same configuration than in Fig. 6. Subspace $D_0$ is superimposed with dashed orange lines whereas subspace $D_1$ is displayed in blue line.

With this convention, $P_D$ is the probability of a submerged source to be classified under the hypothesis $H_1$ and $P_{FA}$ is the probability of a surface source to be classified under the hypothesis $H_1$. The compromise between $P_D$ and $P_{FA}$ allows one to choose the decision threshold based on desired performances.

Like in Ref. [14], these probabilities are numerically estimated using Monte Carlo methods. The results are then displayed using Receiver Operating Characteristic (ROC) curves which provides a general representation of the relevance of the decision metric. The performance of the classifiers along with the choice of the decision thresholds are presented for the ideal case of perfectly known experimental conditions, as well as in degraded experimental conditions, where the environment and the array geometry are partially unknown (environmental and array mismatch). Since the ISA requires a priori knowledge of the source/array range, which in practice cannot be perfeclty known, our approach is also tested against ranging errors.

### *B. Numerical simulation specification*

For the simulations, we consider a HLA of aperture $L = 1$ km and an inter-sensor spacing of $\Delta r = 10$ m. The HLA is placed at depth $z_r = 200$ m. We also consider a white noise source with a flat spectrum $S(f)$ over a bandwidth $B = 20$Hz with central frequency $f_0 = 100$Hz. Gaussian white noise is added on the recieved signal so that the signal-to-noise ratio

$$SNR = \frac{\|s(r, f; z_s)\|^2}{E\{\|\eta(f)\|^2\}} \tag{31}$$

is constant for all source/receiver combinations, and set to 0 dB. Last but not least, please note that a smaller $\Delta r = 5$ m is chosen for the wavenumber spectral analysis method, in order to respect the Nyquist-Shannon sampling theorem.

The scenario is repeated for different source positions, with ranges from 1 to 99 km with 0.5 km steps, and depths from 1 to 400 m, with 1 m steps. This defines a grid of almost $80,000$ observations with different source/array configurations. Only one noise realization is computed for each configuration but considering the amount of source/array configurations of the experimental grid, it still provides a good assessment for the ROC curves.

It is important to note that in the following the generalized likelihood ratio test (LRT) along with the wavenumber spectral analysis (WSA) will be computed only at the central frequency $f_0$, whereas the interference spectral analysis (ISA) uses the whole signal bandwidth. However, this is enough to illustrate the sensitivity of the LRT and of the WSA to respectively the environmental mismatch and the array geometry mismatch.

As a first step, simulated data will be generated using a modal code (KRAKEN [32]) for both the observations and the replica needed for the LRT. In particular, in Sec. IV-C, the LRT, WSA and the ISA will be validated against this KRAKEN's data under perfectly known experimental conditions. In the same section, the decision thresholds $\nu$ for the different methods are determined in order to ensure a desired maximal $P_{FA}$ of $5\%$. The decision threshold along with the decision metric fully define a classifier for each approach. Next, the observations will be synthesized using the RAM code [34]. The RAM code is a fully numerical model that does not use the concept of modes, but is based on Parabolic Equations (PE). It allows one to test the classifiers (based on a mode decomposition of the acoustic field) against numerical data generated by a model that does not use modes, preventing the inverse crime [35]. Also, realistic array geometry mismatch and environmental mismatch will be considered.

### *C. Classifications with known experimental conditions*

The considered environment is a range and azimuth independent waveguide fully defined by its SSP. The seabed and the water column characteristics are given in TABLE I. In this section we use the same environmental model for the direct problem (generating the data) and for the inversion (performing the depth discrimination). This allows one 1) to obtain a proof of concept for the depth discrimination methods and 2) to chose the decision threshold.

As described in Sec. IV-B, the decision metrics for the three methods are calculated for each source/array configuration defined by the experimental grid. Their empirical distributions under the two hypothesis are estimated based on the $79,600$ observations and are displayed in Fig. 9. The color correspond to the considered hypothesis (blue for $H_1$ and orange for $H_0$). The line style is associated to the method, especially, the probability density function of $\tau^{(\mathrm{LRT})}$ is displayed with dotted line, $\tau^{(\mathrm{WSA})}$ with dashed line and the proposed $\tau^{(\mathrm{ISA})}$ with solid line. This convention will be kept along the section.

In the case of a perfectly experimental condition, the LRT is known to be optimal. As displayed by Fig. 9, the density probability functions (PDF) of the associated metric decision under the two hypothesis are very well separated. For WSA and ISA, the PDFs are also separated, validating the approaches in deep-water.

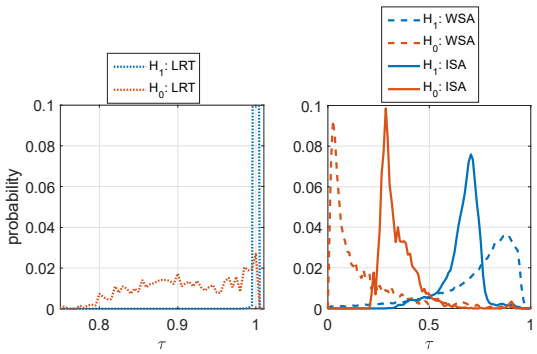


Fig. 9. PDFs of $\tau$s under the two hypotheses $H_1$ (blue) and $H_0$ (orange). They are estimated using Monte Carlos methods for the three approaches : LRT (dotted line), WSA (dashed line) and the ISA (solid line). The probability axis has been cut off at 0.1 on purpose for representation convenience. However, the sharply peacked distribution function for the LRT under the hypothesis $H_1$ does not fit the scale.

The relevance of the different approaches are compared using the ROC curves. They are directly estimated using the repartition functions previously evaluated. Fig. 10 gives a representation of these ROC curves. Please note that for the ISA, two ROCs have been calculated 1) for relatively short range configurations ($r < 10$km) and 2) for long range propagation ($r \geq 10$km). The differences between the two ROC curves is due to the long range approximation used in Eq. (4). This approximation is not valid for ranges $r < 10$ km, which impacts the waveguide invariant [30] and negatively affects our method. Beyond this range, the decision metric is more stable. In the following, the ISA classifier is designed to guarantee its performance only for long range configurations ($r > 10$ km). Moreover, as expected, the LRT approach has results that are nearly perfect.

For the targeted false alarm probability of $P_{FA} = 5\%$ the classifiers achieve respectively detection probabilities of $P_D^{(\mathrm{LRT})} = 99.9\%$, $P_D^{(\mathrm{WSA})} = 84.3\%$ and $P_D^{(\mathrm{ISA})} = 93.8\%$. The associated decision thresholds are $\nu^{(\mathrm{LRT})} = 0.995$, $\nu^{(\mathrm{WSA})} = 0.56$ and $\nu^{(\mathrm{ISA})} = 0.45$.

By inserting $\nu$ into Eq. (9), the classifications is processed for the three methods and the results are displayed on Fig. 10. The sources classified under $H_1$ are displayed with black color whereas the ones associated to $H_0$ are in white. The LRT method almost provides a perfect classification for every source/array configuration. It is interesting to note that for the ISA, most of the false detections are done for ranges of few kilometers where, by design, the performances are not guaranteed. Considering only $r > 10$ km (denoted by the gray dashed line in Fig. 10), all the missed detections and the false alarms are concentrated around the discrimination depth, which leads to a satisfactory classification result. Actually, the effective discrimination depth is largely influenced by the water column SSP. In the case under study, the depth where TMs are really dominating the acoustic field is located around 50 m, so that many classification errors occur between $z_{lim} = 20$ m and 50 m. This can also be seen on the WSA results, because the method is based on the same physics. However, at ranges $r > 10$ km, the WSA method

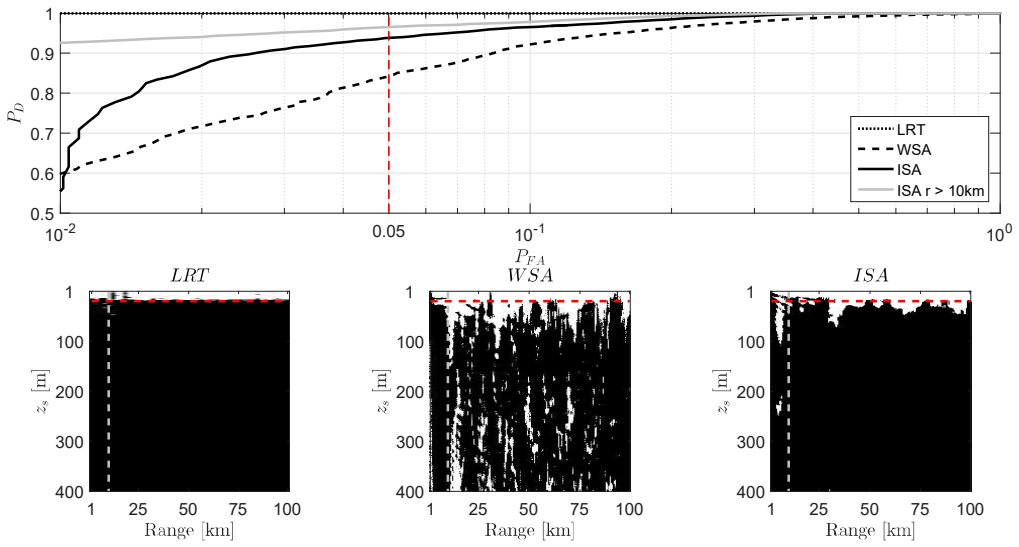


Fig. 10. Top panel: ROC curves for the scenario presented in Sec. IV-B. The red dashed line represent the targeted $P_{FA}$. Bottom panels: Observations classified under the hypothesis $H_1$ are displayed in black and in white for the hypothesis $H_0$. The desired discrimination depth is plotted with a red dashed line.

has extra errors at every depth.

In the following, the classifiers and detection thresholds defined in this section will be used in more realistic scenarios.

### *D. Results with environmental mismatch*

Although range-dependent models are convenient to model underwater acoustic propagation, a true ocean is always more complicated. In this section, we consider the impact of range-dependent effect on our localization method. Data are now simulated in a range-dependent waveguide, with a varying SSP in the water column and a varying bathymetry. As stated before, simulations are perfomed with RAM, a numerical PE code. In these simulations, the range-dependent SSP is defined using three different SSPs at three specific ranges (20km, 30km and 75km), and linear interpolation between these ranges. A zoom on the first 400 m of these three SSPs and the one used by the classifier is plotted in Fig. 11(a). The 2-D representation of the environment is displayed in Fig. 11(b), which notably shows the range varying bathymetry. A zoom on the first 400 m of the water column is presented in Fig. 11(c), which allows a good assessment of the SSP variability.

As previously, the data are generated at each node of the experimental grid, using the full 2-D SSP. On the other hand, the classifiers are ran assuming the previous simple 1-D SSP (dashed curve on Fig. 11(a)), which created environmental mismatch. The obtained ROC curves are plotted in Fig. 12.

The LRT performance drastically falls with environmental mismatch. The associated ROC curve tells that the decision metric is not discriminating. As it can be seen on Fig. 12, the LRT nearly behaves as a random classifier.

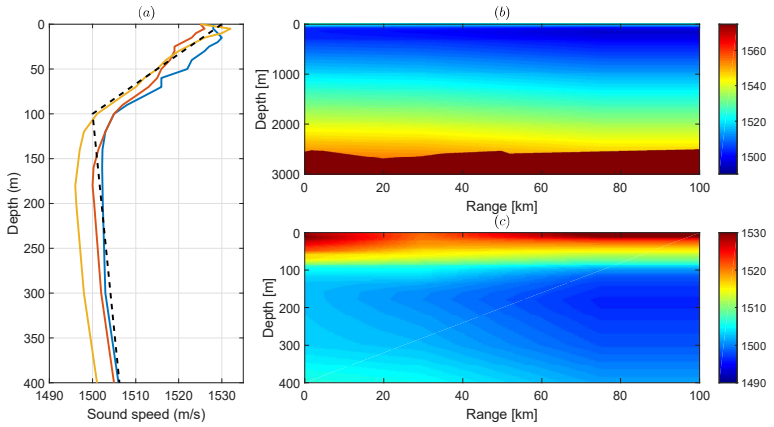


Fig. 11. $(a)$ The SSPs that define the environment in which data have been simulated. The dashed black plot is the SSP used by the classifier. $(b)$ 2-D representation of the environment with range variations of the SSP and the bathymetry. $(c)$ Zoom on the first 400 m of the SSP where the range variations are relatively high. Note that the colorscales for $(b)$ and $(c)$ are different.

On the other hand. the WSA results are relatively equivalent to the previous case with perfectly known experimental conditions. However, the classification features better results around the array depth (i.e. around the minimum of sound speed of the environment), concentrating the errors around the discrimination depth and far beyond the array depth. The performances of the ISA are degraded but not as seriously. It still provides good results for the depth discrimination problem. In particular one finds a $P_D = 90\%$ for the targeted $p_{FA} = 5\%$, which outperforms the WSA, as can be seen on the ROC curve in Fig. 12.

As earlier, ISA errors are mostly located around the discrimination depth or at short range. Moreover, because the environment is range-dependent, the effective discrimination depth induced by the environment may vary with the source/receiver configuration. This can be seen in Fig. 12. The ISA performance improves after 75 km, because the minimum of the sound speed channel becomes more pronounced (see Fig. 11). This comment also applies to the WSA.

### *E. Results with array geometry mismatch*

Most of the time, in operational conditions, the relative positions of the sensors of the array are not completely determined. In this section we investigate the robustness of the depth discriminations to array geometry mismatch. The experimental conditions are the same than in the previous Sec. IV-C (perfectly known environment). However, WSA and ISA are performed assuming a noisy inter sensor spacing $\widehat{\Delta r} \sim \mathcal{N}(\Delta r, 0.1\Delta r)$ and one uses the RAM to generate the data. Because the LRT already failed to discriminate the source depth with environmental model mismatch, it is not computed here.

As before, the decision metrics are computed for the WSA and the ISA, and ROC curves are displayed in Fig. 13. It notability shows that the WSA is very sensitive to array geometry mismatch. The associated decision metric is not relevant. This is due to a translation of the wavenumber subspaces Eqs.(16) (15) on the measured wavenumber spectrum. The same translation also affect the ISA. However, the wavenumber difference $\kappa_{mn}$ is less impacted by

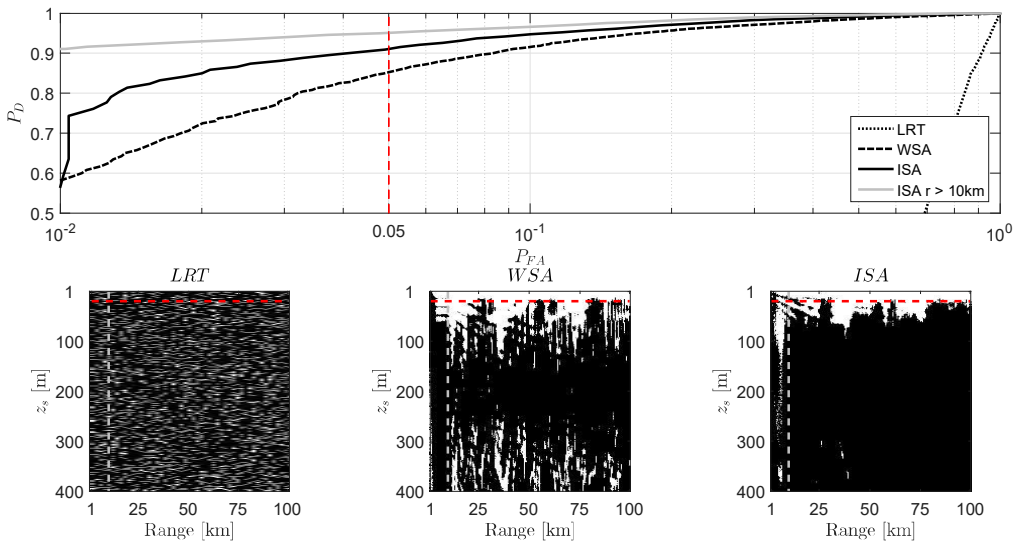


Fig. 12. Top panel: ROC curves for the scenario with environmental mismatch. The data are generated by the RAM code using the environment with values listed in TABLE I. The vertical red dashed line represents the targeted $P_{FA}$. Bottom panels: Observations classified under the hypothesis $H_1$ are displayed in black and in white for the hypothesis $H_0$. The desired discrimination depth is plotted with a horizontal red dashed line.

this translation than the wavenumber themselves. As a result, the ISA still provides good performances. One finds a $P_D = 89\%$ for the targeted $P_{FA}$ of 5%.

### *F. Results with source/array range errors*

As demonstrated by the study, the proposed depth discrimination based on ISA does not need a thorough model of the array or of the environment. However, the knowledge of the source/array range is required to build the zones $D_0$ and $D_1$ via Eq. (29). If this quantity is unknown, it can be estimated using, for instance, the interference pattern [27] or the waveguide invariant [31]. An algorithm, based only on the waveguide invariant, that simultaneously assesses the source/array range and performs the depth discrimination should be built. The performance of such a classifier will be obviously dependent of the range estimation accuracy.

Here, we consider that a localization method is able to provide a range estimation $\hat{r} \sim \mathcal{N}(r, 0.1r)$. This is typically the range estimation errors that one obtains with range estimation methods that use the waveguide invariant [31], [27]. Thus, the decision metric is calculated for each source-array configuration of the experimental grid by replacing $r$ in Eq. (29) by $\hat{r}$. Also, the observation are computed with the RAM code using the 2-D environment displayed in Fig 11 with the noisy inter-sensors spacing $\widehat{\Delta r}$ as in the previous subsection. Then, this study considers all at once, the environmental/array/range mismatch in order to simulate a more realistic scenario.

Following the same procedure as previously, the ROC curves for the ISA are evaluated and they are presented in Fig. 14(a). The performances are degraded but still decent. The method still provides a $P_D = 84\%$ for all the ranges, and $P_D = 88\%$ for the long range propagation case ($r > 10$km). The decision metric appears to be robust

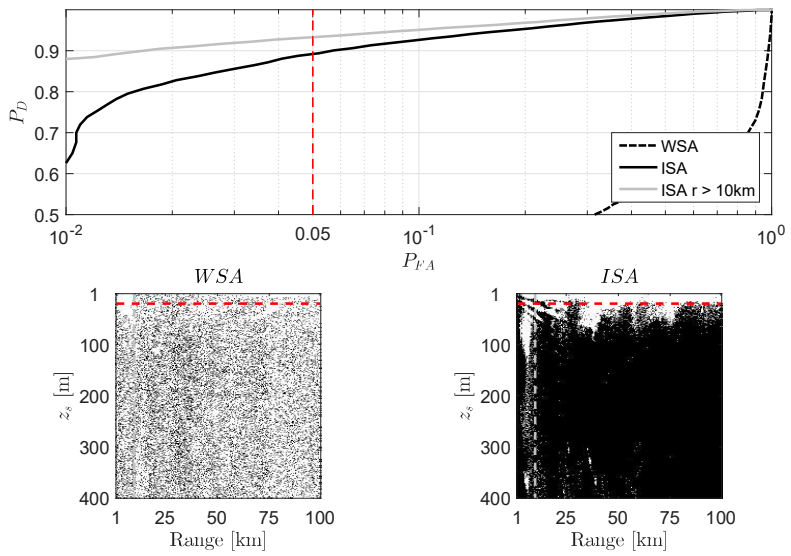


Fig. 13. Top panel: ROC curves for the scenario with both environment and array geometry mismatches. The data are generated by the RAM code using the environment with values listed in TABLE I. The vertical red dashed line represents the targeted $P_{FA}$. Bottom panels: Observations classified under the hypothesis $H_1$ are displayed in black and in white for the hypothesis $H_0$. The desired discrimination depth is plotted with a horizontal red dashed line.

to source/array range errors and more precisely to a combination of environmental/array/range mismatch. As an example, the classifier is performed on these data and the classification results are displayed by Fig. 14(b). It still provide a satisfactory discrimination.

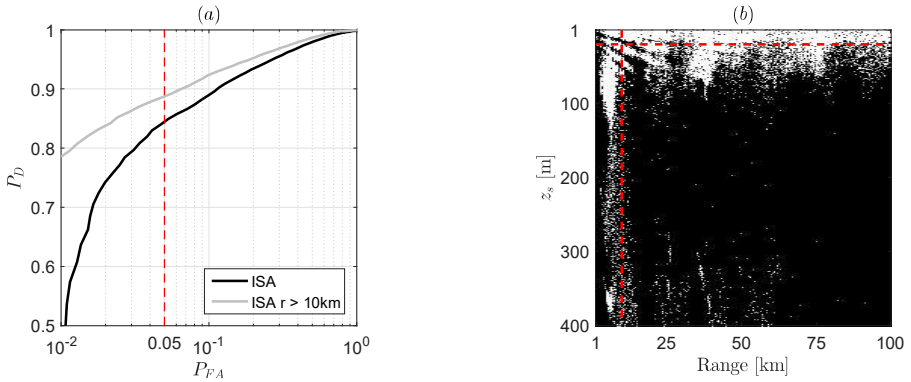


Fig. 14. (a) ROC curves for the scenario with environment, array geometry, and range mismatches. The data are generated by the RAM code using the environment with values listed in TABLE I. The vertical red dashed line represents the targeted $P_{FA}$. (b) Observations classified under the hypothesis $H_1$ are displayed in black and in white for the hypothesis $H_0$. The desired discrimination depth is plotted with a horizontal red dashed line. The vertical red dahed line is the limit r=10 km.

## V. Discussion

### A. Importance of the environment

In this section we emphasize the limitations of the proposed ISA for the depth discrimination problem. Indeed, the method is dedicated to several specific contexts. First of all, because the approach is based on the mode trapping concept, it is evident that the method needs the environment to allow mode trapping. In deep water this is always true, if the frequency of the signal is high enough so that the first modes can be trapped. Typically, with a water column of few thousands meters, mode trapping happens for signals with frequencies higher than few Hertz. Moreover, the method requires some of the modes to be trapped around the depth of the sources of interest, i.e. one needs a sound speed minimum (called sound channel) at a few hundred meters. This is (nearly) always the case in the Mediterranean sea. In some other oceans, such as the Atlantic, the sound channel is usually deeper, around $1000$ m. In this case, a large amount of TMs are not excited by submerged sources, if those are not very deep. As a result, it is expected that the method would not perform as well as illustrated on a Mediterranean scenario. In polar regions, the problem would be different, because the sound speed channel is usually at the surface. In this specific case, a surface source would excite more TMs than a submerged source, and the ISA method would need to be adapted.

Another requirement of the method is that the receiving array must be located in the vicinity of the sound channel, so that it received all mode types. In particular, considering the environment defined in TABLE I, further studies (not presented) here have shown that the array should be submerged between $75$ to $500$ m. Moreover, as already discussed, the arrayshould be far enough from the source, so that the derivation of the waveguide invariant is valid. As a reminder, this area is schematically illustrated by the area $(4)$ of Fig. 1.

### B. Coherence loss on long arrays

In real life, when considering a long HLA (1km), the signal is subject to coherence loss, which may impact the processing performance. Coherence loss can be due to both deterministic and stochastic phenomena. In our context, modal dispersion induces a deterministic coherence loss, which negatively impacts classical processing such as plane-wave beamforming. However, the modal dispersion is taken into account in our model, and thus the method is barely impacted by the deterministic coherence loss. Also, the performance study shows that ISA is robust to environmental mismatch, which is particularly important when complex environmental model are used within localization schemes.

On the other hand, stochastic coherence loss is also a major issue in underwater acoustics. If coherence loss is small (e.g. unsaturated regime), it is expected that ISA still performs well. This is also illustrated in the performance study, when we show robustness to array position (which is equivalent to a multiplicative phase noise, which in turn can be used to model coherence loss [36]). If coherence loss becomes too strong (e.g. fully saturated regime), then ISA will become ineffective. However, in such scenario, no signal processing method is doing magic. The development of source localization methods that perform well under strong stochastic coherence losses is still an open question, even if some solutions are being proposed in easier contexts, such as plane-wave without multipath

[37]. Unfortunately, adaptive optics methods (that reduce the effect of wavefront distortions) are not easy to transfer in the underwater acoustic context.

## VI. Conclusion

A depth discrimination method based on the deep-water waveguide invariant has been proposed. The classifier is validated on simulation with ideal experimental conditions and also successfully tested against realistic model errors. The robustness of the classifier to environmental model errors suggests that it could be successfully applied to real data. The proposed method proves to be reliable to environmental mismatch, unlike more convenient processing such as MFP-based LRT.

The robustness of the proposed method to array geometry errors is an important result. In fact, in an operational context, the soundscape arises from various sources. Hence, the interference pattern is the sum of the contributions of many sources emanating from different direction. Nevertheless, it has been shown that the interference pattern belonging to one specific source of interest can be recovered as the output of a horizontal beamformer [38]. In fact, if the source (and/or the array) is moving over time, the broadband output of a horizontal beamformer $I(\theta(t), f)$ is equivalent to the interference pattern $I(r(t), f)$, provided that the source is alone in its beam $\theta(t)$. Thus, one can use beamformed data and synthetic range aperture as an input of the depth discrimination algorithm proposed in this paper. However, the synthetic inter-sensors spacing relies on the relative source/array range variations between each snapshot which is never accurately known. In that case, being robust to array geometry mismatch is of paramount importance. The resilience of our method to such a model errors suggests that the proposed classifier could perform well in a multi-source context using the output of a horizontal beamformer.

## Acknowledgment

This work was funded by Delegation Général de l'Armement and by Thales Underwater Systems.

## References

[1] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*. Springer Science & Business Media, 2011.

[2] J. L. Krolik, "The performance of matched-field beamformers with mediterranean vertical array data," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2605–2611, 1996.

[3] Y. Le Gall, F.-X. Socheleau, and J. Bonnel, "Matched-field performance prediction with model mismatch," *IEEE Signal Processing Letters*, vol. 23, no. 4, pp. 409–413, 2016.

[4] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on signal processing*, vol. 42, no. 8, pp. 2146–2157, 1994.

[5] C. F. Mecklenbräuker, P. Gerstoft, J. F. Böhme, and P.-J. Chung, "Hypothesis testing for geoacoustic environmental models using likelihood ratio," *The Journal of the Acoustical Society of America*, vol. 105, no. 3, pp. 1738–1748, 1999.

[6] G. Casella and R. L. Berger, *Statistical inference*. Duxbury Pacific Grove, CA, 2002, vol. 2.

[7] J. Tabrikian, G. S. Fostick, and H. Messer, "Detection of environmental mismatch in a shallow water waveguide," *IEEE transactions on signal processing*, vol. 47, no. 8, pp. 2181–2190, 1999.

[8] J. C. Preisig, "Robust maximum energy adaptive matched field processing," *IEEE transactions on signal processing*, vol. 42, no. 7, pp. 1585–1593, 1994.

[9] ——, "A minmax approach to adaptive matched field processing in an uncertain propagation environment," *IEEE transactions on signal processing*, vol. 42, no. 6, pp. 1305–1316, 1994.

[10] T. Yang, "Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing," *The Journal of the Acoustical Society of America*, vol. 87, no. 5, pp. 2072–2084, 1990.

[11] Y. Le Gall, F.-X. Socheleau, and J. Bonnel, "Performance analysis of single-receiver matched-mode localization," *IEEE Journal of Oceanic Engineering*, no. 99, pp. 1–14, 2017.

[12] V. E. Premus and M. N. Helfrick, "Use of mode subspace projections for depth discrimination with a horizontal line array: Theory and experimental results," *The Journal of the Acoustical Society of America*, vol. 133, no. 6, pp. 4019–4031, 2013.

[13] E. Conan, J. Bonnel, T. Chonavel, and B. Nicolas, "Source depth discrimination with a vertical line array," *The Journal of the Acoustical Society of America*, vol. 140, no. 5, pp. EL434–EL440, 2016.

[14] E. Conan, J. Bonnel, B. Nicolas, and T. Chonavel, "Using the trapped energy ratio for source depth discrimination with a horizontal line array: Theory and experimental results," *The Journal of the Acoustical Society of America*, vol. 142, no. 5, pp. 2776–2786, 2017.

[15] J. Bonnel, G. Le Touze, B. Nicolas, and J. I. Mars, "Physics-based time-frequency representations for underwater acoustics: Power class utilization with waveguide invariant approximation," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 120–129, 2013.

[16] K. Xu, J.-G. Minonzio, D. Ta, B. Hu, W. Wang, and P. Laugier, "Sparse svd method for high-resolution extraction of the dispersion curves of ultrasonic guided waves," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 63, no. 10, pp. 1514–1524, 2016.

[17] A. Drémeau, F. Courtois, and J. Bonnel, "Reconstruction of dispersion curves in the frequency-wavenumber domain using compressed sensing on a random array," *IEEE J. Ocean. Eng*, vol. 42, no. 4, pp. 914–922, 2017.

[18] E. Conan, "Traitements adaptés aux antennes linéaires horizontales pour la discrimination en immersion de sources ultra basse fréquence," Ph.D. dissertation, Ecole nationale supérieure Mines-Télécom Atlantique, 2017.

[19] V. Premus, "Modal scintillation index: A physics-based statistic for acoustic source depth discrimination," *The Journal of the Acoustical Society of America*, vol. 105, no. 4, pp. 2170–2180, 1999.

[20] D. Fattaccioli and P. Danet, "A predictive physical model for modal scintillation due to source depth modulation," in *4th International Conference and Exhibition on Underwater Acoustic Measurements: Technologies & Results*, 2011.

[21] A. Turgut and L. T. Fialkowski, "Depth discrimination using waveguide invariance," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 2054–2054, Sep. 2012.

[22] R. McCargar and L. M. Zurk, "Depth-based signal separation with vertical line arrays in the deep ocean," *The Journal of the Acoustical Society of America*, vol. 133, no. 4, pp. EL320–EL325, 2013.

[23] R. Duan, K. Yang, H. Li, and Y. Ma, "Acoustic-intensity striations below the critical depth: Interpretation and modeling," *The Journal of the Acoustical Society of America*, vol. 142, no. 3, pp. EL245–EL250, 2017.

[24] J.-B. Weng and Y.-M. Yang, "Experimental demonstration of shadow zone localization using deep water interference patterns measured by a single hydrophone," *IEEE Journal of Oceanic Engineering*, 2017.

[25] R. Duan, K. Yang, H. Li, and Y. Ma, "Acoustic-intensity striations below the critical depth: Interpretation and modeling," *The Journal of the Acoustical Society of America*, vol. 142, no. 3, 2017.

[26] S. Chuprov, "Interference structure of the sound field in stratified ocean," *Ocean Acoustics*, 1982.

[27] K. L. Cockrell and H. Schmidt, "Robust passive range estimation using the waveguide invariant," *The Journal of the Acoustical Society of America*, vol. 127, no. 5, pp. 2780–2789, May 2010.

[28] Y. Le Gall and J. Bonnel, "Passive estimation of the waveguide invariant per pair of modes," *The Journal of the Acoustical Society of America*, vol. 134, no. 2, pp. EL230–EL236, 2013.

[29] D. Rouseff and R. C. Spindel, "Modeling the waveguide invariant as a distribution," in *AIP Conference Proceedings*, vol. 621. AIP Publishing, Jun. 2002, pp. 137–150.

[30] R. Emmetière, J. Bonnel, M. Géhant, X. Cristol, and T. Chonavel, "Understanding deep-water striation patterns and predicting the waveguide invariant as a distribution depending on range and depth," *The Journal of the Acoustical Society of America*, vol. 143, no. 6, pp. 3444–3454, 2018.

[31] K. A. Sostrand, "Range localization of 10-100 km explosions by means of an endfire array and a waveguide Invariant," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 1, pp. 207–212, Jan. 2005.

[32] M. B. Porter, "The kraken normal mode program," Naval Research Lab Washington DC, Tech. Rep., 1992.

[33] G. L. DSpain and W. A. Kuperman, "Application of waveguide invariants to analysis of spectrograms from shallow water environments that vary in range and azimuth," *The Journal of the Acoustical Society of America*, vol. 106, no. 5, pp. 2454–2468, 1999.

[34] M. D. Collins, "Userś guide for ram versions 1.0 and 1.0 p," *Naval Research Lab, Washington, DC*, vol. 20375, 1995.

[35] A. Wirgin, "The inverse crime," *arXiv preprint math-ph/0401050*, 2004.

[36] R. Dashen, W. H. Munk, and K. M. Watson, *Sound transmission through a fluctuating ocean*. Cambridge University Press, 2010.

[37] R. Lefort and A. Drémeau, "Sub-antenna sparse processing for coherence loss in underwater source localization," in *Signal Processing Conference (EUSIPCO), 2017 25th European*. IEEE, 2017, pp. 2413–2417.

[38] T. C. Yang, "Beam intensity striations and applications," *The Journal of the Acoustical Society of America*, vol. 113, no. 3, pp. 1342–1352, Mar. 2003.