HAL
open science

Passive stochastic matched filter for Antarctic blue whale call detection

Léa Bouffaut, Richard Dréo, Valerie Labat, Abdel-O. Boudraa, Guilhem Barruol

To cite this version:

Léa Bouffaut, Richard Dréo, Valerie Labat, Abdel-O. Boudraa, Guilhem Barruol. Passive stochastic matched filter for Antarctic blue whale call detection. Journal of the Acoustical Society of America, 2018, 144 (2), pp.955 - 965. 10.1121/1.5050520 . hal-01867897

HAL Id: hal-01867897

https://hal.univ-reunion.fr/hal-01867897v1

Submitted on 4 Sep 2018

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.

Passive stochastic matched filter for Antarctic blue whale call detection[a)]

Léa Bouffaut,[1,b)] Richard Dréo,[1] Valérie Labat,[1] Abdel-O. Boudraa,[1] and Guilhem Barruol[2]
[1]*Institut de Recherche de l'Ecole Navale, EA3634, Ecole Navale/Arts et Metiers ParisTech, BCRM Brest CC600, 29240 Brest Cedex 9, France*
[2]*Institut de Physique du Globe de Paris, Sorbonne Paris Cite, UMR 7154 CNRS, 1 rue Jussieu, 75238 Paris Cedex 05, France*

(Received 16 March 2018; revised 26 July 2018; accepted 26 July 2018; published online 23 August 2018)

As a first step to Antarctic blue whale (ABW) monitoring using passive acoustics, a method based on the stochastic matched filter (SMF) is proposed. Derived from the matched filter (MF), this filter-based denoising method enhances stochastic signals embedded in an additive colored noise by maximizing its output signal to noise ratio (SNR). These assumptions are well adapted to the passive detection of ABW calls where emitted signals are modified by the unknown impulse response of the propagation channel. A filter bank is computed and stored *offline* based on *a priori* knowledge of the signal second order statistics and simulated colored sea-noise. Then, the detection relies on *online* background noise and SNR estimation, realized using time-frequency analysis. The SMF output is cross-correlated with the signal's reference (SMF + MF). Its performances are assessed on an ccean bottom seismometer-recorded ground truth dataset of 845 ABW calls, where the location of the whale is known. This dataset provides great SNR variations in diverse soundscapes. The SMF + MF performances are compared to the commonly used MF and to the Z-detector (a subspace detector for ABW calls). Mostly, the benefits of the use of the SMF + MF are revealed on low signal to noise observations: in comparison to the MF with identical detection threshold, the false alarm rate drastically decreases while the detection rate stays high. Compared to the Z-detector, it allows the extension of the detection range of $\simeq 30$ km in presence of ship noise with equivalent false discovery rate.

I. INTRODUCTION

Blue whales worldwide were driven close to extinction by commercial and illegal whaling during the 20th century. As an endangered species, they have been protected internationally since 1965.[1] This decision has been endorsed in the southern hemisphere by the creation of the Indian Ocean Whale Sanctuary in 1979, the first area where all types of commercial whaling are banned.[2]

The investigation of such a large open-water area requires adequate methods, complementary to visual observations. Passive acoustic monitoring is an economical, non-intrusive, durable, and efficient technique for blue whale survey using their loud (between 180 and 190 dB[3,4]) and low frequency (<50 Hz) sounds. Whale monitoring is usually performed using acoustic measurements from the sound fixing and ranging channel.[3,5,6] However, seismic data recorded from the seafloor have been shown to provide valuable information, especially when considering passive acoustic monitoring of blue whales[7–9] with a detection range easily exceeding 100 km.[10]

Data recorded by this mean offer a complete immersion in the sensor's area soundscapes. The Antarctic blue whale (ABW) stereotyped call (named Z-call) occurring in the [18–26.8] Hz bandwidth, is a good candidate for automatic detection.[6,11–13] However, numerous sound sources overlap in the same frequency band such as high energy noises due to seismic activity or ship radiated tonal noises.[14]

Passive acoustic monitoring automatic detection strategies are usually based on signal cross-correlation functions such as temporal MFs[12] or spectrogram-based template matching.[5] They rest on strong assumptions of white background noise and deterministic signals. Due to the propagation, these assumptions are hard to fulfilled in passive acoustic monitoring. Consequently, those methods are not well adapted to ABW call detection on data recorded in low signal to noise ratio (SNR) environments, whether it is due to high background noise in the signal's bandwidth or distant call detection.

To overcome the aforementioned methods limitations, the approach presented in this paper is based on the stochastic matched filter (SMF). The SMF was originally derived from the MF for active sonar processing to maximize the SNR of stochastic signals embedded in colored noise. These assumptions are also well suited for passive detection of stereotyped marine mammals calls[15,16] where emitted signals are modified by the unknown impulse response of the propagation channel. Therefore, at the reception, they can be considered as stochastic. Preliminary results on ABW call

[a)]Preliminary results of this work were presented at EUSIPCO, Kos Island, Greece, 2017.
[b)]Electronic mail: lea.bouffaut@ecole-navale.fr

detection and scuba-divers breathing detection have illustrated the potential of this filter at low SNR, with various types of noises.[13] The work presented in this paper aims to evaluate the SMF performances with strong SNR variations. It is realized using a ground truth dataset, where a set of 845 Z-calls are manually annotated by an expert. Results are compared to classical temporal cross correlation (matched filters) and to the *Z-detector*, a method recently developed based on subspace-detection.[6,17]

To present the SMF for ABW call detection in a passive context and evaluate its performances, the remainder of the paper is organized as follows. Section II presents the experimental database with a special attention to soundscapes and Z-call description. Section III focuses then on the detection strategy. Section IV deals with performances assessment using a ground truth dataset. Finally, Section V discusses the results and some current limitations of the method and how they might be mitigated.

II. EXPERIMENTAL DATA

A. Dataset

1. RHUM-RUM network and the SWIR array

Data used in this work were recorded in the southwest Indian Ocean by the RHUM-RUM (Réunion Hotspot and Upper Mantle—Réunion's Unterer Mantel)) seismic network[18–20] and are currently hosted at the RESIF French national seismic archive center.[21,22] Part of the RHUM-RUM project was to characterize the micro-seismicity associated to an active seamount located on the Southwest Indian Ridge (SWIR). A dense sub-array of eight autonomous ocean bottom seismometers (OBSs), denoted as the SWIR array, was deployed locally between October 2012 and November 2013. Each OBS was equipped with a three component seismometer and a hydrophone, recording data continuously during the immersion time with a 100 Hz sampling frequency. Only hydrophone-recorded data are used in this work. The SWIR array covered an area of 70 km × 40 km with depth varying from 2822 to 5430 m. Its inter-sensor distance of about 20 km is convenient for multiple sites observation of baleen whale songs and therefore for their localization and tracking (Fig. 1).[9] For further information, the complete RHUM-RUM network and OBSs technical descriptions are detailed in the RHUM-RUM performance report.[23]

2. Whale localization and tracking

The description of remote ABW songs (Sec. II C) and the evaluation of the passive SMF performances on real data (Sec. IV) are performed using the result of the tracking of a calling ABW swimming through the SWIR array on May 31st, 2013 (Fig. 1). This individual's song is recorded continuously, for more than 21 h (from 01:20 to 22:40 on OBS RR48). The tracking method, fully detailed in Ref. 9, is based on time difference of arrival. The localization takes into account the acoustic propagation context by including environmental characteristics such as the region's complex bathymetry.

B. Soundscape

In addition to ABW songs, diverse sound sources occur in the SWIR array area below 50 Hz and are interpreted as soundscapes.[14] Their origin can be geophysical (sound sources generated by non-biological natural sources, e.g., meteorological noises, earthquakes, iceberg tremors, etc.), biological (non-human biological sound sources, e.g., whale songs) or anthropogenic (sound generated by humans[24]). As an example, Fig. 2 displays two spectrograms recorded on May 28th, 2013, at two different moments (respectively 06.00 and 15.30), on OBS RR47. These data are representative of the soundscapes wealth offered by this area, under

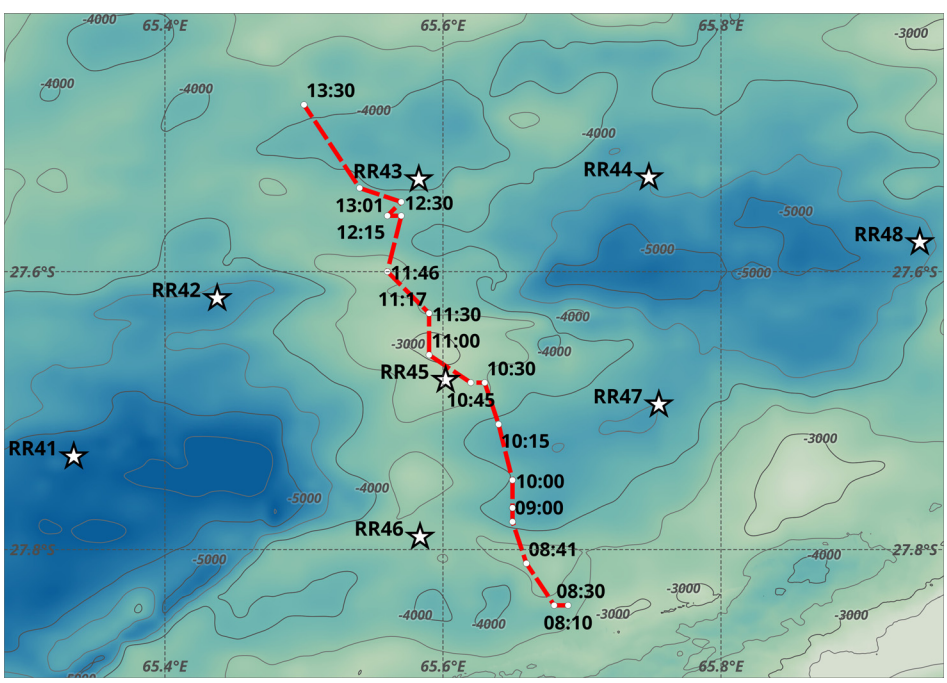

FIG. 1. (Color online) ABW passive acoustic monitoring tracking through the SWIR array. Star denotes the OBS positions (Ref. 9).

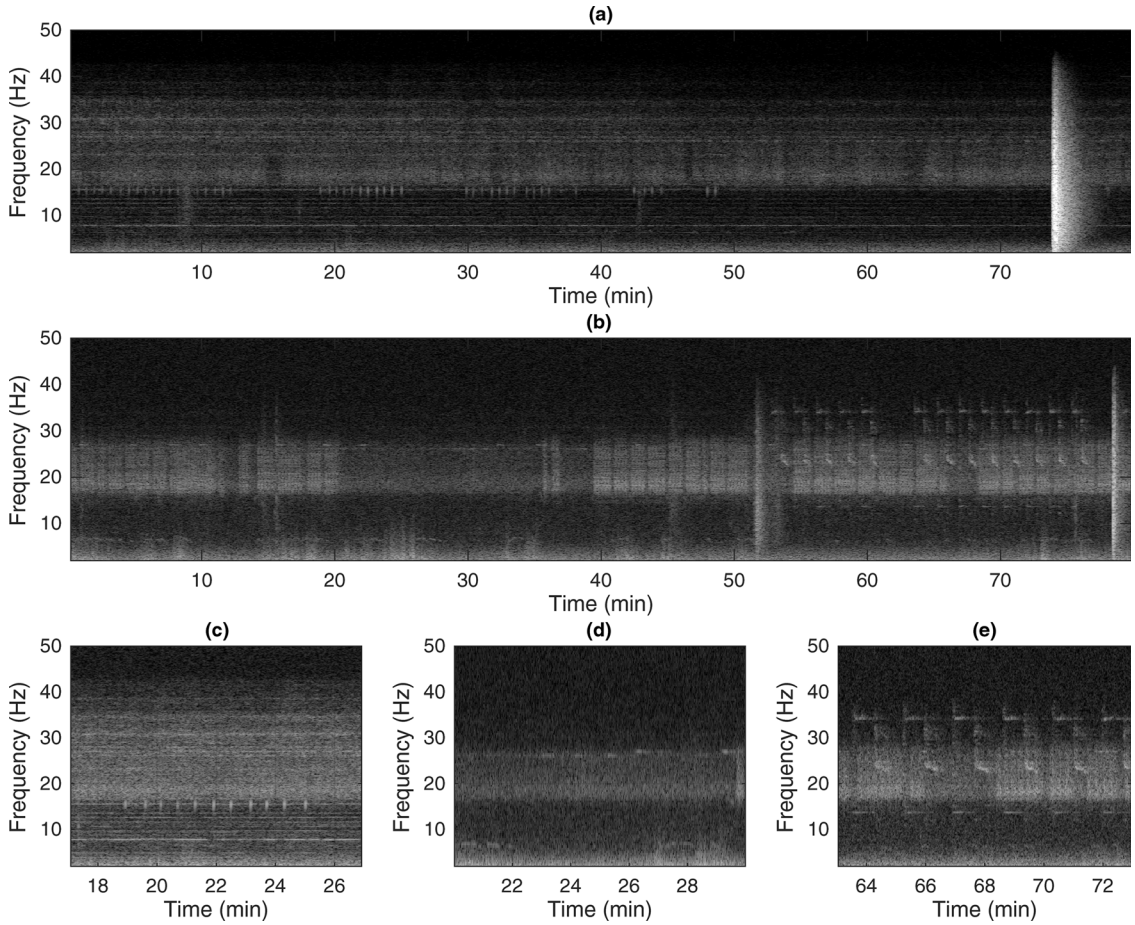

FIG. 2. Soundscape examples (a), (b) and zoomed sections (c)–(e). Records by the station RR47 on May 28th, 2013 represented as gray-scaled spectrograms. Spectrogram parameters (a), (b): Hanning window, $w = 2048$, overlap = 98%. (a) Anthropogenic dominated soundscape at 06.00. (b) Biological dominated soundscape at 15.30. (c) Zoomed section of spectrogram (a) around 22 min ($w = 1024$, overlap = 98%)—Highlights: ship noise and fin whale backbeats (14–16 Hz). (d) Zoomed section of spectrogram (b) around 25 min ($w = 256$, overlap = 98%)—Highlights: Z-call unit A ($\simeq$26 Hz), P-calls (27 Hz) and chorus. (c) Zoomed section of spectrogram (b) around 68 min ($w = 512$, overlap = 98%)—Highlights: Madacascan pygmy blue whale calls, fin whale pulses and chorus.

50 Hz. Due to the passage of a ship near the SWIR array, anthropogenic sound sources are predominant in Fig. 2(a). However, mostly biological sound sources are displayed in Fig. 2(b). To ease spectrogram understanding, soundscapes elements are sorted here between two categories: continuous or short-duration sounds.

1. Continuous sounds

A wideband noise occurs between 16 and 27 Hz on both soundscape spectrograms [Figs. 2(a) and 2(b)]. It is called the chorus effect, the contribution of remote whales' songs.[6] It is most likely due to baleen whale calls, e.g., fin whales pulsed calls and ABW Z-calls.[9]

Furthermore, ship radiated tonal noise is present for the total observation duration [Fig. 2(a)], denoted by an acoustic frequency comb[14] [Fig. 2(c)]. One of the tonal frequencies at $\simeq$ 26 Hz, occurs at the same frequency than Z-call unit A, detailed in Sec. II C.

2. Short-duration sounds

Several short-duration sounds with different characteristics appear on both soundscapes. Most of them can be attributed to baleen whales wandering in the area.

Several ABW Z-call unit A (more detailed in Sec. II C) appear on both observations at $\simeq$26 Hz, all along Fig. 2(a), mixed with one of the ship tonals. They also appear in Fig. 2(b) between $\simeq$5 and 30 min [zoomed—(d)] and probably around 50 and 63 min. Unknown caller "P-call"[9,25] at 27 Hz is also present in Figs. 2(b) and 2(d).

Fin whale song units are also visible on both soundscapes. They are described as short-duration pulses: the most common, the 20 Hz-pulse usually covers the [13–31.5] Hz frequency band. Remote fin whale 20 Hz pulses are visible between 16 and 28 Hz[9,26] on both soundscapes, but mostly on Fig. 2(b) before 20 min and starting again at 40 min. On Fig. 2 zoomed section (c), another fin whale unit occurs on a smaller frequency scale in the 14–16 Hz frequency band, denoted as fin whale "backbeats."[27]

Another short-duration sound is attributed to blue whales. Madacascan pygmy blue whale type call starts at 52 min on Fig. 2(b) [zoomed section (e)]. They are composed of two long units. The first one is composed of 3 to 4 simultaneous tonals between $\simeq$13.5 and 34.0 Hz, with $\simeq$7 Hz intervals. Most of the energy of the second unit is concentrated on a "triple down sweep" from $\simeq$24.4 to 21.6 Hz with a maximum at $\simeq$23.25 Hz.[9]

Furthermore, different broadband noises occur in Fig. 2(a) at 9 min, 43 min, and with higher energy from 73 to 77 min as well as on Fig. 2(b) at 14.6, 15.7, and 45.5 min and with higher energy between 51.4 and 53.6 min and 78.3 and 79.8 min. They are due to the frequent micro-seismic and seismic events that where originally studied in the area by the RHUM-RUM project.

C. ABW call songs and call description

The ABW song is described using blue whale songs specific nomenclature.[3,9] Its call is named Z-call because of its recognizable Z-shape in the time-frequency domain. It is constituted of three short successive units. Unit A is a slightly modulated tone lasting $\simeq$12 s with a maximum at 26.2 Hz, followed by unit B, a 2 s frequency-modulated down-sweep that joins units A and C. Unit C is also a slightly modulated tone lasting 12.2 s with a maximum at 18.7 Hz. They are regularly repeated in series with an inter-call interval of 66.4 ± 0.4 s. Series are separated by longer intervals of 206.4 ± 9.9 s corresponding to the breathing of the animal.[9]

Two observations of 15 min from May 31st, 2013 recorded on OBS RR43 are used to plot Fig. 3. Figure 3(a) is the time-frequency representation of a complete series of seven high SNR ABW calls followed by three calls of the beginning of the following sequence. This record starts at 12:22, when the whale is estimated to be $\simeq$5 km away from the sensor by the tracking data (Fig. 1). Because the ABW is singing continuously for several hours, its song can be recorded even long after its last measured location plotted on Fig. 1. Figure 3(b) is the time-frequency representation of several remote and low SNR Z-calls for which only unit A remains. It was recorded on OBS RR43 at 21:52. Assuming constant swim speed, at that time the ABW is considered $\simeq$ 100 km away in the Northwest direction. Figure 3(c) is the comparative representation of both observations power spectrum densities (PSD), normalized by the high SNR PSD maximum between 15 and 30 Hz. For the high SNR observation [PSD of Fig. 3(a)], despite the presence of fin whale pulses between 15 and 35 Hz, the Z-call spectral content stands out of the PSD. There is a 16 dB difference between units A and C, underlining that unit A conveys more energy than the rest of the call. With the distance, unit A loses 27 dB. Under the assumption that the propagation answer to linear acoustics, unit C would be at about −43 dB on the PSD. Yet, instead of the attenuated peak, there is a bump, likely due to the effect of remote singing fin whales chorus.[9,26,27] The 27 Hz tonal on the remote observation is generated by P-calls.

III. DETECTION STRATEGY

The detection strategy is based on the SMF, an extension of the MF where the signal of interest is never perfectly known. This filter-based denoising method[28] enhances stochastic signals embedded in an additive colored noise by maximizing the filter's output SNR.[29] Unlike matched filtering methods, where the correlation is the optimum for the detection of a known signal in white Gaussian noise,[30] these assumptions are well suited for detection in a passive context.

The SMF has been used for detection in the time domain in diverse fields such as modulated wideband signal detection in active sonar[31,32] or audio pattern detection in automatic speech recognition.[33] Another formulation of the SMF as a time-varying linear filter provides faster *online* data processing and significant representations of the method's filters.[32] Recently, the SMF has been used for sperm whale click and echo detection[28] that even outperforms the Teager-Kaiser-Mallat filter method.[16] To continue investing the potential of the SMF as a passive acoustic monitoring tool, it is explored for ABW call detection. First results and real-data application are presented in Ref. 13, but with no performance assessments.

The SMF complete method is carefully detailed in Ref. 28, where the time-varying linear filter approach, chosen for our passive context application is described. However, Sec. III A presents a general overview of the method's keypoints, to introduce a two stages approach: first, the *offline* computation of the filter bank (Sec. III B) and then, the *online* application where the denoising occur (Sec. III C).

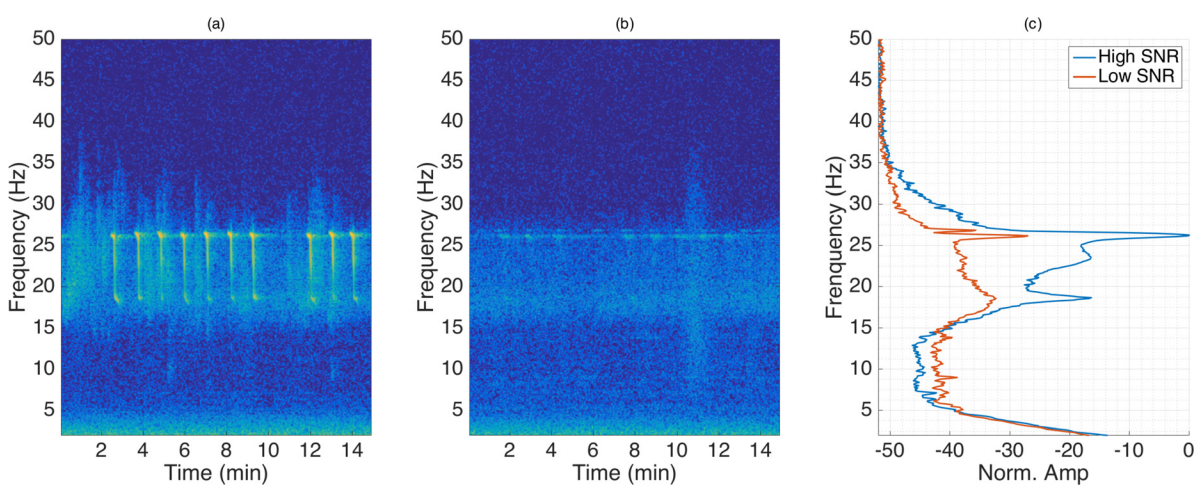

FIG. 3. (Color online) ABW call and propagation effects. Records from May 31st, 2013 on OBS RR43. Spectrogram parameters: Hanning window, $w = 1024$, overlap = 98%. (a) 12:22—High SNR. (b) 21:52—Low SNR. (c) Comparative PSD normalized by the high SNR PSD maximum between 15 and 30 Hz.

A. SMF overview

In the discrete time domain, samples are denoted by the index $i \in \{1, 2,\ldots,I\}$ and the observation Z is composed of a signal of interest S embedded in the additive noise N. Both S and N are supposed to be realizations of random functions whose second order statistics are known as covariance matrices. Using the signal and noise variances, respectively, σ_S^2 and σ_N^2, the observation is written as $Z = \sigma_S S_0 + \sigma_N N_0$, with $\mathrm{E}\{S_0^2\} = 1$ and $\mathrm{E}\{N_0^2\} = 1$. S_0 and N_0 are both assumed centered, second-order stationary and mutually independent.[32]

The SMF theory[29] shows that it is possible to expand Z into a sum of known vectors Ψ_i weighted by uncorrelated random variables z_i such as $Z = \sum_{i=1}^{I} z_i \Psi_i$ and where $\{\Psi_i\}_{i=1,2,\ldots,I}$ is a I-dimensional deterministic basis. Therefore, depending on the choice of basis vectors, some random variables are carrying more signal than noise. It is possible to find the appropriate basis vectors to improve the output SNR of the filter.

The random variables z_i are determined using the scalar product between Z and deterministic vectors Φ_i such as $z_i = Z^{\mathrm{T}}\Phi_i$. The basis $\{\Phi_i\}_{i=1,2,\ldots,I}$ (of I-dimensional deterministic vectors) that ensures the maximization of the SNR is solution to the generalized eigenvalue problem that links the respective signal and noise reduced covariance matrices $\Gamma_{S_0S_0}$ and $\Gamma_{N_0N_0}$. It is written

$$\Gamma_{S_0S_0}\Phi_i = \lambda_i \Gamma_{N_0N_0}\Phi_i, \tag{1}$$

with λ_i and Φ_i the eigenvalues and associated eigenvectors. It is possible to show[28,32] that the output SNR ρ_{out_i} directly depends on the product between λ_i and the input SNR, represented by the power ratio σ_S^2/σ_N^2 : $\rho_{out_i} = \lambda_i(\sigma_S^2/\sigma_N^2)$.

In practice, to overcome the stationary issue of the noise, the observation is processed through an odd sliding window Z_k of size K, centered on the kth sample, denoted $Z[k]$. An approximation of the signal of interest can be reconstructed by keeping only $Q[k]$ components associated to the eigenvalues greater than a certain threshold such as

$$\tilde{S}_{Q[k]}[k] = \sum_{i=1}^{Q[k]} z_{i,k}\Psi_i\left[\frac{K+1}{2}\right], \tag{2}$$

with $z_{i,k} = Z_k^T\Phi_i$ and where $Q[k]$ is the dimension of the basis $\{\Psi_i\}_{i=1,2,\ldots,I}$ that minimizes the mean square error between the signal of interest and its approximation. It corresponds to the number of eigenvalues λ_i times the input SNR of the kth sample ρ_k greater than one[28] such as

$$Q[k] = \#(\rho_k\lambda_i \geq 1). \tag{3}$$

The approximation of the signal of interest can be written for each sliding window's center sample as the product between the Z_k observation and a $Q[k]$-dimensioned filter such as

$$\tilde{S}_{Q[k]}[k] = Z_k^{\mathrm{T}} h_{Q[k]}, \tag{4}$$

the filter being expressed as

$$h_{Q[k]} = \sum_{i=1}^{Q[k]} \Psi_i\left[\frac{K+1}{2}\right]\Phi_i, \tag{5}$$

and where $\{\Psi_i\}$ and $\{\Phi_i\}$ are K-dimensional basis. $Q[k]$ taking values between 1 and K, Equation (5) allows one to compute K vectors h_Q $(1 \leq Q \leq K)$, from h_1 ensuring the maximization of the SNR, to h_K whose bandwidth corresponds to the whole useful signal bandwidth[28] (Sec. III B 2). Equation (5), shows that it is possible to compute *offline* a h_Q filter bank, depending only on $\{\Psi_i\}$ and $\{\Phi_i\}$ basis estimation, that therefore relies on *a priori* knowledge of the signal and noise covariances. This *offline* stage is developed in Sec. III B, to calculate the filters for ABW call denoising and detection. The *online* stage that consists in finding $Q[k]$ for each observation window center sample, and to apply the proper number of filters is dealt with in Sec. III C.

B. Offline SMF

The *offline* stage of the SMF aims to generate a filter bank, matching the signal and adaptable to any type of noise. To solve the generalized eigenvalue problem Eq. (1), the signal covariance matrix is required (Sec. III B 1). The noise covariance matrix is calculated over a classical simulated underwater colored noise, with higher energy at low frequencies. The resulting filter bank is presented in (Sec. III B 2).

1. Signal's covariance matrix estimation

The SMF requires an accurate knowledge of the signal's covariance matrix which provides insights on the variation between samples. Computing $\Gamma_{S_0S_0}$ requires the signal's probability density function. It is obtained either by the estimation of an acoustic pattern from a dataset of several high SNR representative signals[5,11] or spectrum modeling using a mixture Gaussian model.[4,34] For signal with a known instantaneous frequency, it is also possible to realize a corresponding parametric model that describe the temporal variations of the amplitude and phase.[6,17]

Here, for ABW call detection, the parametric model described in Ref. 17 is used [Eq. (6)]. It is based on the complex form of an acoustic signal $s(n) = a(n)e^{j\varphi(n)}$, with $a(n)$ the time-varying amplitude and $\varphi(n)$ the time-varying phase. From the definition of the instantaneous frequency and its parametric expression as a function of the (continuous) time

$$f(t) = f_c + \frac{1}{2\pi}\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} = f_c + L + \frac{U-L}{1+e^{\alpha(t-M)}}, \tag{6}$$

it is possible to derive the expression of the time-varying phase $\varphi(n)$, where n denotes the discrete time, as

$$\varphi(n) = 2\pi \quad L\frac{n}{f_s} + \frac{U-L}{\alpha}\ln\left(\frac{1+e^{-\alpha M}}{1+e^{\alpha((n/f_s)-M)}}\right)\Bigg) + \varphi_0, \tag{7}$$

where $f_c = 22.6$ Hz is the central frequency in the[15–30] Hz bandwidth, L and U are, respectively, linked to the lower and upper asymptotes of the Z-call, M represents the time shift and α the grow rate. The amplitude $a(n)$, is set to vary in accordance with the energetic difference between unit A and C (Sec. II C). To compensate observed annual and

seasonal frequency variations within the call, the signal is built as a summation of several frequency modulated signals with the following parameters: $f_s = 100\,\text{Hz}$, $T_{Zcall} = 20\,\text{s}$, $L = [-4.5;\ -4;\ -3.5]\,\text{Hz}$, $U = [3.2;\ 3.6;\ 4]\,\text{Hz}$, $M = [T_{Zcall}/2;\ (T_{Zcall} + 0.5)/2;\ (T_{Zcall} + 1)/2]$ s, and $\alpha = 1.8$. This temporal signature is then used to compute the signal's reduced covariance matrix $\Gamma_{S_0S_0}$. The dimension of the signal is also chosen to be the size of the Z_k sliding window: therefore $K = 2001$ bins.

2. Filter bank

The need here is to generate a filter bank that matches the signal and is applicable at all time without any *a priori* knowledge on the noise instantaneous variations. It is realized, by computing the filter bank using a synthetic sea-colored noise. Those filters are saved and stored for further use in the *online* application of the SMF. As seen in Sec. III A, it is $Q[k]$ that determines *online* the number of filters to apply. It depends on the eigenvalues and the current SNR (Sec. III C).

Figure 4 illustrates the frequency response of filters h_1, h_{10}, and h_{Qmax}, where h_Q denotes the superposition of the Q-first filters of the optimal linear filter bank for ABW call detection and capital H_Q denotes their spectral representation. They are compared to the spectral representation of the signal's reference, the optimal filter, HOpt. To maximize ρ_{out}, the first filter H_1 is a short-band filter centered on the most energetic component of the call, unit A at ~26.3 Hz. It is applied when the estimation of ρ_k indicates there is no signal and realizes "noise cancellation." The superposition of the first ten filters H_{10} leads to two slightly larger band-pass filters, respectively, centered on units A (~26.3 Hz) and C (~18.6 Hz) of the call. H_{Qmax} represents the superposition of the maximum number of filters that is applied when the estimated input ρ_k (Sec. III C 2) is high enough. The filter's pattern is then close to HOpt but when applied, it band-pass filters the observation in the exact signal frequency band: it realizes the full signal reconstruction.

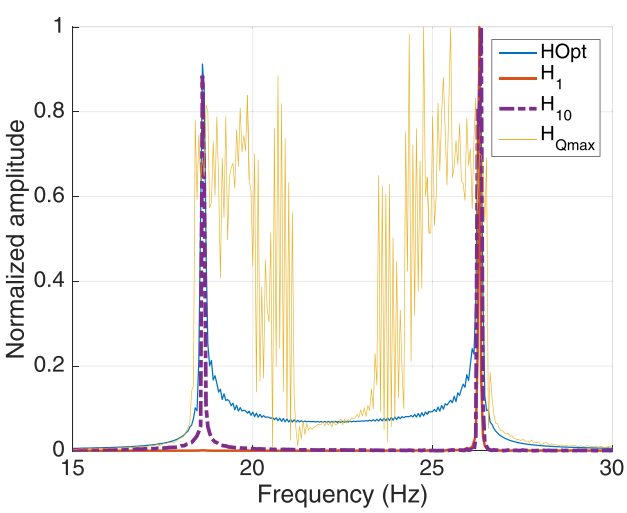

FIG. 4. (Color online) Spectrum of three filters (H_1, H_{10}, H_{Qmax}) of the permanent filter bank compared to the spectral representation of the reference signal (HOpt).

C. Online SMF

The *online* stage of the SMF is where the reconstruction and the detection occur. The observation background noise is first estimated using time-frequency analysis (Sec. III C 1). It provides both the real noise covariance matrix and a noise reference for the *online* input SNR estimation (Sec. III C 2). Knowing all the inputs, the signal and noise covariance matrices [Eq. (1)], it is possible to solve the generalized eigenvalue problem. For each sample of the observation $Z[k]$, it is then possible to evaluate $Q[k]$, and apply the proper number of filters to the current observation.

1. Online noise's covariance matrix estimation

The SMF has been used for the passive acoustic detection of whales clicks.[15,28] In those studies, the noise covariance matrix is estimated on annotated signal free samples. Although it is quite convenient for supervised detection on relatively short records, it becomes tedious and unpractical for automatic detection on several-hours-long passive acoustic monitoring datasets with highly varying background noise. Consequently, the following development deals with finding a way to blindly take into account noise variations despite frequency dependence and high energy events occurrence, to perform accurate estimation of $\Gamma_{N_0N_0}$ even in presence of the signal of interest S.[13]

The strategy for the noise's covariance matrix estimation relies on time-frequency analysis. As for a spectrogram, the observation is segmented in time using weighted overlapping windows. The PSD is calculated for each one of the observation's time segment. This representation of the observation in the time-frequency domain is denoted $\gamma_{ZZ}(k',f)$. An odd-lengthen median filter is then applied through time on each frequency canal of the $\gamma_{ZZ}(k',f)$. Its duration is longer than a Z-call duration in the time-frequency representation. This median filter temporally smoothen outliers (impulsive events of duration shorter than the median filter size) effects for each frequency band, such as the presence of whale calls or short seismic events. The median filter having any effect on the overall observation level, this step provides a time-frequency estimate of the background noise, denoted $\gamma_{\bar{N}\bar{N}}(k',f)$. Then, the Wiener-Khintchine theorem that links the PSD and the autocorrelation of the signal is applied, to estimate *online* the noise covariance matrix $\Gamma_{N_0N_0}$. It is used to solve *online* the generalized eigenvalue problem and find the eigenvalues λ_i.

2. Time-dependent SNR estimation

The estimation of the observation's current input SNR ρ_k, is essential for the *online* application of the SMF [Eq. (3)]. This time variable has a strong impact on the calculation of $Q[k]$ that indicates the number of filters to apply to the kth sample. Its classical definition is the ratio between the instant power of the signal and the noise's. In practice, as the power of the received signal of interest is never known, it is usually estimated as the difference between the instant power of the observation and the noise's ($\sigma_{S_k} = \sigma_{Z_k} - \sigma_{N_k}$).[32] However, passive

acoustic monitoring records can be noisy (Sec. II B) and therefore, there is a need for a more accurate ρ_k estimation.

To sharpen its estimation when dealing with real underwater noises and decrease the false alarm rate, the strategy is again to use the time-frequency representation of the observation. $\gamma_{ZZ}(k',f)$ is compared $\gamma_{\tilde{N}\tilde{N}}(k',f)$ to provide valuable information of the energetic variations in the known signal's bandwidth. Yet, other signal of no interest might occur in the same bandwidth (other biological sources, distant ship noise or transient noise), so the previous calculation is compared to the nearest frequency bands.[35]

For ABW Z-call detection investigated in this work, the ρ_k estimation strategy results in three steps, using the previously computed time-frequency estimate of the PSD $\gamma_{ZZ}(k',f)$ and the noise's $\gamma_{\tilde{N}\tilde{N}}(k',f)$ (Sec. III C 1) described by the following algorithm on each time segment (denoted by k').

Step 1: The signal presence is evaluated in the Z-call unit A frequency band ($A = [25.5–26.5]$ Hz) by the absolute ratio between the maximum value of the observation's PSD and the mean value of the estimated background noise,

$$zcall_{k'} = \frac{\max_f \{\gamma_{ZZ}(k',f)\}}{\overline{\gamma_{\tilde{N}\tilde{N}}(k',f)}}, \quad f \in A. \tag{8}$$

When this is ratio greater than 1, it indicates the presence of a short-duration signal in unit A frequency band but does not differentiate signal from "non-signal" short-duration events.

Step 2: The false alarm due to energetic transient wide band noises is estimated by measuring the ratio between the observation PSD and the estimated background noise, outside the Z-call frequency band ($f \in [0, 15[\cup]27, f_s/2]$).

$$trans_{k'} = \max_f \left| \frac{\gamma_{ZZ}(k',f)}{\gamma_{\tilde{N}\tilde{N}}(k',f)} \right|. \tag{9}$$

Step 3: The time dependent SNR ρ_k, is then determined in dB as the ratio between the two previous steps

$$\rho_{k'} = 20 \log \left(\frac{zcall_{k'}}{trans_{k'}} \right). \tag{10}$$

To take into account global changes in the acoustic environment, mostly due to continuous sounds, ρ_k is enhanced with a value β ($\rho_k = \rho_k - \beta$) for whole observation duration, depending on the background noise estimation as

$$\beta = \begin{cases} 0 & \text{if } M > 0, \\ M & \text{elsewise}, \end{cases} \tag{11}$$

with $M = (1/K) \sum_{k=1}^{K} (zcall_k - trans_k)$. To return to real time k, data are interpolated. Only positive values of ρ_k trigger the reconstruction of the observation [Eq. (3)].

IV. PERFORMANCE ANALYSIS

Passive detection algorithm performances assessment is a tricky subject.[36] The formal expression of a detector's performances is derived from the probability of false alarm and the probability of *true* detection information, when they are available.[30] However, in the absence of this knowledge, theoretical analysis of the performances cannot be easily performed. A solution to this problem is to resort to *ad hoc* approaches such as Monte Carlo simulations. It relies on an extensive computation over an artificial dataset, and has been realized for the classic SMF in several publications.[15,28,31] Yet, performance simulations might not be representative of the method's robustness to noise in real conditions. The third option is to confront the algorithm to a ground truth dataset,[17,35,37] keeping in mind that the annotating process has been shown to be subjective and quite variable, especially in low SNR conditions.[36] This is why, choice was made to use additional information from multisensor observation and whale localization to reduce the subjectivity of the annotation process and evaluate the SMF's performances on a robust ground truth dataset. The SMF's performances are assessed on real marine signals by the comparison of the method output with human inspected and annotated data (Sec. IV A). Results are compared to the commonly used temporal MF and to the Z-detector.[17]

A. Groundtruth dataset context

Over the passage of a singing ABW through the SWIR array on May 31st, 2013 (Fig. 1), 845 Z-calls have been annotated, emitted during slightly more than 21 h (01:20–22:40). This dataset, recorded by OBS RR48, is corrupted with different noise sources presented in Sec. II B:

- Tonal noise radiated from a close ship is present between 01:20 and 08:00.
- P-calls are emitted from 20:10 to the end.
- Strong fin whale pulses appeared between 10:40 and 18:40.
- The "chorus"[6] phenomenon is present in the whole dataset.
- More than 50 seismic events occur.

This dataset provides great variations of the SNR from first contact to the last [Fig. 5(a)]. Considering a mean speed swim of 10 km/h, the whale trajectory is extrapolated as range indicator. At 1:30, the ABW is estimated to be 100 km away in the South direction from the OBS RR48 ($\overline{\rho_k} = 1.6$ dB) and 130 km North West at 22:40 ($\overline{\rho_k} = 3.4$ dB). The closest point of approach is measured by the localization at 10:03, and is 35 km away from the sensor ($\overline{\rho_k} = 13.45$ dB).

B. Method

This day-long record is divided by portions of 80 min for the analysis. As it was described in Sec. III B 1, the signal reference for both the SMF and the Z-detector is quite similar. This signal is also used for inter-correlation for the MF and on the SMF-output cleaned data (SMF + MF) for similar comparison. It is a 20 s-long signal of 2001 samples.

For the SMF, $\gamma_{ZZ}(k',f)$ is realized using Hanning window of 2048 samples, with 98% overlapping. The median filter applied to each frequency canal for the background noise estimation has 201 samples (Sec. III C 1). Its equivalent duration is 83 s, four times longer than a Z-call. After

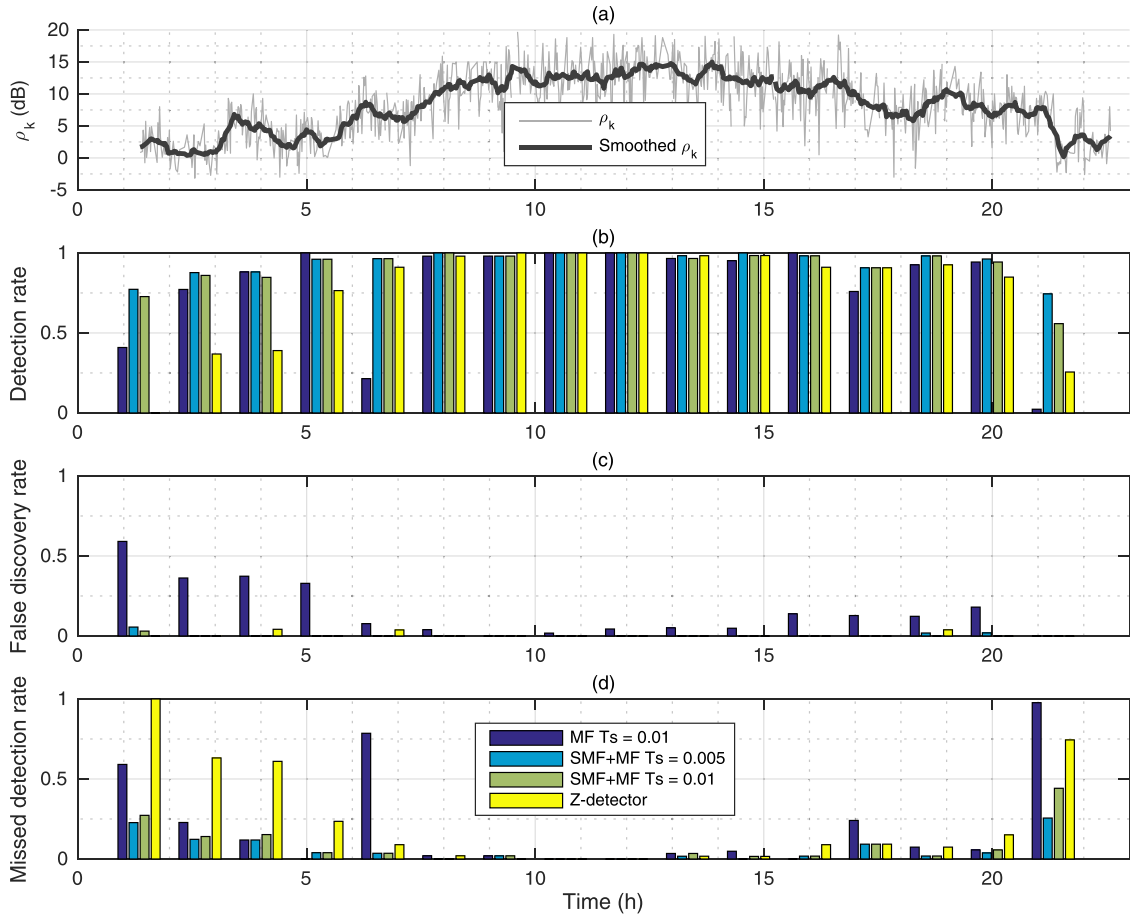

FIG. 5. (Color online) Time-dependent compared performance analysis of MF, SMF + MF, and the Z-detector on a ground truth dataset of 845 annotated calls, on May 31st, 2013. (a) Detected call measured input SNR ρ_k (Sec. III C 2). Performance criterion are measured on 80 min portions of observation containing on average 53 calls (min. 43 to max. 62). (b) DR, (c) FDR, (d) MDR.

retrieving real-time k (Sec. III C 2), the size of the sliding window used to assume the stationarity of the noise is set to match the signal with $K = 2001$ samples. The output is analyzed with two detection thresholds (Ts) set at the output of the correlation (SMF + MF): Ts = 0.01 and Ts = 0.005.

For the MF, the observation is previously filtered by the application of a band-pass filter in the Z-call frequency band (15–30 Hz). The detection threshold is set at Ts = 0.01. As for the SMF + MF, Ts = 0.005 has also been tested. This threshold was to low for detection: most of the MF output were occurring above this threshold, this is why it is not presented here.

Before applying the Z-detector, the observation is down-sampled to a base-band signal for better computation time. The detection threshold of the Z-detector automatically adapts to deal with noise variations in the environment (e.g., transient signal) and allows Z-call variations in frequency, amplitude or duration.[6,17] The threshold τ to separate the signal and the noise is fixed at 0.15% and the worst-case user-defined Probability of false alarm is fixed at 3%, as in the work presented in Ref. 17.

For each annotated Z-call, the maximum of the SMF ρ_k estimation (Sec. III C 2) is measured and the outcome of the detection methods are checked. An arbitrary confidence index assigned by the operator is also added, that reveals to be in accordance with SNR variations (the lower the SNR is, the harder it is to classify an event). Number of false alarms is counted. Different criteria are defined to evaluate the performances.

The detection rate (DR) indicates the rate of detected true calls overall detections

$$\mathrm{DR} = \frac{\text{nb. of true positives}}{\text{nb. of true positives} + \text{nb. of false negatives}}. \tag{12}$$

The missed detection rate (MDR) indicates the rate of true calls that are not detected

$$\mathrm{MDR} = \frac{\text{nb. of false negatives}}{\text{nb. of true positives} + \text{nb. of false negatives}}. \tag{13}$$

The false discovery rate (FDR) indicates the rate of detections that are not calls

$$\mathrm{FDR} = \frac{\text{nb. of false positives}}{\text{nb. of false positives} + \text{nb. of true positives}}. \tag{14}$$

Note that DR + MDR = 1.

C. Results

Figure 5(a) highlights the correspondence between May 31st, 2013 timing and SNR variations introduced by

the whale movement and background noise evolutions. Figure 5(b) presents the estimated DR (12), (c) FDR (14), and (d) MDR (13), per 80 min portions of the observation for the three compared methods, the MF Ts = 0.01, the SMF + MF Ts = 0.005 and Ts = 0.01 and the Z-detector. Figure 5 shows that for all methods is that the DR increases with the measured input SNR ρ_k (when the whale is the closest) while the FDR and the MDR decrease. The MF has higher FDR than the other two methods. The Z-detector has higher MDR (and therefore lower DR) than the SMF + MF or the MF (except on the observation 5:20–6:40 and the last one 21:20–22:40).

Before 08:00, ρ_k is <10 dB and ship noise is present. For the first three portions of observation, the best DR is reached for the SMF + MF for both thresholds, then for the MF. The smallest DR is obtained by the Z-detector, that does not detect any call before 2:40. On the 4th portion starting at 05:20, DR of the MF is greater than the other two methods: the SMF + MF misses a detection occurring at the same time as a seismic event. For the observation starting at 06:40, DR of the MF is really low. It might be due to the change of background noise when the ship noise fades while the whale calls are still quite low: the detection threshold is not adapted. One explanation for the SMF + MF false alarms that occur between 01:20 and 2:40, as well as some of the MF's, is that they are triggered by continuous remaining signal in the unit A band that probably comes from distant calls (Sec. II B) but could not be strictly assigned to one. The other MF false alarms were mostly due to seismic events. The whale is estimated to be 65 km away from OBS RR48 at 05:00 and 130 km away at 22:00. Yet, due to the ship noise present at the beginning of the record, ρ_k is of the same order (Sec. III C 2). Those too low or negative estimations of ρ_k impact not only the SMF + MF, but also the other methods by increasing the number of missed detections.

Between 08:00 and 16:00, the ρ_k is >10 dB and the three methods perform similarly. However, the MF FDR is higher. There are only few missed detections methods likely due to the simultaneous occurrence of Z-call and short larger band noise (fin whale pulse or seismic event).

After 16:00, ρ_k decreases to values smaller than 10 dB. MDR raises as the DR decreases but the methods FDR are still lower than before 08:00: there are lesser continuous sound-sources. However, DR and MDR results are quite similar to the ones before 08:00. For the last observed portion, the MF's really low DR indicates that MF-processed data reach the limit of the signal excess.

It is important to give some perspectives for the analysis of the comparison between the SMF + MF and the Z-detector. Although the SNR definition remains the same between the SMF and the Z-detector (power ratio between the signal and the noise present in the observation), its estimation diverge. As presented in Sec. III C 2, the SMFs input SNR ρ_k is measured continuously in the time-frequency domain, with a parameter to prevent positive values on non-Z-call events. This is made so that in the presence of a call, even incomplete, ρ_k might be measured positive to satisfy Eq. (3). However, for the Z-detector, the input SNR is

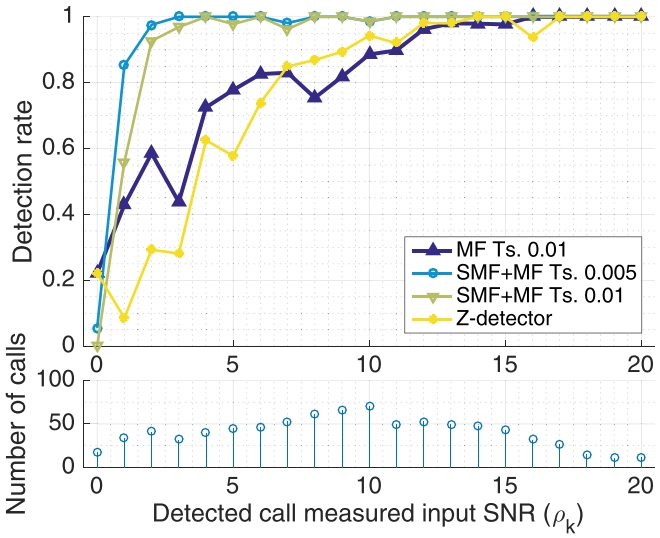

FIG. 6. (Color online) Comparative performance analysis between the MF, the SMF + MF and the Z-detector on real data. (a) DR against measured SNR with two thresholds (Ts = 0.005 and Ts = 0.01) applied on the correlation output. (b) Stacked bar representation of the number of annotated calls per SNR for DR estimation.

measured using the energy of the observation in the baseband representation (15 Hz bandwidth centered around 22.5 Hz) and the diagonal values of the noise covariance matrix:[17] the values can be negative. The SMF-detected calls estimated ρ_k varies between 0 and 20 dB while the Z-detector measured input SNR (for the same detections) varies between −15 and 15 dB.

The Z-detector is designed with an adaptive detection threshold that satisfies a user-chosen "worst" probability of false alarm of 3%.[17] This, might explain the higher MDR when $\rho_k < 10$ dB, compared to the SMF + MF and the MF, where the threshold is fixed and is supported by the results of Fig. 6.

Figure 6(a) displays detection results, as DR against SNR for the MF Ts = 0.01, the SMF + MF Ts = 0.005 and 0.01 and the Z-detector. Figure 6(b) (bar representation) highlights the number of annotated calls on which the DR is estimated. A higher number of annotated calls gives more credit to the associated DR estimation. DR reaches one at a lower SNR for the SMF + MF (≃3 and 4 dB, respectively, for Ts = 0.005 and 0.01) than for the Z-detector (14 dB) or the MF (≃16 dB).

TABLE I. Experimental performances of MF, SMF + MF and Z-detector on a 22 h long noise-corrupted database with 845 annotated Z-calls. Noise types: Boat tonal noise, seismic events, other biological sound sources. Are displayed, the detection rate (DR), the missed detection rate (MDR), and the false discovery rate (FDR). Best results are indicated in bold font.

	Ts	DR	MDR	FDR
MF	0.01	81.4%	18.6%	18.1%
SMF + MF	0.005	**94.1**%	**5.9**%	0.50%
	0.01	92.2%	7.8%	**0.13**%
Z-detector	—	77.6%	22.4%	0.76%

To complete this analysis, Table I presents comparative performances of the methods over the entire dataset. It highlights that the SMF + MF Ts = 0.005 performs great results with the highest DR–smallest MDR (respectively, 94.1% and 5.9%), and with a lower FDR than the other methods (0.5% against 0.76% for the Z-detector and 18.1 for the MF). At an equivalent threshold Ts = 0.01, it still performs better than the MF with higher DR (92.2% against 81.4%), and the best FDR, 0.13%. The choice of the threshold should result from a compromise between the DR and FDR wanted that relies on the context. The Z-detector FDR is still reasonable since <1% and quite close to the SMF + MF Ts = 0.005: the Z-detector has only one more false alarm than SMF + MF over the whole dataset. However, their DR (SMF +MF Ts = 0.005 94.1%, Z-detector 77.6%) and MDR (SMF + MF Ts = 0.005 5.9%, Z-detector 22.4%) are quite different since more than a fifth of the annotated Z-calls are not detected by the Z-detector. Even if it might be due to the set of the Z-detector threshold at a worst probability of false alarm of 3%, the SMF + MF Ts = 0.005 performs better: if the threshold was lowered to reach the same FDR, the DR that is already better than the Z-detector's would increase and the MDR would decrease. Drawn from the results, it is possible to notice that when the ship noise is present and at a fixed DR > 75%, the SMF + MF Ts = 0.005 expands the detection range of the sensor up to ≃90 km against ≃ 60 km for the Z-detector, with an equivalent FDR.

To conclude, because it strongly reduces the noise, the SMF+MF achieves great FDR even with really low thresholds and. It performs better than the MF in real conditions, when dealing with real marine signals corrupted by noises. It is more robust than the MF and the Z-detector in low SNR conditions, improving the detection range of ≃30 km, compared to the other methods, with an equivalent FDR.

V. DISCUSSION AND PERSPECTIVES

The method's limitations are reached when ρ_k is estimated equal to zero or negative [Eq. (3)]. In that case the SMF applies the first filter and therefore the observation is not considered as signal. Due to the definition of the estimation of ρ_k used Eq. (10), it might occur in different context. If the call is completely embedded in short-duration but same bandwidth noise, e.g., seismic event or high intensity fin whale pulses, the noise estimation will overcome the estimated presence of the call. Yet, this combination of events is quite unusual and might be hard to detect even by an experimented analyst. The second source of error occurs on highly remote calls. As discussed in Sec. II C, only unit A remains from long distance propagation. The emitted calls follow multiple-path spreading that are denoted on recorded data by multiple echoes. In some circumstances, it leads to an almost uninterrupted signal in the unit A frequency band, where the dissociation of singular calls is complicated, even for an automatic detection algorithm. It might lead to missed detections.

Overall, in passive contexts, it is very difficult to control the content of recorded data. Therefore, the evaluation of detection algorithm performances has to be either assessed using simulations or confronting the detector to a ground truth datasets. The first one, related to the detection theory, provides a probabilistic approach of the method's performances and satisfies classic signal processing techniques. Yet, it is often not representative of the method's robustness against noise, to the detection of degraded signals etc. Ground truth datasets provide a large variety of observations and set of events that would be hard to recreate. But, data annotation has been shown to be subjective and highly variable between data analysts, and even one cannot always be consistent.[36] It would be a great asset to create a common open-source dataset for algorithm training and testing, and then be able to compare detection algorithm performances to one another.

The use of the SMF + MF increases greatly the detection range, by providing robust detection performances up to a 130 km radius (when no ship noise and 90 km when ship noise is present): there might be several acoustically active ABW in such an area. Two different approaches exist to attribute one song to the right whale. The first one is to separate the recorded songs, as in blind source separation. However, for the moment, no call differences between two blue whales of the same sub-species have been noticed. The other is to associate a song to a trajectory. This is the aim of our future work, by automatically measuring signal's times of arrival using the SMF, to feed the input database of detection algorithm.[9]

VI. CONCLUSION

In this work, an extended version of the SMF for SNR maximization and detection in a passive context is presented. The method, originally based on the MF, is derived from the second order statistics of the signal and the noise. It can be seen as a time-varying linear filter adjusted in accordance with the estimation of the input SNR. The background noise and the time varying SNR are estimated using time-frequency analysis. For ABW call detection, the signal's covariance matrix is derived from a predefined parametric model. Performances are assessed using a ground truth dataset containing 845 calls, spread on more than 22 h, recorded over the passage of a singing ABW in the vicinity of an OBS array in the southwest Indian Ocean. This record presents great variations of SNR and is quite representative of region's sounds diversity in the 0–50 Hz bandwidth. The combination SMF + MF outperforms the MF on low SNR observations: the detection rate is higher, the number of missed detections is lower, and the number of false alarms is drastically decreased. When dealing with ship-noise corrupted data, the SMF + MF also provides a greater detection range than the Z-detector (+30 km), for a detection rate >75% and with an equivalent false discovery rate.

ACKNOWLEDGMENTS

The authors would like to thank the Z-detector team,[17] especially François-Xavier Socheleau, Emmanuelle Leroy, and Maëlle Torteroto, for their help with the method's performance comparison and their great advices. They also

want to thank the reviewers for their thorough reading and their relevant remarks that helped improve the manuscript.

[1]T. A. Branch, K. M. Stafford, D. M. Palacios, C. Allison, J. L. Bannister, C. L. K. Burton, E. Cabrera, C. A. Carlson, B. Galletti Vernazzani, P. C. Gill, R. Hucke-Gaete, K. C. S. Jenner, M.-N. M. Jenner, K. Matsuoka, Y. A. Mikhalev, T. Miyashita, M. G. Morrice, S. Nishiwaki, V. J. Sturrock, D. Tormosov, R. C. Anderson, A. N. Baker, P. B. Best, P. Borsa, J. R. L. Brownell, S. Childerhouse, K. P. Findlay, T. Gerrodette, A. D. Ilangakoon, M. Joergensen, B. Kahn, D. K. Ljungblad, B. Maughan, R. D. McCauley, S. McKay, T. F. Norris, S. Rankin, F. Samaran, D. Thiele, K. Van Waerebeek, and R. M. Warneke, "Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the southern hemisphere and northern Indian ocean," Mammal Rev. **37**(2), 116–175 (2007).

[2]IWC, "Chairman's report of the thirty-first annual meeting," Rep. Int. Whaling Comm. **30**, 25–41 (1980).

[3]M. A. McDonald, S. L. Mesnick, and J. A. Hildebrand, "Biogeographic characterization of blue whale song worldwide: Using song to identify populations," J. Cetacean Res. Manag. **8**(1), 55–65 (2006).

[4]F. Samaran, C. Guinet, O. Adam, J.-F. Motsch, and Y. Cansi, "Source level estimation of two blue whale subspecies in southwestern Indian Ocean," J. Acoust. Soc. Am. **127**(6), 3800–3808 (2010).

[5]N. E. Balcazar, J. S. Tripovich, H. Klinck, S. L. Nieukirk, D. K. Mellinger, R. P. Dziak, and T. L. Rogers, "Calls reveal population structure of blue whales across the southeast Indian Ocean and southwest Pacific Ocean," J. Mammal. **96**(6), 1184–1193 (2015).

[6]E. C. Leroy, F. Samaran, J. Bonnel, and J.-Y. Royer, "Seasonal and diel vocalization patterns of Antarctic blue whale (*Balaenoptera musculus intermedia*) in the southern Indian Ocean: A multi-year and multi-site study," PLoS One **11**(11), e0163587 (2016).

[7]R. A. Dunn and O. Hernandez, "Tracking blue whales in the eastern tropical Pacific with an ocean-bottom seismometer and hydrophone array," J. Acoust. Soc. Am. **126**(3), 1084–1094 (2009).

[8]D. Harris, L. Matias, L. Thomas, J. Harwood, and W. H. Geissler, "Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic," J. Acoust. Soc. Am. **134**(5), 3522–3535 (2013).

[9]R. Dréo, L. Bouffaut, E. Leroy, G. Barruol, and F. Samaran, "Baleen whale distribution and seasonal occurrence revealed by an ocean bottom seismometer network in the western Indian Ocean," Deep Sea Res. Part II: Top. Stud. Oceanogr. in press (2018), available at https://www.sciencedirect.com/science/article/pii/S0967064517303995.

[10]S. D. Frank and A. N. Ferris, "Analysis and localization of blue whale vocalizations in the Solomon Sea using waveform amplitude data," J. Acoust. Soc. Am. **130**(2), 731–736 (2011).

[11]K. M. Stafford, D. R. Bohnenstiehl, M. Tolstoy, M. Chapp, D. K. Mellinger, and S. E. Moore, "Antarctic-type blue whale calls recorded at low latitudes in the Indian and eastern Pacific Oceans," Deep Sea Res. Part I: Oceanogr. Res. Pap. **51**(10), 1337–1346 (2004).

[12]F. Samaran, O. Adam, and C. Guinet, "Detection range modeling of blue whale calls in Southwestern Indian Ocean," Appl. Acoust. **71**(11), 1099–1106 (2010).

[13]L. Bouffaut, R. Dréo, V. Labat, A.-O. Boudraa, and G. Barruol, "Antarctic blue whale calls detection based on an improved version of the stochastic matched filter," in *EUSIPCO 2017*, Greece (2017), pp. 2383–2387.

[14]W. S. D. Wilcock, K. M. Stafford, R. K. Andrew, and R. I. Odom, "Sounds in the ocean at 1–100 Hz," Ann. Rev. Marine Sci. **6**, 117–140 (2014).

[15]F. Bénard, H. Glotin, and P. Giraudet, "Highly defined whale group tracking by passive acoustic stochastic matched filter," in *Advances in Sound Localization* (InTech, London, 2011).

[16]F. Caudal and H. Glotin, "Stochastic matched filter outperforms Teager-Kaiser-Mallat for tracking a plurality of sperm whales," IEEE PASSIVE 2008 (2008), pp. 1–9.

[17]F.-X. Socheleau, E. Leroy, A. Carvallo Pecci, F. Samaran, J. Bonnel, and J.-Y. Royer, "Automated detection of Antarctic blue whale calls," J. Acoust. Soc. Am. **138**(5), 3105–3117 (2015).

[18]"RHUM-RUM web page," www.rhum-rum.net (Last viewed 3/15/2018).

[19]G. Barruol and K. Sigloch, "Investigating La Réunion hot spot from crust to core," EOS, Trans. Am. Geophys. Union **94**(23), 205–207 (2013).

[20]G. Barruol, K. Sigloch, and R.-R. group, "RHUM-RUM experiment, 2011-2015, code YV (Réunion Hotspot and Upper Mantle–Réunion's Unterer Mantel), funded by ANR DFG, CNRS-INSU, IPEV, TAAF, instrumented by DEPAS, INSU-OBS, AWI and the Universities of Muenster, Bonn, La Réunion," (2017), doi: 10.15778/RESIF.YV2011.

[21]"Rhum-rum dataset," http://dx.doi.org/10.15778/RESIF.YV2011 (Last viewed 3/15/2018).

[22]"RESIF," http://seismology.resif.fr (Last viewed 3/15/2018).

[23]S. C. Stähler, K. Sigloch, K. Hosseini, W. C. Crawford, G. Barruol, M. C. Schmidt-Aursch, M. Tsekhmistrenko, J.-R. Scholz, A. Mazzullo, and M. Deen, "Performance report of the RHUM-RUM ocean bottom seismometer network around La Réunion, western Indian Ocean," Adv. Geosci. **41**, 43–63 (2016).

[24]B. C. Pijanowski, L. J. Villanueva-Rivera, S. L. Dumyahn, A. Farina, B. L. Krause, B. M. Napoletano, S. H. Gage, and N. Pieretti, "Soundscape ecology: The science of sound in the landscape," BioScience **61**(3), 203–216 (2011).

[25]E. C. Leroy, F. Samaran, J. Bonnel, and J.-Y. Royer, "Identification of two potential whale calls in the southern Indian Ocean, and their geographic and seasonal occurrence," J. Acoust. Soc. Am. **142**(3), 1413–1427 (2017).

[26]W. A. Watkins, P. Tyack, K. E. Moore, and J. E. Bird, "The 20-Hz signals of finback whales (*Balaenoptera physalus*)," J. Acoust. Soc. Am. **82**(6), 1901–1912 (1987).

[27]B. Koot, J. K. Ford, D. Hannay, and A. W. Trites, "Fin whales (*Balaenoptera physalus*) in British Columbia sing a consistent song," J. Acoust. Soc. Am. **134**(5), 3988–3988 (2013).

[28]P. Courmontagne, "The stochastic matched filter and its applications to detection and de-noising," in *Stochastic Control* (InTech, London, 2010).

[29]J.-F. Cavassillas, "Stochastic matched filter," Int. Conf. Sonar Sign. Process. **13**, 194–199 (1991).

[30]H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory* (Wiley, New York, 2002), p. 578.

[31]J.-L. Mori and P. Gounon, "The use of stochastic matched filter in active sonar," in *Signal Processing Conference, 2000 10th European* (IEEE, Piscataway, NJ, 2000), pp. 1–4.

[32]P. Courmontagne, G. Julien, and M. Bouhier, "An improvement to the pulse compression scheme," in *OCEANS 2010 IEEE—Sydney* (2010), pp. 1–5.

[33]J. Bonnal, P. Danes, and M. Renaud, "Detection of acoustic patterns by stochastic matched filtering," in *International Conference on Intelligent Robots and Systems* (2010), pp. 1970–1975.

[34]F. Samaran, O. Adam, J.-F. Motsch, and C. Guinet, "Definition of the Antarctic and pygmy blue whale call templates. Application to fast automatic detection," Can. Acoust. **36**(1), 98–103 (2008), available at: https://jcaa.caa-aca.ca/index.php/jcaa/article/view/1997.

[35]A. N. Gavrilov and R. D. McCauley, "Acoustic detection and long-term monitoring of pygmy blue whales over the continental slope in southwest Australia," J. Acoust. Soc. Am. **134**(3), 2505–2513 (2013).

[36]E. C. Leroy, K. Thomisch, and I. V. Opzeeland, "Variability in ground-truth data sets and the performance of two automated detectors for Antarctic blue whale calls in different soundscape conditions," J. Acoust. Soc. Am. **141**(5), 3604–3605 (2017).

[37]D. K. Mellinger and C. W. Clark, "Recognizing transient low-frequency whale sounds by spectrogram correlation," J. Acoust. Soc. Am. **107**(6), 3518–3529 (2000).