Highly defined whale group tracking by passive acoustic Stochastic Matched Filter

Frédéric Bénard-Caudal and Hervé Glotin

Systems & Information Sciences Laboratory (LSIS) - UMR CNRS 6168,

University of South-Toulon-Var BP 20132 83957 La Garde Cedex France

Email: caudal@univ-tln.fr; glotin@univ-tln.fr

Abstract

This paper compares two real-time passive underwater acoustic methods to track multiple emitting whales using at least four omnidirectional hydrophones. Passive acoustic localization permits to study whales' behavior in deep water (hundreds of meters) without inferring with the environment. For this purpose, we show that the Stochastic Matched Filter (SMF) allows to be completed by an echo removal and then generates higly defined whale tracks. It is compared to the Teager-Kaiser-Mallat (TKM) filter method. Firstly, we briefly review the SMF and TKM theories. Then we describe how rough Time Delays Of Arrival are calculated, selected and filtered, and used to estimate the positions of whales by a constant or a linear sound speed profile. SMF shows higher performance than TKM with more positions estimated. Our results are validated by partial similar results from the US Navy and Hawaii univ labs in the case of one whale, and by similar whales counting from the Columbia univ. ROSA lab, and other partial preliminary results from SOEST Hawaii in the case of multiple whales. Our results on multiple whales are likely (speed, depth, tracks) and are on our knowledge the first complete one on such difficult multiple whales record. Moreover our model is realtime. We evaluate the *a priori* performance of the system via the Cramér-Rao Lower Bound (CRLB) and Monte Carlo simulations. We show that the model is validated by good performances with these theoric CRLBs and the computed confidence ellipses. Localization potentiality is discussed and widen to study the whale's behavior thanks to an automatic labeling of the signal and to extract different features such as the speed, energy of the clicks, Inter-Click-Interval (ICI), etc. These features allow us to cross-analyse the whales behavior (foraging, hunting, ingestion) and to see the influence

of each parameter. The complete localization method is processed on real data from the Naval Undersea Warfare Center and the Atlantic Undersea Test And Evaluation Center.

PACS numbers: 43.80.n, 43.80.Ka, 43.80.Ev

Keywords: Bioacoustics, multiple clicking whales, passive acoustic localization

Nomenclatura

AUTEC: Atlantic Undersea Test And Evaluation Center.

CRLB: Cramér-Rao Lower Bound.

CC: Cross-Correlation.

D1, D2: Dataset 1, Dataset 2.

H1, H7: Hydrophone 1, Hydrophone 7.

ICI: Inter-Click-Interval.

MM: Marine Mammal.

NUWC: Naval Undersea Warfare Center.

SMF: Stochastic Matched Filter.

SNR: Signal to Noise Ratio.

SOEST: School of Ocean and Earth Science and Technology.

TK: Teager-Kaiser.

TDOA: Time Delay Of Arrival.

TKM: Teager-Kaiser-Mallat.

s(n): Sample n of the raw signal.

x(n): Signal of interest (click).

u(n): Wide Sense Stationary (WSS) Gaussian noise.

w(n): Random gaussian process.

$T(i, j)$: TDOA between hydrophones i and j.

$T(\tilde{i}, j)$: Estimate of T(i,j).

I. INTRODUCTION

In this paper, we compare two low cost time-domain tracking algorithms based on passive acoustics. The problem consists in tracking an unknown number of sperm whales (Physeter catodon). Clicks are recorded on two datasets of 20 and 25 minutes on an open-ocean widely-spaced bottom-mounted hydrophone array. The output of the method is the track(s) of the MM(s) in 3D space and time. Firstly, we briefly review studies of the SMF detector and its performances with an echo cancellation, the TKM filtering, the source separation methods and the main characteristics of MM signals. Then, we propose a real-time algorithm for MM transient call localization. We also briefly recall the CRLB[1] and the confidence ellipses theory to predict the reachable accuracy and compare it to the tracking results. In Section 3 we show and compare results of track estimates with results from specialized teams and compare SMF versus TKM localization. Then, the system is evaluated with the confidence ellipses on the trajectories. Finally, we discuss on the possible dynamic behavior analysis of the whale that these localizations offer, like hunting and foraging strategies.

This paper deals with the 3D tracking of MM using a widely-spaced bottom-mounted hydrophone array in deep water. It focuses on sperm whale clicks. There were previous algorithms developed in the state of art[2–5] but none of them has satisfying results for multiple tracks and most of them are far from being real-time. Our main goal is to build a robust and real-time tracking model, despite ocean noise, multiple echoes, imprecise sound speed profiles, an unknown number of MM, and the non-linear time-frequency structure of most MM signals. Background ocean noise results from the addition of several noises: sea state, biological noises, ship noise and molecular turbulence. Propagation characteristics from an acoustic source to an array of hydrophones include multipath effects (and reverberations, Fig. 1), which create secondary peaks in the Cross-Correlation (CC) function that the generalized CC methods cannot eliminate. In[6,7], we gave an extension of[2] that shows multiple tracking using TKM. Here we improve this model using SMF which also allows an efficient echo removal process developed higly defined tracking and inter-click-interval analysis.

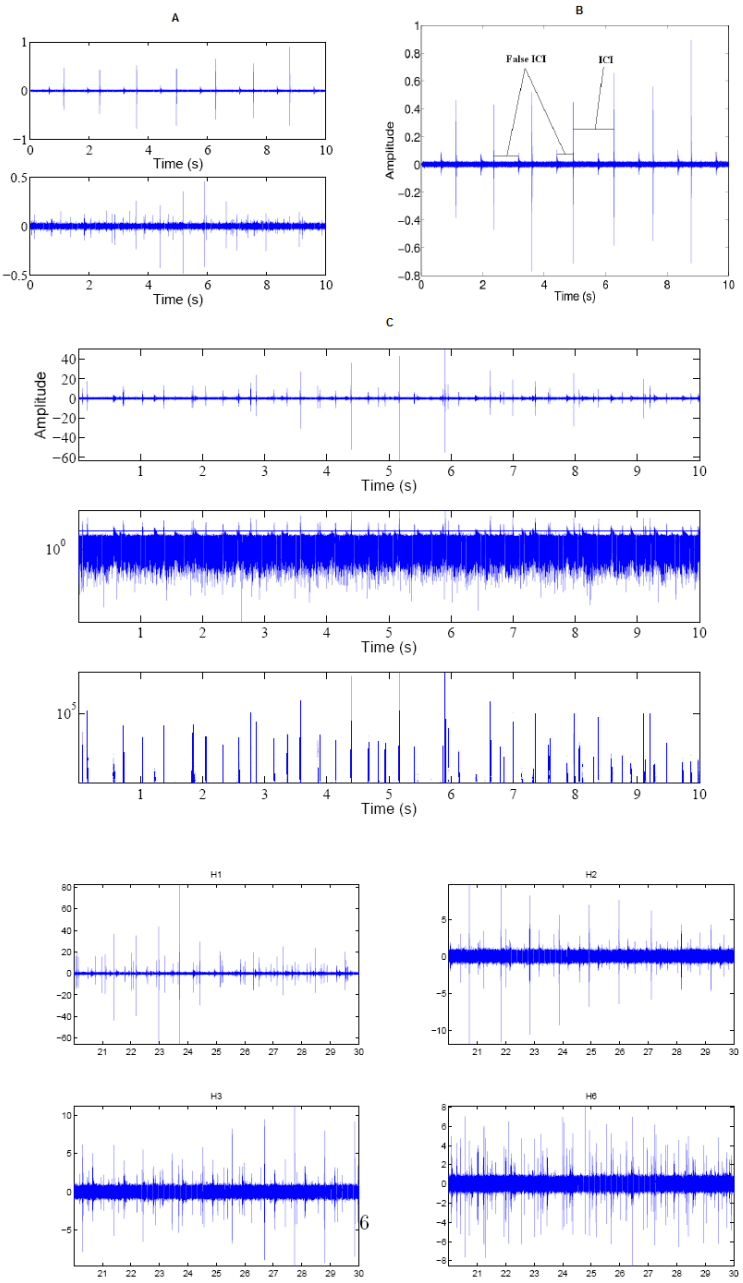

FIG. 1. (A): on the top, a raw signal from dataset2 (D2) and hydrophone 7 (H7) during the first 10s of recording, containing 7 clicks and their echoes. At the bottom, 10s (from D1, H1) containing several (4) simultaneous emitting whales click and the echoes. (B): a click train with echoes from a single sperm whale. We can see an ICI, and two false ICI. (C): Example of a raw multipe whales' signal on H1 (10s) (top) and the corresponding $\Lambda(x)$ with the threshold in a log-scale (middle) and the thresholded signal (bottom). H1 to H6: plots of a 10 sec samples of raw signals from the four hydrophones of the dataset D1 (time in sec).

II. HYDROPHONE ARRAY CARACTERISTICS

A. Records Settings

The signals are records of March 2002 from the ocean floor (about 1500m) near Andros Island - Bahamas (Tab.I), provided with celerity profiles. Datasets are sampled at 48 kHz and contain MM clicks and whistles, background noises like distant engine boat noises. Dataset1 (D1) is recorded on hydrophones 1 to 6 during 20 min length (see Fig.1 for a sample view) while the dataset 2 (D2) is recorded on hydrophones 7 to 11 with 25 min length. We will use a constant sound speed with $c = 1500ms^{-1}$ or a linear profile with $c(z) = c_0 + gz$, where z is the depth, $c_0 = 1542ms^{-1}$ is the sound speed at the surface and $g = 0.051s^{-1}$ is the gradient[6]. Sound source tracking is performed by continuous localization in 3D using Time Delays Of Arrival (T) estimation from four hydrophones (Tab.I).

B. Cramér-Rao Lower Bound from the hydrophone array geometry

For each hydrophones array, the Cramér-Rao Lower Bound (CRLB) provides the maximum accuracy for the estimation of any source position. Considering a constant sound speed profile, the function model of the TDOA is defined by:

$$s(\theta) = \frac{1}{c_s}\big[||X_i - \theta|| - ||X_j - \theta||,\ ||X_i - \theta|| - ||X_k - \theta||,\ ||X_i - \theta|| - ||X_l - \theta||\big]^T, \quad (1)$$

where $||\ ||$ denotes the euclidian norm, X_i is the vector coordinate of hydrophone i, θ is the unknown parameters vector $[x\ y\ z]^T$ and c_s the celerity. Here $i = 1$, $j = 2$, $k = 3$, $l = 4$. Thus, considering the TDOA noise as a Gaussian process and B its variance-covariance matrix, the Fisher Information matrix is:

$$I_\theta = \nabla_\theta s(\theta) B^{-1} \nabla_\theta^T s(\theta). \quad (2)$$

Then, the CRLB is $B_\theta = I_\theta^{-1}$. The solution error ellipses are contours of constant value of the inner product $\theta I \theta$.

We compute the CRLB (in meter) in the space (x,y,z) and plot the values for both datasets (Fig.2). We consider that the standard deviation of the noise is equal to the quantification noise with a sampling frequency of 480Hz. The main dependencies of the bounds are the noise and the array configuration. In figure 2.A to F, the CRLB on y and z is shown for a depth of 500m, and is just about the same for a depth of 1000m as shown in figure 2.G-H.

III. FILTERS DESIGN

A. Teager-Kaiser-Mallat Filtering

A sperm whale click is a transient increase of signal energy lasting about 20 ms (Fig.1). Therefore, we use the Teager-Kaiser (TK) energy operator[8] on the discrete data:

$$\Psi[x(n)] = x^2(n) - x(n+1)x(n-1), \quad (3)$$

where n denotes the sample number. Considering the raw signal $s(n)$ as:

$$s(n) = x(n) + u(n), \quad (4)$$

where $x(n)$ is the signal of interest (clicks), $u(n)$ is an additive noise defined as a process realization considered Wide Sense Stationary (WSS) Gaussian during a short time. By applying TK to $s(n)$, $\Psi[s(n)]$ is:

$$\Psi[s(n)] \approx \Psi[x(n)] + w(n), \quad (5)$$

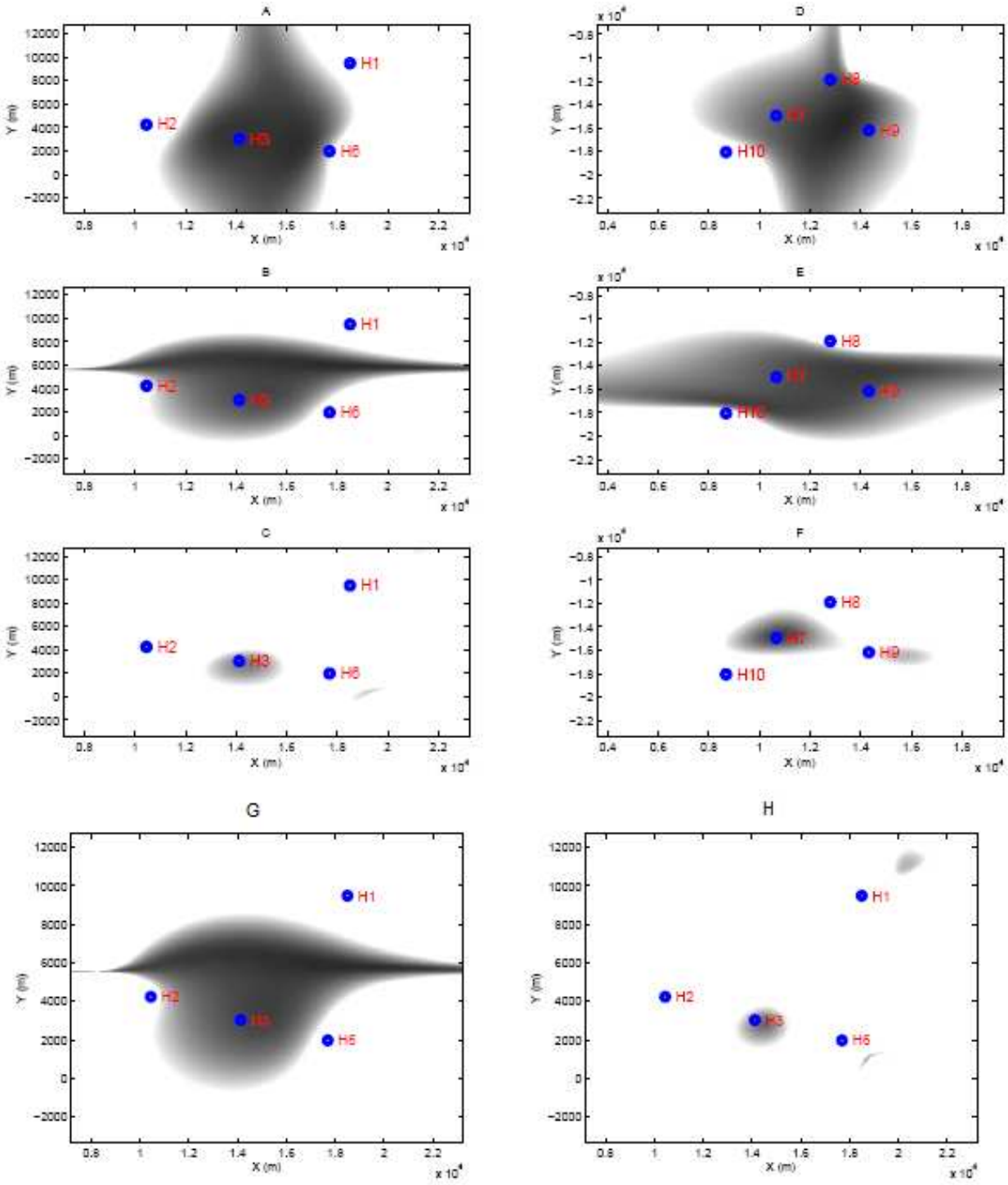

FIG. 2. CRLB values scaled in gray colors with a plan view: black means a null CRLB and white a CRLB≥10 m. For the figures A to F, a depth of 500m was chosen. (A): CRLB on x values, dataset D1. (B): y, D1. (C): z, D1. (D): x, D2. (E): y, D2. (F): z, D2. (G): CRLB on y axis, plan view, dataset D1, depth=1000m. (H): CRLB on z axis, plan view, dataset D1, depth=1000m.

where $w(n)$ is a random gaussian process[8]. The output is dominated by the clicks energy. Then, we reduce the sampling frequency to 480Hz by the mean of 100 adjacent bins to reduce the variance of the noise. We apply the Mallat's algorithm[9] with the Daubechies wavelet[10] (order 3) into the process. The signal is denoised by a universal thresholding[11]. This filtering step is very fast and without any parameter. The Fig.5-b shows the filtered signal of multiple emitting MMs signal (Fig.5-a). The click detection[8] is not a concern. The purpose is to ameliorate the TDOA estimation, but echoes are not removed. The TK operator can be used for detection but need an empirical threshold, thus we need a more adaptive method. The purpose of this paper is not a full comparison between the performances of two detectors, but to offer two methods for the whale tracking and the performance relative to these methods.

B. The Stochastic Matched Filter

The SMF, which is a filtering method, is employed here for detection. The clicks and sea noise are considered as gaussian stochastic process with 0-mean. Considering a stochastic process s (of length N), the covariance matrix is $E(ss^T) = A$. We also consider an additive, centered and independent noise b with the variance-covariance matrix B. Those processes are not correlated to each other and the matrix are supposed positive and full rank. The SMF theory[12,13] says that the linear filtering h of length N that maximizes the SNR (Signal to Noise Ratio) is the eigen vector solution of:

$$Bh = \lambda Ah, \tag{6}$$

associated to the greatest eigen value λ_0. Thus, we are looking for the eigen values and vectors of $B^{-1}A$. The function used for the detection is:

$$\Lambda(x) = [h^T x]^2, \tag{7}$$

with x a N-length window. Denoting ρ as the SNR gain after filtering, and $\tilde{M}$ a matrix normalized by its trace, we have $\rho = \frac{h^T \tilde{A} h}{h^T \tilde{B} h}$. We work on windows of 20ms which correspond

to a mean click length. A is computed with an average of 1000 sperm whale clicks, and B is calculated directly from the hydrophone signals. After h is calculated for each channel, we are able to filter the signal with one bin of shifting. Thus, we obtain $\Lambda(x)$ (Fig.1.C) with a threshold chosen considering the performance that maximizes the synthetic ROC.

We compute the Receiver Operating Characteristic (ROC) curve for a SNR of -10dB which corresponds to an emitting whale at about 5km from the receptor. SNR (dB) is computed by $10log_{10}(\frac{P_s}{P_b})$ with $P_s = \frac{s^T s}{N}$ and $P_b = E(\frac{b^T b}{N})$. The detection rate for 1% false alarm rate of the SMF is 49%. To the contrary of the TKM filter, as we use SMF as a detector, we have a date for each probable click, and thus we can eliminate false detections like echoes.

C. Echo cancellation

In order to generate robust estimates of the TDOA we avoid relying completely on correlation based techniques. In reality, the SMF detector often detects the echoes present after each click (Fig.1.B). To remove it, we work on each detection date given by SMF in a channel, considering it as an echo or a click, and we seek to discriminate between direct path and reflected arrivals. Since the multi-path arrivals pass through the surface layer and are reflected from the sea air interface, they are subject to significant surface reverberation. This causes a temporal elongation (Fig.3.C,[14]). Dan Ellis and al.[15] proposed a cancellation method based on frequential propreties. Here we propose for a simpler temporal discrimination. Therefore, on the raw signal, we smooth each potential click and calculate the sum of the normalized envelope (Fig.3.D). Consequently, the results from direct and reflected paths are significantly different.

A relatively crude threshold allows one to distinguish the majority of reflected signals from the direct arrivals. There is no demand for highly accurate discrimination; subsequent[2] delay estimation algorithm performs well as long as the majority of events surviving discrimination correspond to direct arrivals.

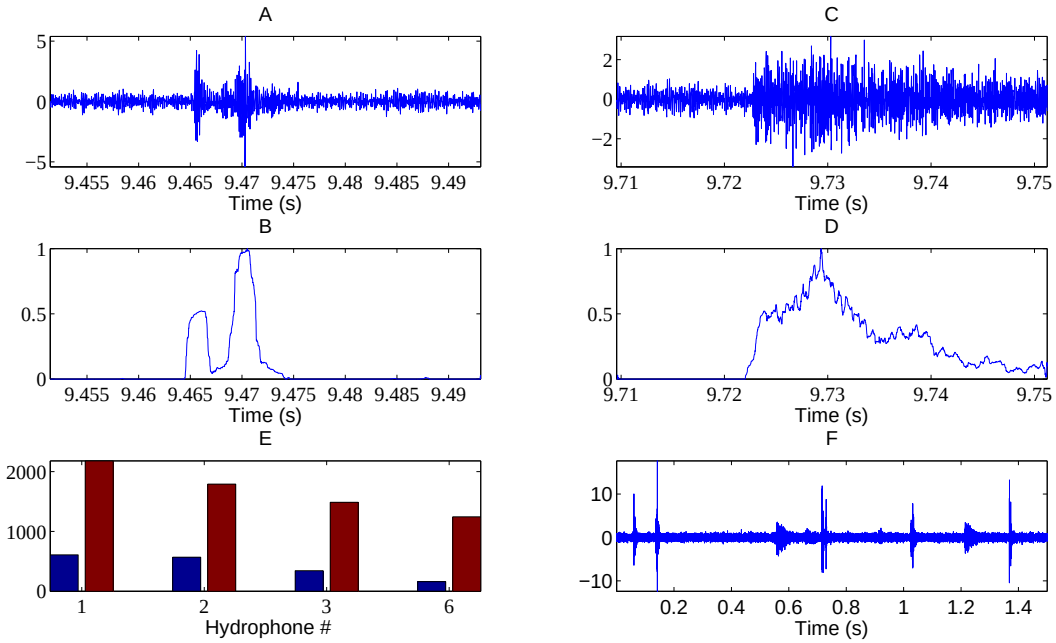

FIG. 3. A direct click (A) with its absolute normalized smoothed envelope (B) compared to an echo (C) and its envelope (D). (E): number of echoes (bar on the left) and clicks (bar on the right) detected for each hydrophone in D1 (H4 and H5 are out of order). (F): Raw signal (H1) showing clicks and echoes with several emitting whales.

The Fig.4 summarizes the echoes and clicks detected in both datasets. The total number of clicks in D2 are manually detected. We denote that the number of clicks detected with the SMF is varying with the hydrophone because of the different SNR on each hydrophone, and that the echoes are partly rejected (considering one echo per click).

IV. ROBUST AND FAST LOCALIZATION METHOD

A. Rough $\widetilde{T}$ Estimation and Selection

Either after TKM or after [SMF + Echo cancellation], we process rough TDOA estimation and selection. The Fig.5-a,c are respectively a raw multiple emitting MMs signal, and the corresponding complete filtering signal. First, TDOA estimates are based on MM

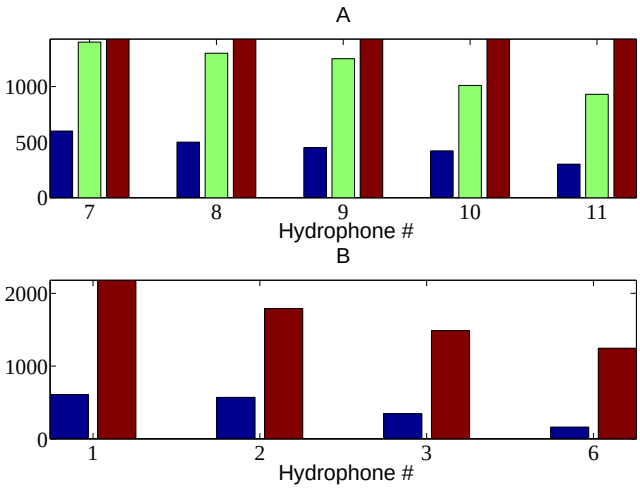

FIG. 4. SMF + echo cancellation statistics. In figure A, we can see, on each hydrophone, the echoes detected in dataset D2 (bar on the left), and the clicks detected (bar in the middle). The real number of clicks in D2 is shown on the bar on the right. In figure B (D1), bar on the left: echoes detected, bar on the right: clicks detected. We do not know yet the exact number of clicks in D1.

click realignment only. Every 10s, and for each pair of hydrophones (i, j), the difference between times t_i and t_j of the arrival of a click train on hydrophones i and j is referred as $T(i, j) = t_j - t_i$. Its estimate $\widetilde{T}(i, j)$ is calculated[2,3] by CC of 10s chunks (2s shifting) of the filtered signal for hydrophones i and j. We keep the 15 highest peaks on each CC to determine the corresponding $\widetilde{T}(i, j)$. The filtered signals give a very fast rough estimate of TDOA (precision $\pm$ 2ms). The Fig.5.d shows the CC with the raw signal, and the Fig.5.e,f with respectively the TKM filtered signal and the SMF. The number of common peaks between TKM, SMF and raw filtering methods are in Tab.II. At most, only 23% of the TDOA are common between the 2 filters, the 23% corresponding to the TDOA of the high SNR clicks. Without any filtering, CC generates spurious delays estimates and the tracks are not correct. Finally, thanks to the $\widetilde{T}$ transitivity constrained system described in[2], we keep $\widetilde{T}$ triplets coming from the same source.

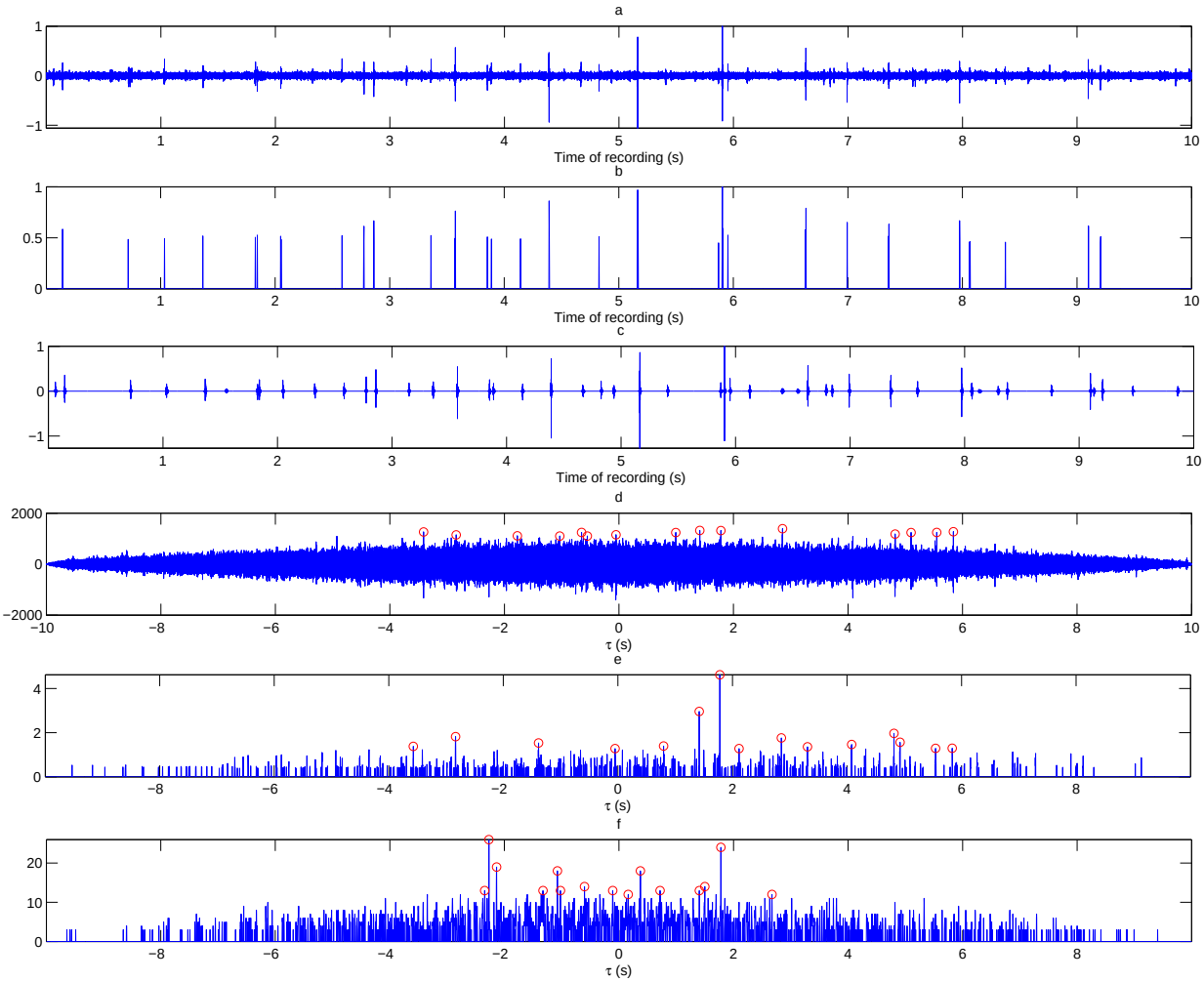

FIG. 5. (a): raw signal for dataset D1 of H1 during the first 10s of recording showing multiple emissions. (b): (a) after TKM filtering. (c): (a) after the SMF. (d): CC between (a) and corresponding raw signal chunk of H3. (e): idem than (d) but with (b). (f): idem than (d) but with (c). Circles on the top of some peaks correspond to the 15 maximum peaks (that are used for localization).

B. Localization with a constant and a linear profile

Thanks to the measured delays and an acoustic model based on a constant or a linear sound speed profile, the least squares cost function determines the MM positions using a multiple non-linear regression with the Gauss-Newton method[2] (Levenberg-Marquardt technique[16]). The residuals are approximatively following a Chi-square distribution with

$Nc - d$ degrees of freedom, which is noted X^2_{Nc-d}, where Nc is the number of hydrophones couples considered and d the number of unknowns, that are the coordinates (x, y, z). The position is accepted if the residual is inferior to a threshold x, that is calculated solving $P = prob(X^2_{Nc-d} > x)$ with $P = 0.01$ (we keep 99% of the estimates).

V. RESULTS

A. Tracking Comparisons

In this section, we give the tracks results for TKM and SMF for D1 and D2 dataset. For the dataset D2 (Fig.6), a constant and a linear sound speed profile were used like in[6] and the results are similar to those of Morrissey's[5] and Nosal's[4] methods. The diving profile underlines a bias of about -70m between the linear and the constant profiles results, emphasizing the importance of the chosen profiles. Moreover with the linear sound speed, the results are about the same as those of Morrissey and Nosal, who used profiles corresponding to the period and place of the recordings. The results on D1 are shown in Fig.7,8,9, and are for a linear sound speed profile. The TKM method lets appear an artifact whale (yielding to 5 whales), which is due to the echoes produced by the whale with the (+) symbol which are eliminated thanks to the echo removal with the SMF method (without the echo removal, the same virtual whale appears in the SMF results) but we can not apply an echo removal on TKM because it is not used as a detector. We thus localize 4 MM with the SMF method. The number of positions estimated for each method is in Tab.II. The confidence regions are computed for both datasets with a Monte Carlo method. The ellipses maxima (30m) fit with MM length (20m), which is acceptable.

In the SMF method, there are much more estimated positions, this can be explained that the SMF used in detection, detects partically all the clicks and even the ones with a bad SNR. This last depends on the direction and the distance of the whale to the hydrophone. The signals correlation for the TDOA (Time Delay Of Arrival) estimation are thus binary,

and do not contain the information of distance, whereas the Teager-Kaiser method just enhances the correlation, where the signal is filtered without detection and thus low click energy results in small correlation, compared with the high click correlation. This is why in the 3D region, when the whale is in the opposite direction and/or emits small energy clicks, the SMF method returns more positions. The result for the left whale (note with '+') in the multiple whales case is a good illustration. In the first trajectory part, the whale is off-axis with all the hydrophones, so the SNR is low on pratically all the hydrophones, and thus, TDOA extraction and localization are difficult with the TKM method (SMF is much better). In the last part, the whale is on-axis with 2-3 hydrophones, then the TDOA and positions estimation are facilitated thanks to the high SNR and then the TKM method generates more positions per time sample relatively to the first trajectory part.

To summarize the confrontation, SMF produces more positions than TKM because it is a more efficient detector. But it does not modify the accuracy of the position estimation, as seen below in section V.B.

B. Confidence Region analyses

To compute the ellipses, we apply a Monte Carlo method and a gaussian distribution noise with the standard deviation described above. For each $\widetilde{T}$ realization, the source position is calculated. We deduce the variance and the mean for each position to plot the confidence regions with a confidence level of 0.95, which means that there is 0.95 probability for the whale to be in the ellipse centered on the position. In dataset D2, the mean values of the confidence intervals on X, Y, Z axes are about 18, 16 and 30m (Fig.10). The results confirm that the errors on the vertical axis are meaningfully higher than the other axes because the distance between each hydrophones in this direction is smaller (the maximum difference on the Z axis between hydrophones is 200m). As estimated by the CRLB analysis in section II.B, the farthest whales in dataset D1 from the hydrophones array center have a larger uncertainty with an error of about 20 to 30m on X and Y axes (Fig.10), while the

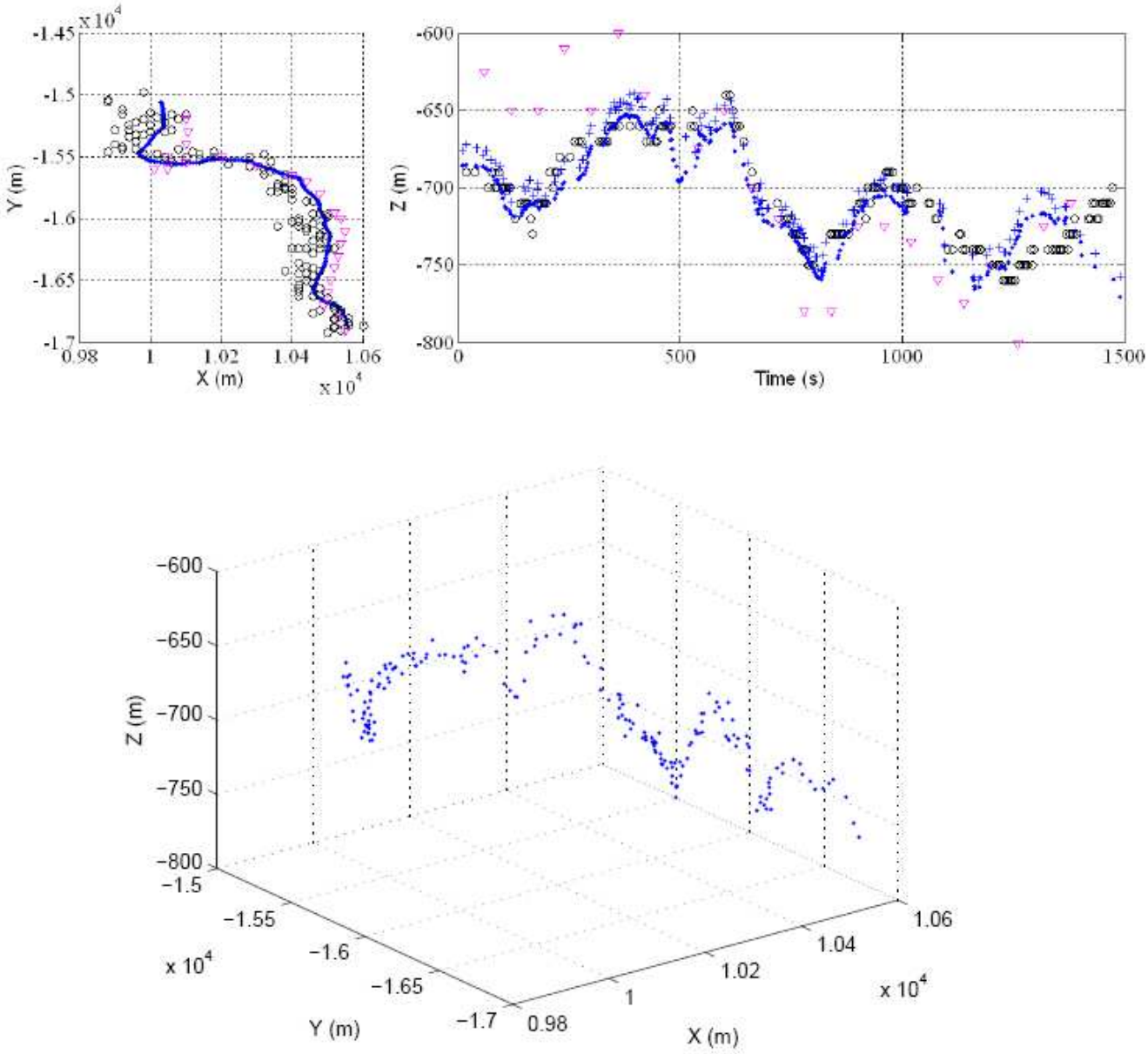

FIG. 6. Top: plan view and diving profile of the MM in D2, our estimates with a linear profile and the SMF method (.) or TKM (+); and estimates of Morrissey's ($\triangledown$) and Nosal's methods (o). The results with a constant profile underline a bias of about -70m[6]. Bottom: 3D plot of the trajectory in D2.

whales close to the center (Fig.10) exhibit an error of about 10 to 20m like for D2 (Fig.7). These uncertainties are reasonable according to the sperm whale length (20 to 30 meter).

Comparing the CRLB with the ellipses, we denote the correlation between maximum

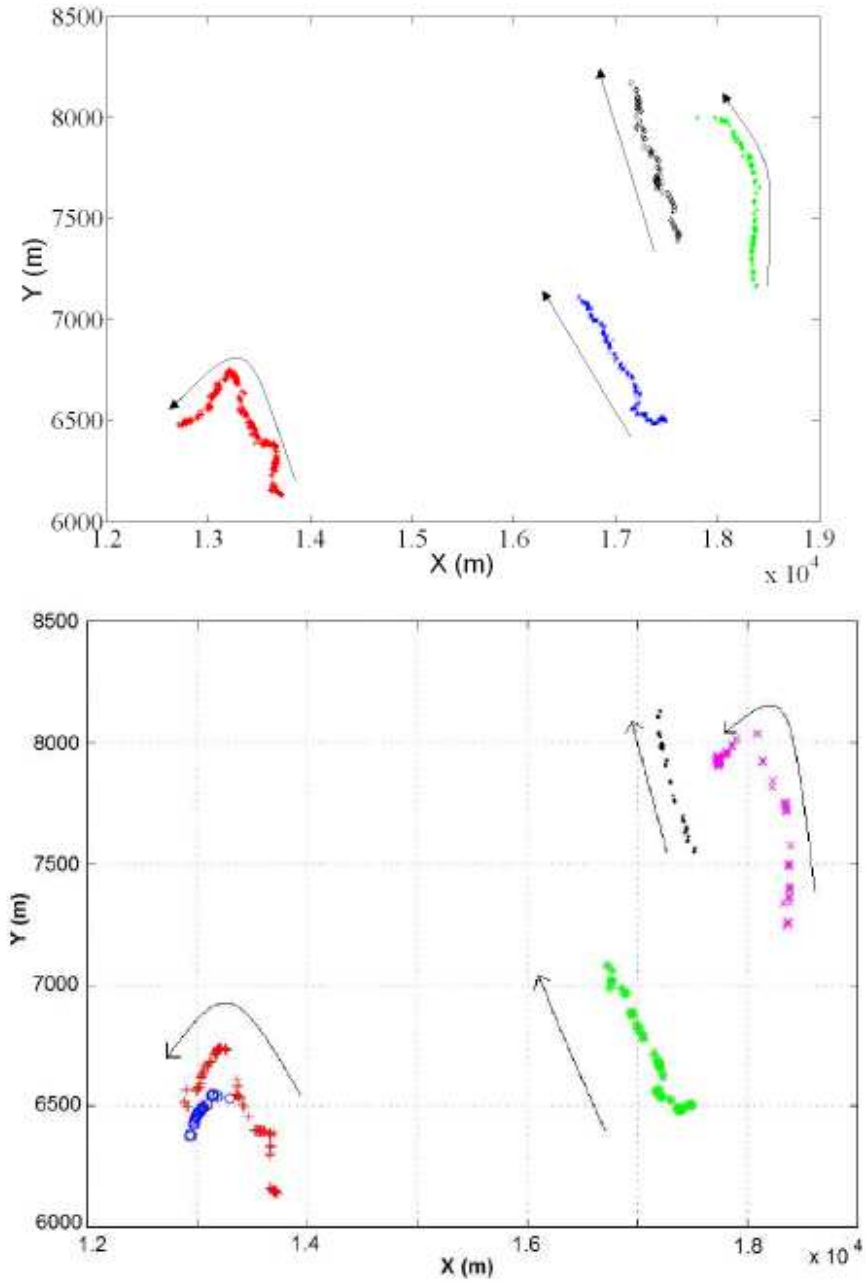

FIG. 7. On the top, plan view in D1 with the SMF method. Each symbol corresponds to one of the 4 whales. The arrows stress the directions. On the bottom, plan view in D1 with the TKM method. Each symbol corresponds to one of the 5 whales. Compared to the SMF, one false whale appears due to the lack of echo removal for TKM. Another great difference is the high definition of the SMF track compared to TKM ones, as discribed in Tab.II. The arrows stress the directions. A 3D plot of the trajectores with the SMF method can be seen on Fig.8.

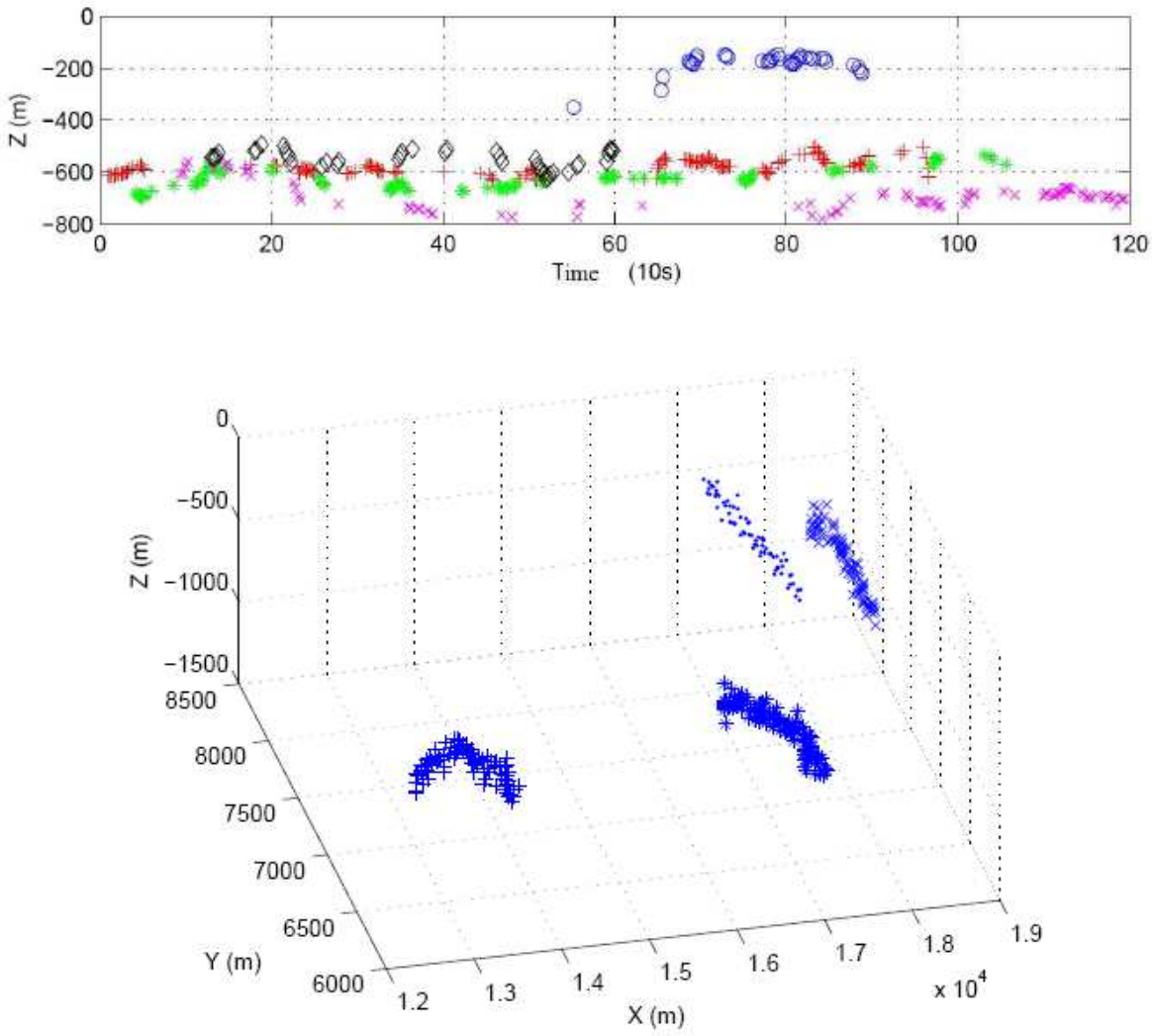

FIG. 8. Top: averaged diving profile in D1 with TKM method. Each symbol corresponds to one whale in Fig.7 (Bottom to top (+), (o), (hexagon), (x), (diamond)). The SMF results in section V.B demonstrate that the 'o' is an echo of the '+' whale (cf Fig.7). Bottom: 3D plot of the trajectories in D1.

accuracy and confidence regions. Figures 2-C and 2-F show the accuracy on z axis is larger than 10 m for both datasets in the tracking regions, which is consistent with the ellipses results, but in the case of D1, the diving profile estimation is not suitable mainly due to the z-component of the CRLB (Fig.2.H). The CRLB (in D1, Fig.2.D-E) also explains that the farthest whale has larger confidence regions. But for both datasets, the CRLB on x and y

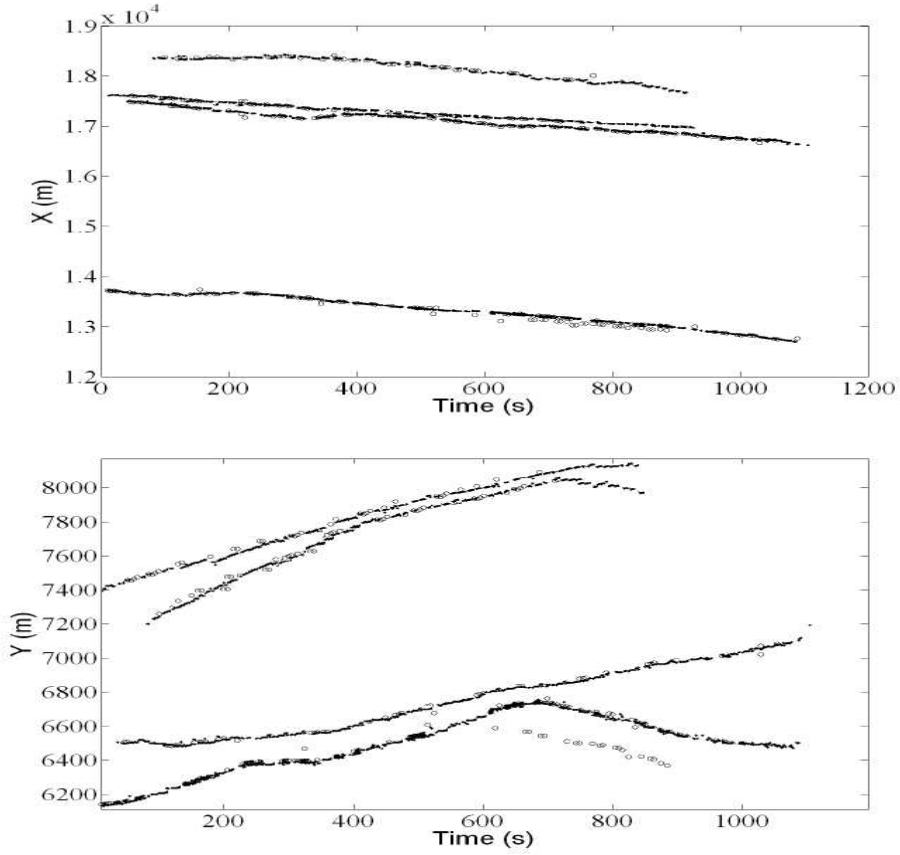

FIG. 9. Comparisons between SMF tracks (.) and TKM tracks (o) in D1, demonstrating that SFM outperforms TKM. X(t) (Y(t)) coordinates are on the top (bottom). We can see that the number of positions is far larger with the SMF than TKM (see Tab.II).

is far inferior to 10m inside the array whereas our ellipses are about 10 to 20m, which is maybe caused by other parameters like the approximated celerity profile are involved in.

VI. PERSPECTIVES ON WHALE BEHAVIOR ANALYSES

Our method gives in real-time multi tracks of a whales group. The Fig.11 shows the speed profile for each whale. These localizations allow to use the true TDOA and to label the signal. This leads to a precise ICI extraction[17]. Other features can also be extracted thanks to this localization, such as the speed, energy of each click, distance to a given hydrophone i

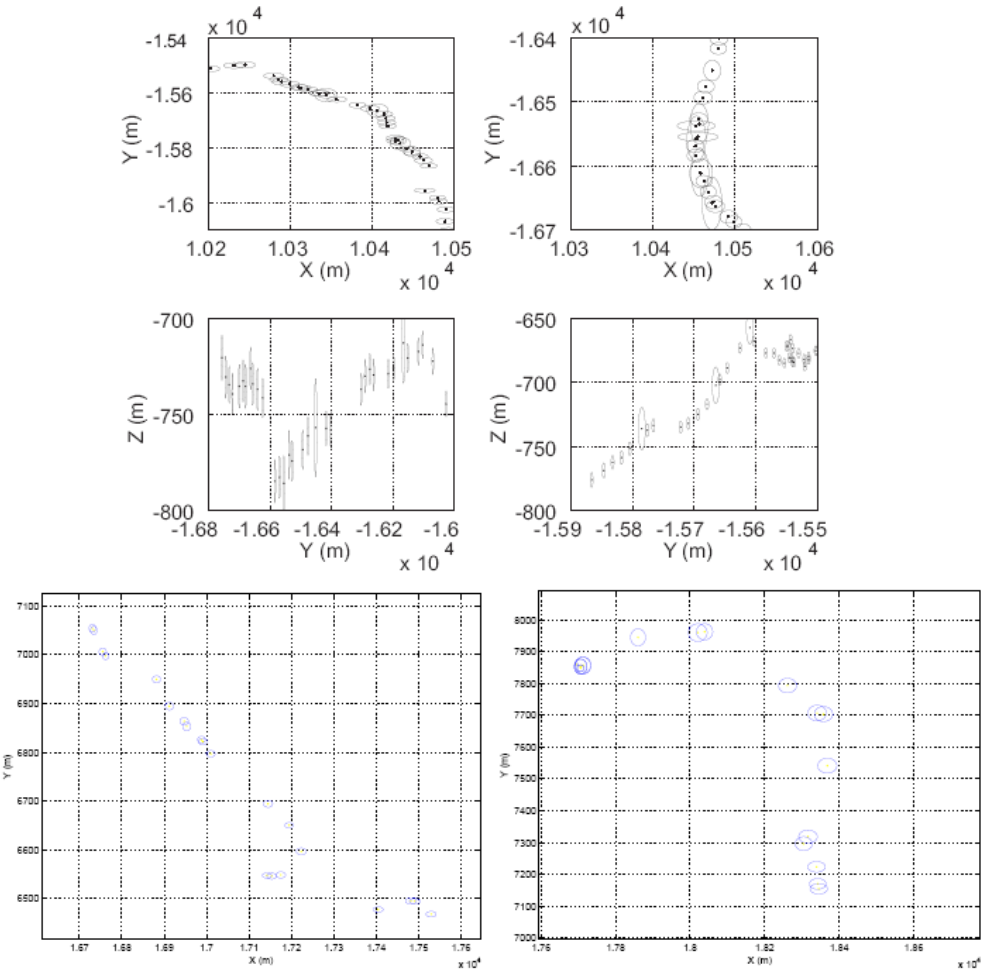

FIG. 10. Top: confidence region projection on X and Y and on Z and Y axes for different trajectories in D2 with the SMF. Bottom left: confidence region projection on X and Y for the whale in the middle (* symbol), D1 dataset, trajectory with the SMF. Bottom right: confidence region projection on X and Y for the whale in the right (x symbol), D1 dataset, trajectory with the SMF.

and head's angles with a given hydrophone. These features would give relevant informations on the whale's behavior when hunting prey at depth. It is admitted that sperm whales make a slow pitch movement and created a faster pitch or yaw movement in synchronization with the clicking activity. The literature[18],[19] suggested that sperm whales would, at depth, make an asymmetric scan of the surrounding water and that during the search phase sperm whales would methodically scan a cone-shaped mass of water when searching for prey. This

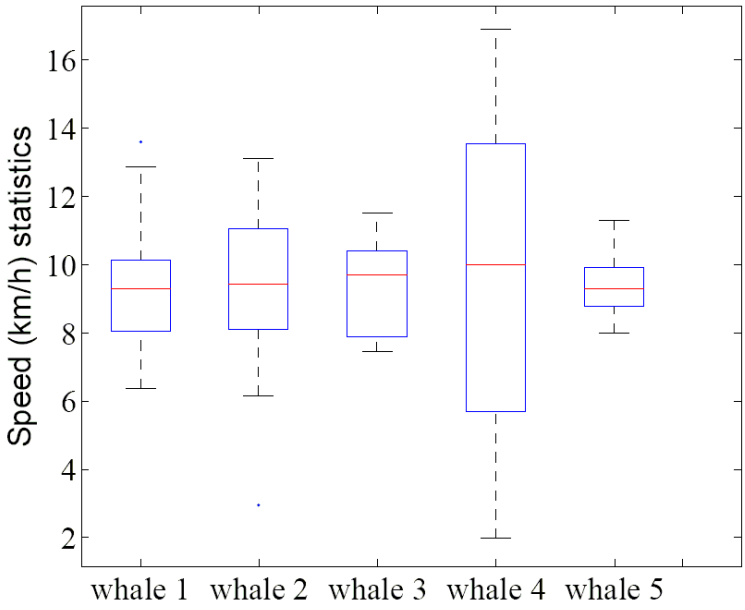

FIG. 11. Speed (averaged on 30s windows) statistics on the whole set for each whale in dataset D1 (whales 1 to 4, numeroted from left to right) and dataset D2 (whale 5). The central line of the box is the median of speed and the lower and higher lines are the quartiles. The whiskers show the extent of the speed (9km/h likely as regular sperm whale).

scan would suggest that each sperm whale click is generated to aim in a specific direction and at a specific range. Sperm whales would move physically to change the click beam direction, and control level and ICI to change the click target range. Authors in[18] then pointed out a correlation between click level variations and ICI. The hypothesis explaining such a correlation would be click level control: sperm whales would click slowly at a high source level and faster at a lower source level. We are currently analysing the dependencies of all these results that will validate or note this model.

The Fig.12 summarizes different possible features computation in our framework, and their connections with our tracking algorithm. Actually, these measurements are correlated with each other and offer a new large research field for whale behavior analysis.

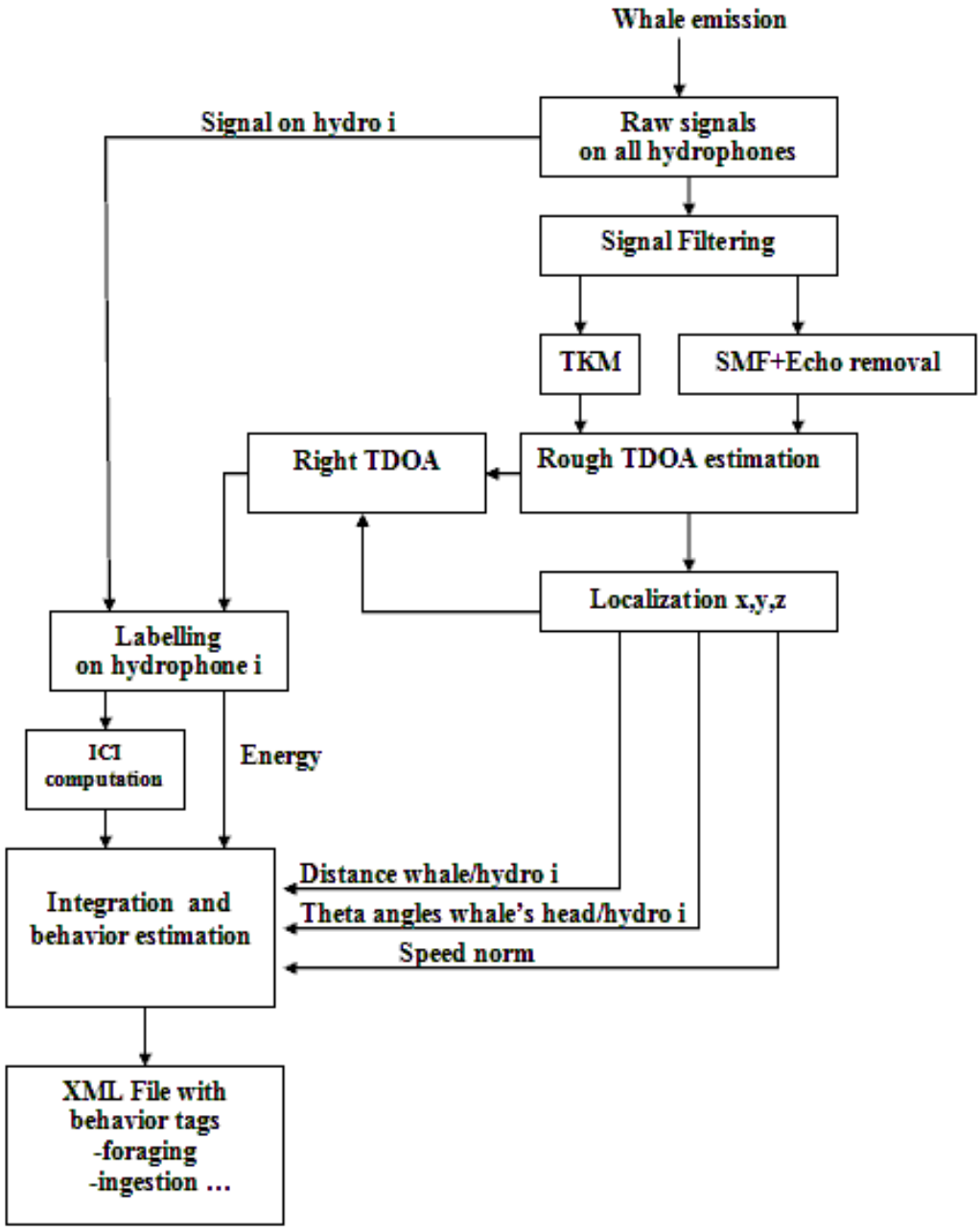

FIG. 12. Block scheme of the possible features extraction.

VII. CONCLUSION

We present in this paper a method using a real-time algorithm for tracking one or multiple emitting sperm whales, in order to analyse whale behavior and to index hydrophones audio files. All the derived features (speed, direction changes) are likely according to the

cetologists knowledge. Part of the results presented in this paper are in video demonstrations on http://glotin.univ-tln.fr/PIMC/DEMO.

The localization step is run in real-time on a standard laptop, and works for one or multiple emitting sperm whales. The depth results with a constant speed contain a bias error due to the refraction of the sound paths from the MM to the receivers which a linear speed corrects. Another way to tackle the speed profile issue[7] would be to estimate it as a fourth unknown in the regression. However, SMF perfoms better than TKM method (see Tab.II, 103 positions estimated for TKM versus 387 for SMF in D1). The SMF provides a simple way for detecting the sperm whales with good performance, where all the thresholds are learned online and no parameters are needed considering the database. In D2, results indicate that only one sperm whale was emitting in the area, as also analysed by Morrissey[5] and Nosal[4], but not in real-time. Moreover, according to ROSA Lab estimation[15] based on click clustering, and also preliminary results from the SOEST, the number of MM for each 5min chunks on D1 (Tab.III) is similar to ours ($[4; 4; 4; 2]$). The localization accuracy is computed via the CRLB and the confidence ellipses which are correct considering the MM length. The depth error is mainly due to the low precision reachable with the CRLB considering the array configuration.

Our method allows to label the signal and then will be used to extract features presented in Fig.12. Other characteristics such as inter-pulse-interval (a click is composed of multiple pulses), which contain informations about the whale's pitch and yaw behavior. The features discussed above would allow us to analyse whale's foraging and hunting behavior with a good resolution. Finally, we can index the files thanks to features extracted. Therefore, a XML structure can be generated and include some behavior tags for a rapid access to the acoustic data. Our method offers facilities for robust online passive acoustics behavior studying of clicking MM in open ocean. More research will be conducted to reveal high level features (ICI, behavior, hunting) from the tracks.

Acknowledgments

We thank the AUTEC and the NUWC for having provided the dataset. This research was conducted within the international sea *pôle de compétitivité* at Toulon-France, through "Plateform for Integration of Multimodal Cetacean data (PIMC)". Part of this work is funded by the "Conseil régional Provence-Alpes-Côte d'Azur" France, and Chrisar Software Inc. Another part of this work has been previously pattented.[20]. We thank as well O.Adam, R.Morrissey and E.Nosal for the use of their work, and P.Giraudet for the discussions.

References

[1] S. Kay, *Fundamentals of statistical signal processing* (Prentice Hall, PTR) (1993).

[2] P. Giraudet and H. Glotin, "Real-time 3d tracking of whales by echo-robust precise tdoa estimates with a widely-spaced hydrophone array", Applied Acoustics **67**, 1106–1117 (2006).

[3] P. Giraudet and H. Glotin, "Echo-robust and real-time 3d tracking of marine mammals using their transient calls recorded by hydrophones array", in *ICASSP IEEE* (2006).

[4] E. Nosal and L. Frazer, "Delays between direct and reflected arrivals used to track a single sperm whale", Applied Acoustics **62**, 1187–1201 (2006).

[5] R. Morrissey, J. Ward, N. DiMarzio, S. Jarvisa, , and D. Moretti, "Passive acoustics detection and localization of sperm whales in the tongue of the ocean", Applied Acoustics **62**, 1091–1105 (2006).

[6] F. Caudal and H. Glotin, "Multiple real-time 3d tracking of simultaneous clicking whales using hydrophone array and linear sound speed profile", ICASSP IEEE 4p (2008).

[7] H. Glotin, F. Caudal, and P. Giraudet, "Whales cocktail party: a real-time tracking of multiple whales", International Journal Canadian Acoustics **36**, 141–147 (2008).

[8] V. Kandia and Y. Stylianou, "Detection of sperm whale clicks based on the teager-kaiser energy operator", Applied Acoustics **67**, 1144–1163 (2006).

[9] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet represen-

tation", IEEE Transaction on Pattern Analysis and Machine Intelligence **11**, 674–693 (1989).

[10] O. Adam, M. Lopatka, C. Laplanche, and J.-F. Motsch, "Sperm whale signal analysis: Comparison using the autoregressive model and the wavelets transform", International Journal of Information Technology **2**, 1–8 (2005).

[11] D. L. Donoho, "De-noising by soft thresholding", IEEE Trans. IT **41**, 613–627 (1995).

[12] P. Courmontagne and F. Chaillan, "The adaptive stochastic matched filter for sas images denoising", OCEAN 2006 1–6 (2006).

[13] N. Juennard, "Acoustic submarine detection and localisation of very high energy particules", Thèse de doctorat, University of Toulon (2007).

[14] P. White, T. Leighton, D. Finfer, C. Powles, and O. Baumann, "Localisation of sperm whales using bottom-mounted sensors", Applied Acoustics **62**, 1074–1090 (2006).

[15] X. Halkias and D. Ellis, "Estimating the number of marine mammals using recordings from one microphone", in *ICASSP IEEE* (2006).

[16] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters", SIAM Journal on Applied Mathematics **11**, 431–441 (1963).

[17] F. Caudal and H. Glotin, "Automatic inter-click-interval (ici) and behavior estimation for one emitting sperm whale.", in *PASSIVE 08 IEEE* (2008).

[18] C. Laplanche, O. Adam, M. Lopatka, and J.-F. Motsch, "Measuring the off-axis angle and the rotational movements of phonating sperm whales using a single hydrophone", Journal of Acoustical Society of America. **119**, 4074–4082 (2005).

[19] C. Laplanche, O. Adam, M. Lopatka, and J.-F. Motsch, "Male sperm whale acoustic behavior observed from multipaths at a single hydrophone", Journal of Acoustical Society of America. **118**, 2677–2687 (2006).

[20] H. Glotin, P. Giraudet, F. Caudal at Inst. Nat. de la Propriété Intellectuelle, nb 07/06162, (2007).

D	Hydro	Dist (m)	X (m)	Y (m)	Z (m)
D1	H 1	5428	18501	9494	-1687
	H 2	4620	10447	4244	-1677
	H 3	2514	14119	3034	-1627
	H 4	1536	16179	6294	-1672
	H 5	3126	12557	7471	-1670
	H 6	4423	17691	1975	-1633
D2	H 7	1518	10658	-14953	-1530
	H 8	4314	12788	-11897	-1556
	H 9	2632	14318	-16189	-1553
	H 10	3619	8672	-18064	-1361
	H 11	3186	12007	-19238	-1522

TABLE I. Hydrophones positions: Dist=Distance to the barycenter of the set. (H4 and H5 are out of order)

% of common TDOA	Raw	TKM	SMF
Raw	5/13	3%	2%
TKM	12%	103/57	14%
SMF	10%	23%	387/143

TABLE II. In diagonal, the number of positions obtained in D1/D2. Above the diagonal of this table, we give the % of common TDOA in D1, and under it, the % of common TDOA in D2.

5min chunks	0-5	5-10	10-15	15-20
ROSA Lab	4.3	5.3	4	3.6
PIMC TKM	4	4	4	3
SOEST	4	4	4	2
PIMC SMF	4	4	4	2

TABLE III. Counting number estimations of whales in D1. First row is the five minutes chunks of D1, second is the averaged number of whales estimations from ROSA Lab, third is our estimations (PIMC SMF). The estimates of the SOEST lab Hawaii are given by personnal communication from E.M.Nosal (preliminary results form Bellop model).

List of Figures