The Signal-to-Noise Ratio of Human Divers

Edmund Brekke
University Graduate Center
Kjeller, Norway
Email: edmund@unik.no

Oddvar Hallingstad
University Graduate Center
Kjeller, Norway
Email: oh@unik.no

John Glattetre
Kongsberg Maritime
Horten, Norway
Email: jhg@kongsberg.com

***Abstract*— A challenging problem in underwater acoustics is the detection and tracking of human divers. In order to assess the suitability of different tracking methods on this problem, an evaluation of the signal-to-noise ratio (SNR) is required.**

This paper investigates the SNR of a human diver using two approaches. The first approach is a moment-based target power estimator which uses the amplitudes of measurements validated by the tracking method. The second approach is by means of the receiving operating characteristic (ROC).

The integrated Probabilistic Data Association (IPDA) is used to track the diver. To the best of our knowledge, this is the first well-documented implementation of the IPDA on real data.

I. Introduction

Sonar detection and tracking of divers are important in port protection against underwater intruders. This is a challenging problem for several reasons. First, the target strength of a diver can be rather low and unpredictable, being a complicated function of aspect angle and frequency. The human lungs and the oxygen tank are thought to be the largest contributors [1]. Second, the harbor environment is at the same time characterized by abundant reverberation from the seafloor and the sea surface which severely deteriorates the relative target strength. Release of air bubbles from open circuit breathing equipment can make the diver more detectable, but is also likely to increase the amount of clutter in the environment.

The problem of tracking a target in clutter is addressed by *tracking methods* such as the classical Probabilistic Data Association filter (PDAF) [2]. The problem of detecting a target, in the sense of confirming its existence, is addressed by more general methods for *track management* (TM), which address both track initiation, confirmation, maintenance and termination. Although TM is most often done by means of more or less heuristic rules, the last 20 years have seen several more rigorous existence-based TM methods being proposed. These methods, which attempt to evaluate the probability that a target is present, include the integrated Probabilistic Data Association (IPDA) [3], methods based on finite set statistics (FISST) [4] and some track-before-detect (TBD) methods [5]. The multiple hypothesis tracker (MHT) [2 pp. 334-354] also addresses the TM problem.

The tracking of human divers was addressed in [6], where the PDAF was modified to provide enhanced discrimination against measurements caused by the bubble cloud, also referred to as *wake*, released by open circuit breathing equipment. Further improvements of this tracking method have been submitted to this conference in a separate paper [7]. The TM problem was in [6] solved by means of the M/N-logic [2 p. 104].

The efficiency and simplicity of PDAF-based tracking methods come at the price of sub-optimality. The suitability of such methods depends on several scenario characteristics. Among these are the maneuverability of the target as represented by the plant noise, the sensor resolution as represented by the measurement noise, and most importantly the signal-to-noise ratio (SNR). The higher the noises are, the higher the SNR must be for loss of track to be avoided. For low SNR it may be necessary to use amplitude information (AI) to improve the data association, leading to the PDAFAI method [8, 2 p. 247]. It may also be necessary to switch to more robust, but also vastly more expensive, tracking methods such as the MHT. The need for an improved basis for making such decisions is the underlying motivation of this paper.

To the best of our knowledge, this paper presents the first experimental report on the SNR of human divers available in the open literature. This investigation is done using both moment-based estimators and the receiving operating characteristic (ROC) on the same data as were as were treated in [6]. While the authors of [6] focused on the tracking of a diver with an open breathing system, this paper restricts attention to a rebreather with a fully closed breathing system. This allows us to investigate the target return from the diver itself without being bothered by surrounding air bubbles.

In this paper the IPDA is used to establish a track on the diver. To the best of our knowledge, this is the first well documented implementation of the IPDA on real data. This is important, because the recent surge in research on existence-based TM has only to a limited degree been followed by subsequent implementations of these methods on real data. The very few papers on existence-based TM which report implementations in the real world (e.g. references in [4 pp. 613-615]) have not assessed the difficulty of the scenario.

Our usage of the IPDA allows us therefore, for the first time, to confidently state that an existence-based TM method is able to deal with a given SNR in the real world. We are also able to make some statements regarding design of the false alarm rate for both the PDAF and the IPDA. We argue that in order to use the IPDA (or a similar method) for a target such as a human diver, one must allow a substantially higher false alarm rate than what is necessary for the PDAF.

In order to facilitate the IPDA, this paper also introduces a couple of methodological novelties, which include a new clustering method for measurement extraction and a track quality measure which we refer to as track coherence. Both of these contributions are thoroughly explained together with the entire tracking machinery used.

The remainder of the paper is organized as follows: Section II summarizes the assumptions underlying the IPDA. Section III describes this method together with preceding measurement extraction. In Section IV we describe the experiment reported herein. Results from the analysis of this experiment, including evaluation of the SNR, are described in Section V. The implications of these results on the performance of the PDAF and the IPDA are discussed in Section VI. Finally, a brief conclusion is given in Section VII.

II. System setup

The core component in our analysis of sonar data is the IPDA method, which is based on the following assumptions:

A1: There is at most one target present, whose existence or perceivability E_k can attain 3 discrete values

$$\begin{aligned} O_k &= \text{Target exists and is perceivable at time } k \\ I_k &= \text{Target exists but cannot be detected at time } k \\ N_k &= \text{Target does not exist at time } k. \end{aligned}$$

A2: Target existence follows a Markov chain:

$$\begin{aligned} P\{O_k\} &= p_{11}P\{O_{k-1}\} + p_{21}P\{I_{k-1}\} + p_{31}P\{N_{k-1}\} \\ P\{I_k\} &= p_{12}P\{O_{k-1}\} + p_{22}P\{I_{k-1}\} + p_{32}P\{N_{k-1}\} \\ P\{N_k\} &= 1 - P\{O_k\} - P\{I_k\}. \end{aligned}$$

By denoting $\boldsymbol{\eta}_k = [P\{O_k\}, P\{I_k\}, P\{N_k\}]^\mathsf{T}$ and collecting the transition probabilities p_{ij} in a matrix $\boldsymbol{\pi}$, we can write this more compactly as $\boldsymbol{\eta}_k = \boldsymbol{\pi}^\mathsf{T}\boldsymbol{\eta}_{k-1}$.

A3: The target moves according to a linear plant model

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{v}_k, \;\; \boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}).$$

A4: The past information about the target is summarized by

$$p(\boldsymbol{x}_k|\boldsymbol{Z}^{k-1}) \approx \mathcal{N}(\boldsymbol{x}_k\,;\,\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{P}_{k|k-1}).$$

Here $\boldsymbol{Z}^k = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k)$ is the set of all measurement sets up to and including time step k, while $\hat{\boldsymbol{x}}_{k|k-1}$ and $\boldsymbol{P}_{k|k-1}$ are the predicted state and its covariance.

A5: The measurement set $\boldsymbol{Z}_k = \{\boldsymbol{z}_k(1), \ldots, \boldsymbol{z}_k(m_k)\}$ consists of all measurements that fall within a limited validation gate $\mathcal{G}_k$ with volume V_k.

A6: There is at most one validated measurement with probability $P_\mathrm{D}P_\mathrm{G}$ from a perceivable target, where P_D is the detection probability and P_G is the gate probability.

A7: The target measurement is, when it exists, related to the target by a linear measurement model

$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{w}_k, \;\; \boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}).$$

A8: The remaining φ measurements in $\boldsymbol{Z}_k$ are false alarms with i.i.d. uniform spatial distributions and Poisson cardinality distribution:

$$\mu(\varphi) = \mathrm{Poisson}(\varphi\,;\,\lambda V_k) = e^{-\lambda V_k}\frac{(\lambda V_k)^\varphi}{\varphi!}.$$

A1-A8 are assumptions assumed by the IPDA, and not necessary requirements for its implementation. For instance, although A1 only allows one target to be present, it is straightforward to implement the IPDA in a multi-target setting by propagating several tracks in parallel.

III. Methodology

In this section we present the methods used in this paper for measurement *detection*, measurement *extraction* and target tracking. The amplitude models relied upon in the detection phase and SNR analysis are also presented in this section.

A. CFAR detection

Under the hypothesis $\mathcal{H}_0$ that no target is present in cell number j we model the complex background w_k^j as Gaussian and thus its amplitude a_k^j as Rayleigh distributed,

$$p(a_k^j|\eta) = \mathrm{Rayleigh}(a_k^j\,;\,\eta) = \frac{a_k^j}{\eta}\exp\left(\frac{-(a_k^j)^2}{2\eta}\right) \tag{1}$$

where η is the power of the background. Due to its tractability and parsimony [9 p. 669] as well as the central limit theorem (CLT), the Gaussian-Rayleigh model should be preferred unless solid evidence suggests a more refined model.

Under the hypothesis $\mathcal{H}_1$ that a target is present in cell number j we assume that the complex data z_k^j can be written as $z_k^j = s_k^j + w_k^j$ where s_k^j is the target signal. For the Swerling I target model, which models a fluctuating target, the amplitude $a_k^j = |z_k^j|$ is distributed according to

$$p(a_k^j|d,\eta) = \mathrm{Rayleigh}(a_k^j\,;\,d+\eta) \tag{2}$$

where d is the target power, also known as its cross-section. For the Swerling 0 target model, which models a non-fluctuating target, the amplitude is distributed according to

$$\begin{aligned} p(a_k^j|d,\eta) =& \mathrm{Rice}(a_k^j\,;\,\sqrt{2d},\eta) \\ =& \frac{a_k^j}{\eta}\exp\left(\frac{-((a_k^j)^2+2d)}{2\eta}\right) I_0\left(\frac{a_k^j\sqrt{2d}}{\eta}\right). \end{aligned} \tag{3}$$

Here $I_0(\cdot)$ is the modified Bessel function of the first kind with order zero [10 p. 374]. Since this paper only treats sonar data, there is only one pulse per scan.

As long as (1) holds i.i.d. for all resolution cells, a constant false alarm rate (CFAR) detector can be implemented by means of the following test:

$$a_k^i \gtrless T_k^i = \sqrt{c\cdot\widehat{\eta}_k^i} \tag{4}$$

where

$$\widehat{\eta}_k^i = \frac{1}{2M}\sum_{j\in G_k^i}(a_k^j)^2 \;\;,\;\; c = 2M\left(P_\mathrm{FA}^{-1/M} - 1\right). \tag{5}$$

Here $\hat{\eta}_k^i$ is the maximum likelihood estimator (MLE) of the background strength while T_k^i is the detection threshold. G_k^i is a window of M auxiliary cells chosen to lie in the vicinity of the cell under test, although with guardbands to prevent them from being affected by signal leakage. For reasons explained in

Section V-C we use a window which consists of 8 leading and 8 lagging cells in the range direction. Thus, all the auxiliary cells have the same bearing as the cell under test.

For the case of a Swerling I target the detection probability is given by the closed form formula

$$P_{\mathrm{D}} = \left(1 + \left(P_{\mathrm{FA}}^{-1/M} - 1\right)\frac{\eta}{d+\eta}\right)^{-M}. \tag{6}$$

For a Swerling 0 target it can be found by evaluating

$$P_{\mathrm{D}} = \frac{1}{c}\int S\left(\frac{\tau}{\eta}; 2, d\right)\mathrm{Gamma}\left(\frac{\tau}{\eta}; M, \frac{\eta}{M}\right)\mathrm{d}\tau \tag{7}$$

where

$$S(x; n, d) = \int_x^{\infty} \chi_n^2(u; d)\mathrm{d}u \tag{8}$$

is the survival function of the non-central chi-square distribution with n degrees of freedom [10 p. 942] and $\mathrm{Gamma}(\cdot)$ denotes the pdf of the Gamma distribution [9 p. 87].

B. Measurement extraction

The assumption A6, that at most one measurement can originate from the target, is pivotal in the derivation of most tracking methods including the PDAF and the IPDA. In order to ensure the validity of A6 on real data, clustering must be used in the extraction of measurements. Although it is possible to use these methods without any clustering, there are at least two reasons why one should attempt to maintain A6.

First, there is an increased danger that a tracker running on un-clustered measurements will fail to take all the target-affected cells into consideration. This can lead to filter inconsistency as illustrated in Figure 1.

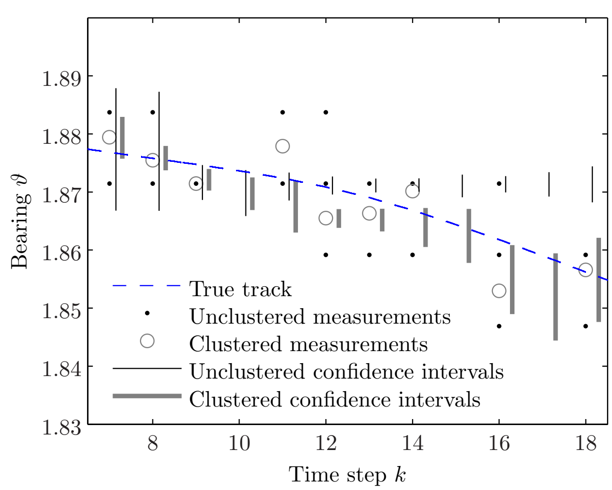

Fig. 1. Implementation of PDAF on simulated data with and without clustering. As long as several detections are received ($k = 7, 8$), the unclustered PDAF hedges on all of these. But when only a single detection is received ($k = 9$), its estimated covariance reduces drastically. It is thereafter ($k = 10, \ldots, 16$) satisfied with subsequent detections at the same spot, and eventually looses track. The clustered PDAF is on the other hand able to spot the motion of the target ($k = 7, 8, 9$), and avoids getting stuck in the same way.

Second, any additional detections from the target will rightfully be interpreted as clutter when A6 is assumed. In other words, a tracker running on un-clustered measurements will perceive an artificially high clutter intensity. This is not so serious for the PDAF, which often comes in a non-parametric version which ignores any knowledge about the clutter intensity. TM methods such as the IPDA are more sensitive to the perceived clutter intensity as the frequency of target measurement versus background measurements necessarily yields implications for the probability that a target is present.

We suggest that a clustering method for multibeam sonar data should attempt to satisfy the following criteria:

- C1: Clustering should primarily be done along the bearing direction.
- C2: Excessively large bearing clusters should be partitioned.
- C3: Only bearing clusters at adjacent ranges which obviously belong together should be joined.

The rationale behind C1 and C3 is that the data are characterized by a high degree of blurring along the bearing direction, and only limited blurring along the range direction. The criterion C2 should be satisfied in order to avoid so-called chaining [11 p. 684].

We address the criteria C1-C3 by devising a three-stage clustering process. During the first stage, detected cells from each row of the sonar image (i.e. constant range) are clustered independently using single linkage [11 p. 681]. These clusters are then split into possibly smaller clusters in order to avoid unreasonably large clusters. These clusters are then fed as input to another clustering method constructing the final clusters which possibly may contain cells from different rows. We refer to the corresponding three kinds of clusters as B-clusters, S-clusters and R-clusters.

The abstract structure is similar for the methods constructing B-clusters and R-clusters. In both cases the method takes a list of objects $\{C_j\}_{j=1}^N$ with values v_j and some kind of content z_j as input. Two objects C_p and C_i, where $v_p \geq v_i$, are clustered together if a boolean function $q(C_p, C_i)$ is satisfied.

For the construction of B-clusters the contents are simply the cell bearings ϕ^j as measured in resolution cells, and the values are the corresponding cell values. The boolean function tests whether two input cells are placed next to each other,

$$q(C_p, C_i) = \begin{cases} \text{true} & \text{if } \|\phi_p - \phi_i\| \leq 1 \\ \text{false} & \text{otherwise.} \end{cases} \tag{9}$$

The splitting of B-clusters into S-clusters is based on the notion of a *local master*. By a local master we mean a cell which is stronger than both its neighbors. Thus we proceed outwards from each local master and associate bearing cells to its new master until the territory of another local master is encountered. Precedence is given to stronger masters. Having done this for all ranges, the original collection of B-clusters has been expanded into a larger collection of S-clusters. Each S-cluster S_i is now dominated by a master cell $m(S_i)$ whose value we denote $v(m(S_i))$.

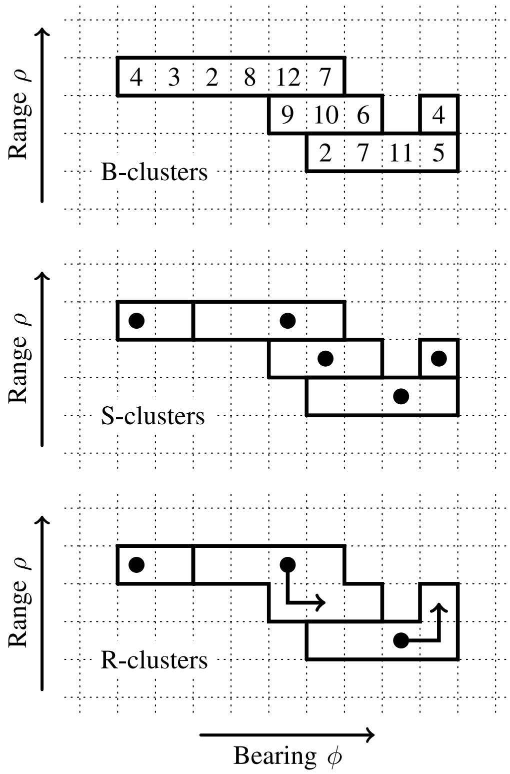

Fig. 2. Illustration of B, S and R-clusters together with master cells and the direct paths required by criterion C3C. While conventional single linkage would have collected these cells into one big lump, our method generates three distinct clusters instead. Situations like this occur frequently for high values of the false alarm rate P_{FA}.

Whether the target is believed to be visible at different ranges depends on the somewhat imprecise criterion C3. We phrase a more precise version of C3 by means of another boolean function $q^*(S_p, S_i)$ which takes two S-clusters S_p and S_i such that $v(m(S_p)) \geq v(m(S_i))$ as input. We define it to be true when and only when the following three criteria hold:

C3A: Master of S_p is adjacent to a cell in S_i.
C3B: Master of S_i is adjacent to a cell in S_p.
C3C: There exists a direct path from master of S_p to master of S_i which only changes range once and where the value of each cell along the path $\geq 0.9v(m(S_i))$.

Applying single linkage to the collection of S-clusters with this asymmetric proximity criterion produces the collection of R-clusters. Their centroids are evaluated by averaging the cell centers with weights proportional to their corresponding amplitudes. We define the value of an R-cluster as the amplitude of its strongest cell in the same way as for S-clusters. This is consistent with the definition of SNR used in e.g. [8].

C. PDAF and IPDA

While the PDAF is developed under the assumption that *exactly one* target moves in the surveillance region, the IPDA is developed under the assumption that *maximally one* target moves in the surveillance region. For the IPDA the full target state $\boldsymbol{s}_k = [\boldsymbol{x}_k^\mathsf{T}, E_k]^\mathsf{T}$ consists of a kinematic part $\boldsymbol{x}_k$, and an existence part E_k. The PDAF can be viewed as a special case of the IPDA obtained when E_k is kept constant. For future reference notice that the kinematic state $\boldsymbol{x}_k = [\boldsymbol{\rho}_k^\mathsf{T}, \mathbf{v}_k^\mathsf{T}]^\mathsf{T}$ in this particular implementation consists of position $\boldsymbol{\rho}_k = [x_k, y_k]^\mathsf{T}$ and velocity $\mathbf{v}_k = [\mathrm{v}_{xk}, \mathrm{v}_{yk}]^\mathsf{T}$.

The IPDA calculates the updated state estimate $\hat{\boldsymbol{x}}_{k|k}$, the predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$ and their associated covariances $\boldsymbol{P}_{k|k}$ and $\boldsymbol{P}_{k|k-1}$ using the same Kalman-filter like formulas as the PDAF. For brevity these formulas are omitted here. The reader can find them in [2 pp. 130-136] or [7].

The IPDA works by evaluating the association weights

$$\beta_k(i) = \frac{P\{O_k, \theta_k(i)|\boldsymbol{Z}^k\}}{P\{O_k|\boldsymbol{Z}^k\} + P\{I_k|\boldsymbol{Z}^k\}} \tag{10}$$

corresponding to the events

$$\begin{aligned}
\theta_k(0) =& \{\text{ no measurement is from the target }\}\\
\theta_k(1) =& \{\text{ measurement 1 is from the target }\}\\
&\vdots\\
\theta_k(m_k) =& \{\text{ measurement } m_k \text{ is from the target }\}
\end{aligned}$$

together with the existence probabilities $P\{O_k|\boldsymbol{Z}^k\}$, $P\{I_k|\boldsymbol{Z}^k\}$ and $P\{N_k|\boldsymbol{Z}^k\}$ in an integrated manner. These are available through the following formulae:

$$\begin{aligned}
P\{O_k|\boldsymbol{Z}^k\} &= \frac{1-\delta_k}{1-\delta_k P\{O_k|\boldsymbol{Z}^{k-1}\}} P\{O_k|\boldsymbol{Z}^{k-1}\} \qquad (11)\\
P\{I_k|\boldsymbol{Z}^k\} &= \frac{1}{1-\delta_k P\{O_k|\boldsymbol{Z}^{k-1}\}} P\{I_k|\boldsymbol{Z}^{k-1}\}\\
P\{N_k|\boldsymbol{Z}^k\} &= 1 - P\{O_k|\boldsymbol{Z}^k\} - P\{I_k|\boldsymbol{Z}^k\}\\
\beta_k(0) &= \frac{(1-P_\mathrm{D}P_\mathrm{G})P\{O_k|\boldsymbol{Z}^{k-1}\} + P\{I_k|\boldsymbol{Z}^{k-1}\}}{(1-\delta_k)P\{O_k|\boldsymbol{Z}^{k-1}\} + P\{I_k|\boldsymbol{Z}^{k-1}\}}\\
\beta_k(i) &= \frac{P_\mathrm{D}P_\mathrm{G}\frac{V_k}{\hat{m}_k}p_k(i)P\{O_k|\boldsymbol{Z}^{k-1}\}}{(1-\delta_k)P\{O_k|\boldsymbol{Z}^{k-1}\} + P\{I_k|\boldsymbol{Z}^{k-1}\}}
\end{aligned}$$

where

$$\delta_k = \begin{cases} P_\mathrm{D}P_\mathrm{G} & m_k = 0\\ P_\mathrm{D}P_\mathrm{G} - P_\mathrm{D}P_\mathrm{G}\frac{V_k}{\hat{m}_k}\sum_{i=1}^{m_k} p_k(i) & m_k > 0 \end{cases} \tag{12}$$

and

$$\begin{aligned}
p_k(i) =& p(\boldsymbol{z}_k(i)|\boldsymbol{Z}^{k-1}, m_k, O_k, \theta_k(i))\\
=& \frac{1}{P_\mathrm{G}}\mathcal{N}(\boldsymbol{z}_k(i)\,;\, \boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{H}\boldsymbol{P}_{k|k-1}\boldsymbol{H}^T + \boldsymbol{R}). \qquad (13)
\end{aligned}$$

In case the clutter density λ is known, we can write $\hat{m}_k = \lambda V_k$. In reality this quantity should be estimated, and [3] suggests the estimator

$$\hat{m}_k = \max\left(0,\, m_k - P_\mathrm{D}P_\mathrm{G}P\{O_k|\boldsymbol{Z}^{k-1}\}\right) \tag{14}$$

which is used in our implementation of the IPDA. More sophisticated estimators have been suggested in [12] and [13].

The IPDA must in addition to the kinematic state also predict the existence probabilities. In accordance with the compact notation of assumption A2 we express the existence prediction using a matrix multiplication:

$$\boldsymbol{\eta}_{k|k-1} = \boldsymbol{\pi}^\mathsf{T}\boldsymbol{\eta}_{k-1|k-1}. \tag{15}$$

D. Use of the IPDA and track coherence in TM

Track management is carried out using a three-stage process where the stages represent increasing track quality. Tracks in the three stages are termed preliminary, confirmed and coherent tracks.

Any two consecutive measurements within a small region generate a preliminary track through two-point differencing. Such a preliminary track is then propagated by the IPDA until it either is confirmed or rejected. The former event occurs if the existence probability $1 - P\{N_k\}$ exceeds a threshold T_c. The latter event occurs if this has not happened during the last 10 time steps, or if the non-existence probability $P\{N_k\}$ exceeds a threshold T_t. A confirmed track is propagated by the IPDA and terminated if $P\{N_k\}$ exceeds T_t. Both these thresholds are determined according to the recommendations of [12] (cf. Table I).

The IPDA is able to filter out a large amount of false tracks, but far from all of them. There will for example be several tracks on stationary clutter. A third stage is therefore necessary in order to distinguish interesting tracks from not so interesting tracks. We suggest what we call *track coherence* as a useful tool for this. Define $\boldsymbol{u}_k = \boldsymbol{\rho}_k - \boldsymbol{\rho}_{k-1}$. Denote the angle between $\boldsymbol{u}_k$ and $\mathbf{v}_k$ by γ_k. The track coherence c_k is then defined as

$$c_k = \frac{1}{k}\sum_{l=1}^{k}\cos\gamma_l = \frac{1}{k}\sum_{l=1}^{k}\frac{\mathbf{v}_l^{\mathsf{T}}\boldsymbol{u}_l}{\|\mathbf{v}_l\|\|\boldsymbol{u}_l\|}. \tag{16}$$

A confirmed track is declared to be a coherent track if c_k exceeds a threshold T_u while at the same time $P\{O_k\}$ exceeds another threshold T_o. A confirmed track must have been maintained as a confirmed track for at least N_c time steps before it is allowed to be considered for this test. A coherent track is also propagated by the IPDA, and terminated if $P\{N_k\}$ exceeds the threshold T_t.

The idea of track coherence is to pick up targets whose motion appears to be goal oriented. The track coherence of a target moving in a purely random pattern will be close to zero, while the track coherence of a target moving along a straight line will be close to one. The usage of track coherence require that the averaging of (16) can be done over several samples. We therefore only use track coherence after the target has moved a minimal distance d_c. This adds to the delay that must be accepted before a track can be declared coherent. The concept of track coherence can most likely be made more rigorous in a proper probabilistic framework, but for our purpose the simple heuristic of (16) is sufficient.

E. Target power estimation

Based on reasoning presented in [8], we suggest the following moment-based estimator $\hat{d}_k$ of the target power d at time step k,

$$\widehat{d}_k = \max\left(\frac{C_k}{D_k}, \frac{B_k - C_k}{D_k}\right) \tag{17}$$

where

$$\begin{aligned}
B_k =& \frac{1}{2L}\left(\sum_{l=1}^{L}\frac{1}{m_{k-l+1}}\sum_{i=1}^{m_{k-l+1}} a_{k-l+1}^2(i)\beta_{k-l+1}(i)\right)\\
C_k =& \frac{1}{L}\left(\sum_{l=1}^{L}\frac{1}{m_{k-l+1}}\sum_{i=1}^{m_{k-l+1}} \hat{\eta}_{k-l+1}(i)\beta_{k-l+1}(i)\right)\\
D_k =& \frac{1}{L}\sum_{l=1}^{L}(1-\beta_{k-l+1}(0)).
\end{aligned}$$

The time lag L must be chosen according to a compromise between accuracy and non-stationarity.

IV. Experimental setup

The experiment reported in this paper was carried out in a roughly $7\,\mathrm{m}$ deep harbor basin in Oslofjorden outside Horten, Norway. Two divers were swimming towards a $90\,\mathrm{kHz}$ narrowband multibeam echosounder as illustrated in Figure 3. The first of these used ordinary open-circuit breathing equipment, and therefore left a trail of air bubbles. The tracking of this diver was thoroughly discussed in [6]. The second diver used closed breathing equipment, and therefore left no footprints except its own backscatter. Both divers were towing buoys with GPS units, so that a coarse ground truth could be obtained.

A. Filter model

The IPDA was implemented using the following matrices:

$$\boldsymbol{F} = \begin{bmatrix} 1 & T & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & T\\ 0 & 0 & 0 & 1\end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix} \tag{18}$$

$$\boldsymbol{Q} =\sigma_v^2\begin{bmatrix} T^3/3 & T^2/2 & 0 & 0\\ T^2/2 & T & 0 & 0\\ 0 & 0 & T^3/3 & T^2/2\\ 0 & 0 & T^2/2 & T\end{bmatrix}. \tag{19}$$

The measurement noise matrix $\boldsymbol{R}$ was given by a conversion from polar to cartesian coordinates as described in [2 pp. 36-41]. In polar coordinates we used the covariance matrix

$$\boldsymbol{R}_{\text{polar}} = C_{\boldsymbol{R}}^2\begin{bmatrix}\frac{\Delta r^2}{12} & 0\\ 0 & \frac{\Delta\vartheta^2}{12}\end{bmatrix} \tag{20}$$

with Δr and $\Delta\vartheta$ being the sizes of a resolution cell along the range and bearing directions and $C_{\boldsymbol{R}}$ being a tuning constant. In addition to these kinematic matrices, the IPDA needs specification of the existence transition matrix:

$$\boldsymbol{\pi}_p = \begin{bmatrix} 1-\alpha & 0 & \alpha\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad \boldsymbol{\pi}_c = \begin{bmatrix} 1-\psi-\zeta & \psi & \zeta\\ 0.48 & 0.5 & 0.02\\ 0 & 0 & 1\end{bmatrix}.$$

The matrix $\boldsymbol{\pi}_p$ was used for preliminary tracks, while the matrix $\boldsymbol{\pi}_c$ was used for confirmed and coherent tracks. All the constants in this section can be found in Table I. The constants σ_v and $C_{\boldsymbol{R}}$ were determined with basis in the discussion of Sections V-D and V-E.

TABLE I
FILTER MODEL PARAMETERS

Parameter	Value	Specification
N_S	1160	Number of time steps
T	$1.089\,\mathrm{s}$	Sampling time
σ_v	$0.05\,\mathrm{m/s^2}$	Plant noise
$C_{\boldsymbol{R}}$	1.5	Scale factor for measurement noise
Δr	$0.66\,\mathrm{m}$	Resolution cell size in range
$\Delta\vartheta$	$3.5\cdot 10^{-3}\,\mathrm{rad}$	Resolution cell size in bearing
P_{FA}	0.01	False alarm rate
P_{D}	0.65	Detection probability
P_{G}	0.99	Gate probability
g	10	Gate size
α	0.02	Death probability (prelim. tracks)
ψ	0.1	Invisibility probability (conf. tracks)
ζ	0.01	Death probability (conf. tracks)
T_c	0.95	IPDA confirmation threshold
T_t	0.74	IPDA termination threshold
T_o	0.99	IPDA threshold for coherence declaration
T_u	0.8	Track coherence threshold
N_c	4	Coherence declaration delay
d_c	$20\,\mathrm{m}$	Coherence minimal distance
L	60	Amplitude estimation time lag

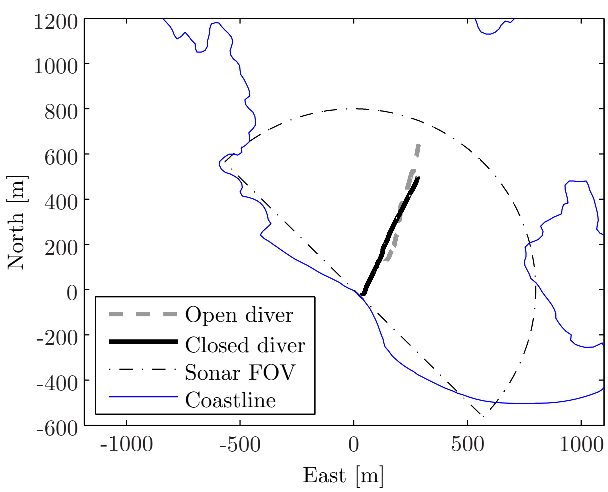

Fig. 3. Geographical setup for the experiment. The sonar's field of view (FOV) covered a half-circle with radius 800m in the inner harbor basin of Horten, Norway. The divers were swimming in south-west direction towards the sonar with an average velocity of slightly less than $0.5\mathrm{m/s}$. The coastline can easily be recognized on any local map of Horten.

V. Experimental results

In this section we present results from the experiment described in Section IV. The main purpose of this section is to assess the SNR of the closed breathing system diver, hereafter only referred to as closed diver. In order to validate the underlying assumptions, and to facilitate a discussion regarding the practical implications of a given SNR, we also assess several other quantities such as the *actual* false alarm rate, the plant noise and the measurement noise.

A. Moment-based SNR analysis

The estimator (17) can be used to evaluate the mean power d of the closed diver. For evaluation of the SNR

$$\mathrm{SNR} = 10\log\frac{d}{\eta} \tag{21}$$

we must also provide an estimate of the background strength η. Since such an estimate $\hat{\eta}_k(i)$ already exists for all the validated measurements $\boldsymbol{z}_k(i)$, we may estimate η by averaging over these measurement-conditioned estimates. If we are willing to assume temporal stationarity, a more stable background estimate can be found as

$$\hat{\eta}_k = C_k/D_k. \tag{22}$$

Moment-bases SNR estimates can be seen in Figure 4, while the underlying background and target power estimates are displayed in Figure 5. The smoothed SNR estimate based on (22) is most of the time located between 10 and $12\,\mathrm{dB}$. The stability of this curve indicates that range-dependent transmission loss had zero impact on the SNR, which again can be taken as evidence that reverberation was the sole dominant contributor to the background noise. Nevertheless, the fluctuations of the raw SNR estimate indicate that the effective SNR went as low as $4\mathrm{dB}$ over short periods of time.

In Figure 6 the accuracy of the raw SNR estimator is investigated using implementation on simulated data. The data were generated using the recipe of [8] as outlined in Section V-E. The SNR estimate is both highly uncertain and slightly biased, but nevertheless accurate enough to give us a rough idea about the difficulty of the tracking problem. Such uncertainties must be accepted in any short-term analysis of SNR.

B. ROC-based SNR analysis

The receiving operating characteristic (ROC), which describes P_{D} as a function P_{FA}, can be evaluated by counting the number of target detections and misdetections. In this context we treat any R-cluster who contains a cell which "touches" the cell in which $\boldsymbol{H}\hat{\boldsymbol{x}}_{k|k}$ is located as a detection (cf. Section V-E). The empirical ROC thus obtained is plotted as a thick dashed curve in Figure 7. Together with it we also display theoretical ROC curves of Swerling I and 0 targets as given by (6) and (7) with SNR values 10 and $12\,\mathrm{dB}$.

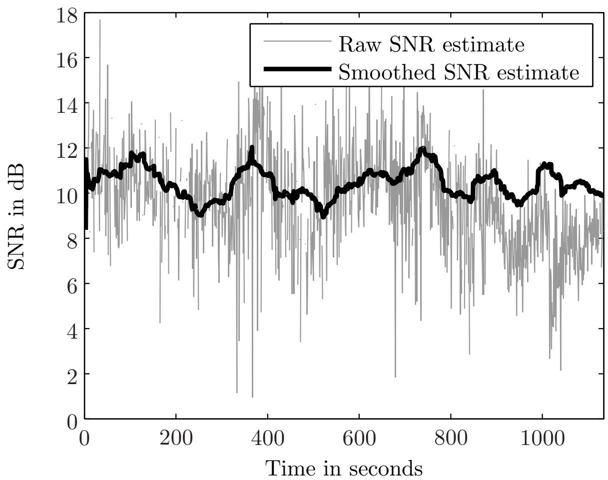

Fig. 4. The SNR for the closed breathing system diver evaluated by means of the estimator proposed in [8] with an estimation lag of 60 samples.

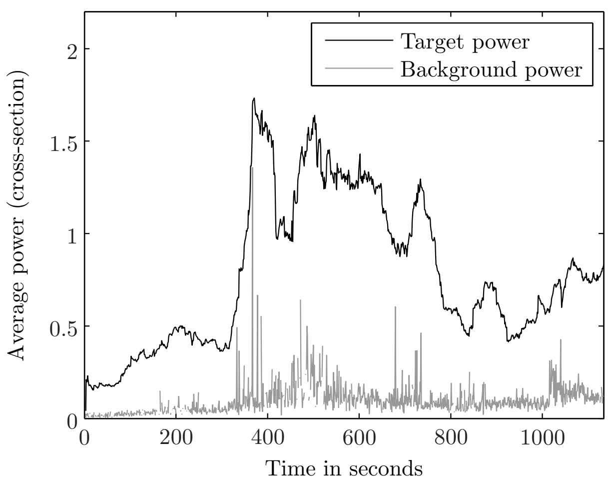

Fig. 5. Target and background power estimates displayed separately.

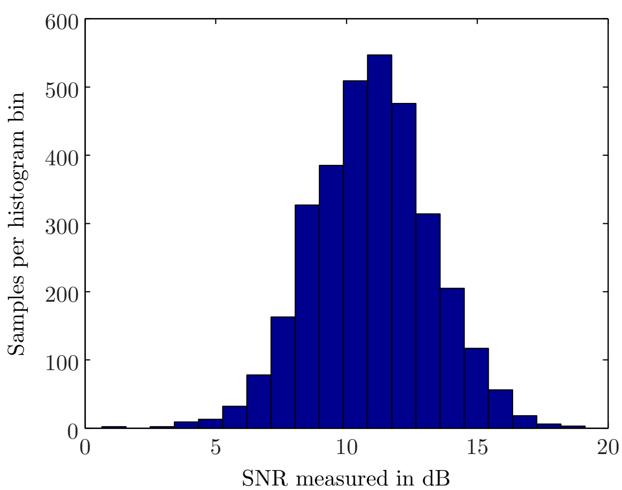

Fig. 6. Distribution of SNR estimate when the true SNR is 10 dB and 60 samples are used to estimate it.

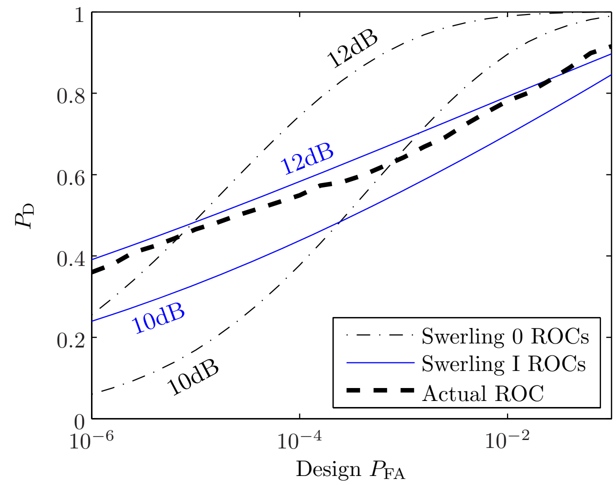

Fig. 7. ROC curve for the closed diver together with ROC curves for idealized Swerling 0 and Swerling I targets in Rayleigh distributed background noise.

This figure supports the assumption of a Swerling I target, since the dashed curve for most false alarm rates is located between the two Swerling I ROC's. It also reconfirms that the average SNR was somewhere between 10 and 12 dB.

C. Assessment of background assumptions

The analysis of Sections V-A and V-B was based on the Rayleigh background assumption (1). A standard approach to validate this assumption is given by the Kolmogorov-Smirnov test, which calculates the maximal distance between a reference cumulative distribution function (cdf) and the empirical cdf corresponding to a set of samples. A problem with this approach is that it systematically neglects the tails of the distributions, which in a detection problem actually are the most important parts of the distributions.

An alternative approach used here is to investigate whether the detector (4) yields the desired false alarm rate. In Figures 8 and 9 this was done for two different false alarm rates and three different CFAR windows. These figures were generated by putting all cells in a large slab of sonar data under the CFAR test and then counting the number of detections.

For the range only window (cf. Section III-A) the window size w refers to the number of auxiliary cells on either side of the cell under test. The dense square windows is square shaped with w rows of auxiliary cells above and below the guard band. The sparse square window is similar to the dense square window, but only with auxiliary cells along its boundary.

Since the increase in actual P_{FA} was negligible for the range only window, the Rayleigh assumption could be verified as far as this window was concerned, and it was therefore chosen as the preferred window to be used in our tracking system. The vastly increased false alarm rates encountered by the square windows could be taken as evidence against the Rayleigh assumption, but are more appropriately interpreted as evidence of non-stationarity along the bearing direction.

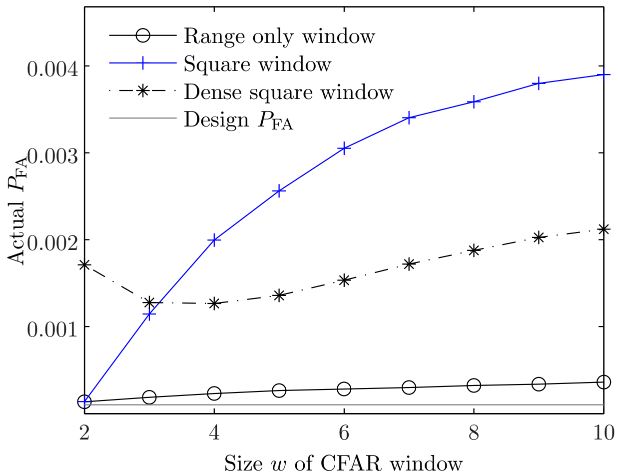

Fig. 8. Actual false alarm rates for design $P_{\mathrm{FA}} = 10^{-4}$.

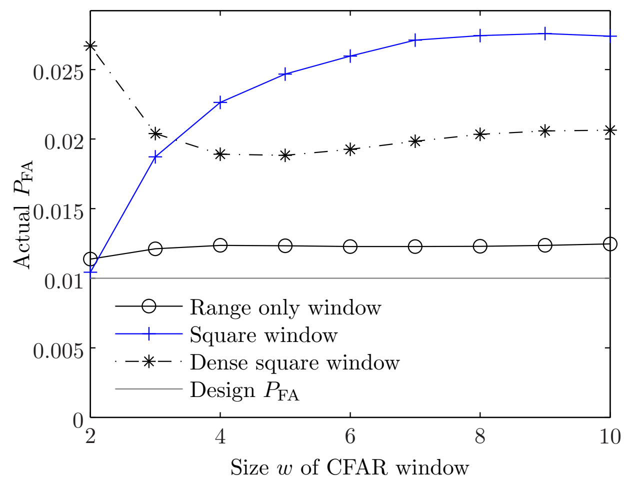

Fig. 9. Actual false alarm rates for design $P_{\mathrm{FA}} = 10^{-2}$.

D. Maneuverability and plant noise

The kinematics of the human diver is difficult to recover from the estimated track due to large estimation uncertainties. However, the kinematics of the GPS buoy is strongly linked to the kinematics of the diver. The GPS position data were smoothed using a Gaussian window in order to remove sporadic artifacts and emphasize the main tendencies of movement. The velocity was then evaluated by differencing the smoothed data.

In Figure 10 we have displayed this GPS velocity together with the IPDA's velocity estimate for a particular time interval when the diver appeared to take a short rest before starting swimming again. The maximal slope of the velocity, i.e. the maximal acceleration maintained over more than a couple of seconds, can be used as a guideline for the plant noise. From visual inspection we have arrived on $\sigma_v = 0.5\mathrm{m/s}^2$ as a reasonable value. This value was also used for the same data in [6]. It should be noted that this value represents an upper bound on the plant noise. We found that the IPDA was able to track the closed diver for values of σ_v down to $0.2\mathrm{m/s}^2$.

E. Measurement noise

For an ideal sensor with no blurring, the measurement of a non-fluctuating point target can be modeled as uniformly distributed inside the resolution cell occupied by the target. Approximating such a pdf by a Gaussian using moment-matching yields a measurement noise matrix on the form

$$\boldsymbol{R}_0 = \begin{bmatrix} \frac{\Delta r^2}{12} & 0 \\ 0 & \frac{\Delta \vartheta^2}{12} \end{bmatrix}$$

where Δr and $\Delta\vartheta$ are the sizes of a resolution cell along the range $[\mathrm{m}]$ and bearing $[\mathrm{radians}]$ directions, respectively.

The concept of measurement noise is less well-defined for a fluctuating point target in the presence of sensor blurring.

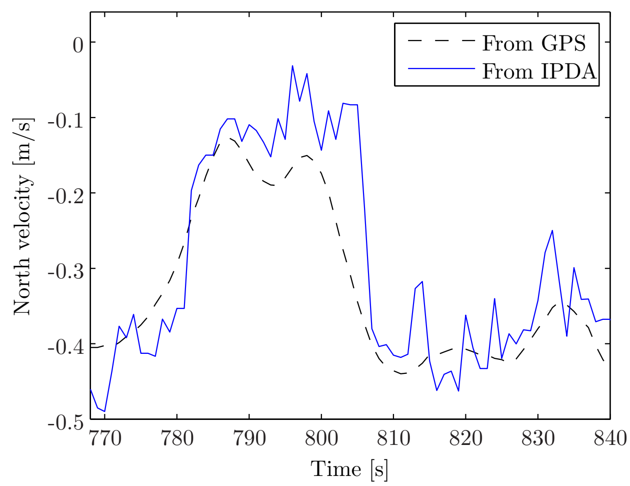

Fig. 10. Target velocity over a limited time interval.

Since a target detection contains both signal from the background and from the target, it is not obvious which detections are to be considered from the target and which are to be considered from the background. In other words, there is a soft transition between measurements which obviously originate from the target and measurements which originate from the background.

In order to make the concept of measurement noise well-defined we insist that the R-cluster which contains the target cell should be considered to be the target detection. If such an R-cluster does not exist, the strongest R-cluster which contains at least one cell adjacent to the target cell will be considered to be from the target.

The measurement noise of an extracted centroid depends on quantities such as the SNR and the sonar's point spread function (PSF) as well as the clustering method [14, 2 pp. 566-583]. In order to investigate how these affect the centroid statistics for the clustering scheme of Section III-B, we carried out Monte-Carlo simulations using the simulator described in

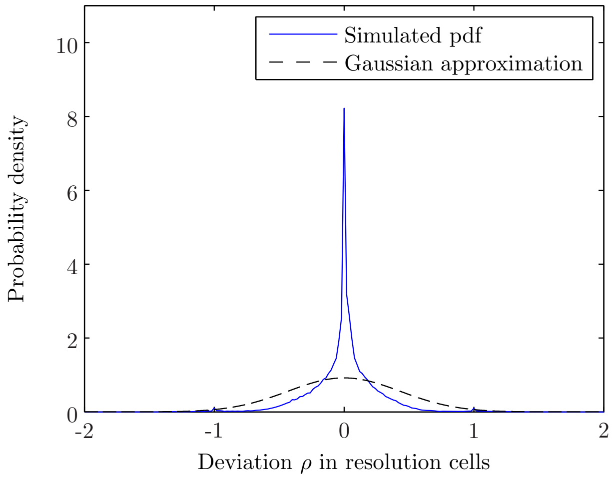

Fig. 11. Range measurement noise.

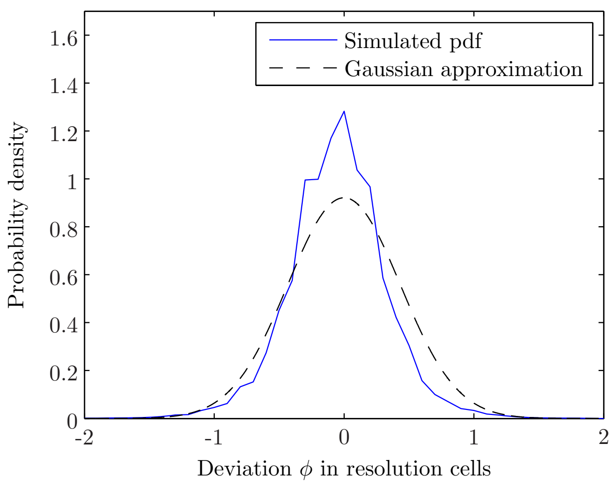

Fig. 12. Bearing measurement noise.

[8]. This simulator is based on the compound K-distribution and generates correlations similar to those found in real radar or sonar data. We simulated a $10\,\mathrm{dB}$ Swerling I target in a K-distributed background with shape parameter 8 and global correlation length 8, yielding a near-Rayleigh output. The PSF was approximated as

$$h(\rho, \phi) = h_\rho(\rho) h_\phi(\phi) = \operatorname{sinc}\left(\frac{\rho}{A}\right) \exp\left(-\frac{\phi^2}{2B^2}\right)$$

where ρ and ϕ are range and bearing relative to the target measured in resolution cells, $A = 1.0$ resolution cells and $B = 1.3$ resolution cells. These values correspond to 3dB beamwidths of 0.8 and 2.2 resolution cells, respectively.

Simulation results are displayed for a $10\,\mathrm{dB}$ target in Figures 11 and 12. We decided with basis in these results to set the filter model measurement noise matrix equal to

$$\boldsymbol{R}_{\text{polar}} = 1.5^2 \boldsymbol{R}_0. \tag{23}$$

The resulting bivariate Gaussian is illustrated by the dashed curves in Figures 11 and 12. We found that reducing the values in $\boldsymbol{R}_{\text{polar}}$ quickly led to loss of track or track termination. Increasing these values had less of an impact on the tracking performance, but caused an excess of false confirmed tracks.

VI. Discussion

Given knowledge about the SNR as well as the covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, one can infer a great deal about how challenging a tracking scenario is. Several tools have been developed for predicting the performance of the PDAF. One of these tools, the modified Riccati equation (MRE), is applied to the diver tracking problem in Section VI-A, before we discuss the detection performance of the IPDA in Section VI-B.

A. The PDAF: MRE with and without AI

The expected RMSE can be calculated for the PDAF by means of the MRE. Recent research [15] has extended this analysis to the PDAFAI, which uses the amplitudes of the backscattered measurements to improve its data association. In this section we implement the MRE with and without AI for a $10\,\mathrm{dB}$ Swerling I target in Rayleigh distributed background noise. For the noise covariances we use

$$\sigma_v = 0.05\mathrm{m/s}^2, \quad \boldsymbol{R} = \begin{bmatrix} (0.289\mathrm{m})^2 & 0 \\ 0 & (0.866\mathrm{m})^2 \end{bmatrix}.$$

The mathematical details of the MRE with and without AI are elaborated in [15], but omitted here for the sake of brevity.

The calculated steady-state RMSE is in Figure 13 plotted as a function of the false alarm rate P_{FA}. These curves justify the rather high false alarm rate 10^{-2}, which appears to be close to the minimizers of both the RMSE with AI and the RMSE without AI. It can also be seen that increasing P_{FA} beyond 10^{-2} leads to rapid deterioration of the PDAF. The PDAFAI is able to beat this divergence for much higher values of P_{FA}, and the usage of AI should therefore be considered if one insists on employing such a high false alarm rate.

The deterioration is much slower when P_{FA} is decreased, indicating that the PDAF can safely be used with lower false alarm rates as done in [6], obliterating the need for AI.

B. The IPDA: Detection performance considerations

Although the detection performance of the IPDA on its own was a bit disappointing, it performed very well when equipped with track coherence. Out of a total of 1663 preliminary tracks, the IPDA confirmed 48 tracks, of which roughly half were on stationary clutter close to the sonar. Only the one true track was declared coherent. This track was established after $5\mathrm{s}$, confirmed after $15\mathrm{s}$ and declared coherent after $46\mathrm{s}$. It was maintained for the next $1138\mathrm{s}$ until the diver climbed out of the water and into a boat.

However, the IPDA failed to maintain a continuous track for lower false alarm rates P_{FA}. The reason for this is that lowering P_{FA} increases the number of subsequent scans without target originating measurements. The number of such empty intervals encountered in our data are displayed for

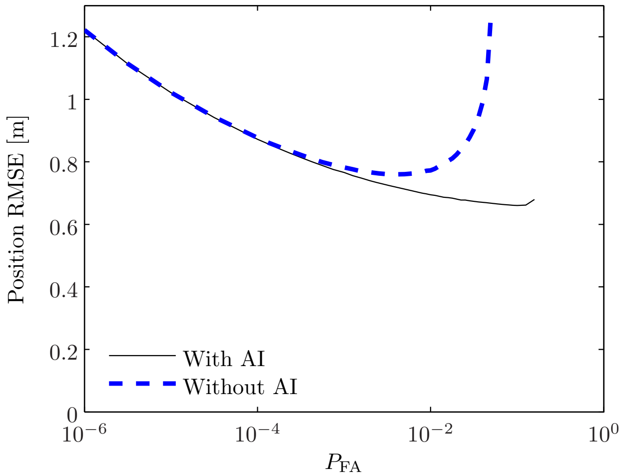

Fig. 13. Steady-state solution of MRE with and without AI.

TABLE II

DURATIONS OF INTERVALS WITHOUT DETECTION

Interval length	0	1	2	3	4	5	6	7	Max
$P_{\mathrm{FA}}\ 1\cdot10^{-4}$	447	98	40	17	8	9	2	6	20
$P_{\mathrm{FA}}\ 5\cdot10^{-4}$	527	100	40	17	7	6	9	1	19
$P_{\mathrm{FA}}\ 1\cdot10^{-3}$	564	98	48	11	8	6	7	1	19
$P_{\mathrm{FA}}\ 5\cdot10^{-3}$	700	100	34	7	10	4	5	2	7
$P_{\mathrm{FA}}\ 1\cdot10^{-2}$	767	102	26	8	7	4	1	1	7
$P_{\mathrm{FA}}\ 5\cdot10^{-2}$	943	76	9	7	1	0	0	0	4
$P_{\mathrm{FA}}\ 1\cdot10^{-1}$	1015	51	6	5	0	0	0	0	3

different false alarm rates and interval lengths in Table II. The rightmost column displays the length of the longest empty interval encountered for each false alarm rate. It can be seen that the false alarm rate $5\cdot10^{-4}$ used in [6] required the tracker to maintain the track without support for up to 19 subsequent scans. As a consequence, the authors of [6] could not terminate a track before 20 empty scans had elapsed. Fortunately, no sudden maneuvers occurred during these intervals, and the PDAF was able to predict the correct location of the diver. As argued in Section VI-A, the PDAF suffers only a negligible performance loss for this false alarm rate, provided that the assumed noise characteristics of $\boldsymbol{Q}$ and $\boldsymbol{R}$ are valid.

The IPDA would on the other hand inevitably terminate the track in such a situation, unless its existence parameters ψ and ζ have been set to extremely conservative values. More precisely, track maintenance would require $\psi \to 1$, reflecting a general belief that a perceivable target most likely will become invisible at the next time step. Clearly, such a tuning of the IPDA would be rather dubious.

The IPDA should therefore only be applied to moderately weak targets (e.g. human divers) if one is willing to employ a sufficiently high false alarm rate, such as the value 10^{-2} used in this paper. Some of the more advanced TM methods which recently have been proposed suffer from more severe misdetection problems [16], and this statement can therefore be extrapolated to these methods as well.

VII. CONCLUSION

The experimental results reported in this paper indicate that the SNR of a human diver in multibeam sonar data can be expected to lie around $10\,\mathrm{dB}$ - $12\,\mathrm{dB}$. Furthermore, the paper has demonstrated that such a target can be tracked by the IPDA insofar as the false alarm rate, and thus the corresponding detection probability, is high enough.

The results presented here are valid for a particular sonar in a particular environment. It would be interesting to compare these results with results from other experiments using one of the other diver detection sonars that currently are available. We also hope this paper can create inspiration for further experimental validation of other existence-based TM methods.

ACKNOWLEDGMENT

The authors would like to thank Kongsberg Maritime for providing the sonar data recorded by the active $90\,\mathrm{kHz}$ narrowband SM2000 multibeam imaging sonar from Kongsberg Simrad Mesotech Ltd.

REFERENCES

[1] S. Sarangapani, J. Miller, G. Potty, D. Reeder, T. Stanton, and D. Chu, "Measurements and modeling of the target strength of divers," in *Oceans 2005 - Europe*, vol. 2, 2005, pp. 952–956.

[2] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT, USA: YBS Publishing, 1995.

[3] D. Musicki, R. Evans, and S. Stankovic, "Integrated Probabilistic Data Association," *IEEE Transactions on Automatic Control*, vol. 39, no. 6, pp. 1237–1241, 1994.

[4] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[5] S. Davey, M. G. Rutten, and B. Cheung, "A Comparison of Detection Performance for Several Track-before-Detect Algorithms," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008.

[6] A. Rødningsby and Y. Bar-Shalom, "Tracking of Divers using a Probabilistic Data Association Filter with a Bubble Model," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1181–1193, Jul. 2009.

[7] E. Brekke, O. Hallingstad, and J. Glattetre, "Target Tracking in State Dependent Wake Clutter," to appear in Proceedings of IEEE Oceans Conference, Sydney, Australia, 2010.

[8] ——, "Tracking Small Targets in Heavy-Tailed Clutter Using Amplitude Information," accepted for publication in IEEE Journal of Oceanic Engineering.

[9] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2002.

[10] M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions*. Dover, 1965.

[11] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 5th ed. Prentice Hall, 2002.

[12] N. Li and X. R. Li, "Target Perceivability and its Applications," *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2588–2604, 2001.

[13] D. Musicki, S. Suvorova, M. Morelande, and B. Moran, "Clutter Map and Target Tracking," in *Proceedings of the 8th International Conference on Information Fusion*, Philadelphia, PA, USA, Jul. 2005.

[14] B. J. Slocumb, "Surveillance radar range-bearing centroid processing," in *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 4473, San Diego, CA, United States, Jul. 2001, pp. 74–85.

[15] E. Brekke, "Clutter Mitigation for Target Tracking," 2010, PhD dissertation, submitted to Department of Engineering Cybernetics, NTNU, Trondheim, Norway.

[16] D. Fränken, M. Schmidt, and M. Ulmke, ""Spooky Action at a Distance" in the Cardinalized Probability Hypothesis Density Filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1657–1664, Oct. 2009.