Joint Passive Detection and Tracking of Underwater Acoustic Target by Beamforming-Based Bernoulli Filter with Multiple Arrays

Zhongyue Chen [1] and Wen Xu [1,2,*]

[1] College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China; joywithjoy@zju.edu.cn

[2] Key Laboratory of Ocean Observation-Imaging Testbed of Zhejiang Province, Zhoushan 316021, China

[*] Correspondence: wxu@zju.edu.cn

Abstract: In this paper, improved Bernoulli filtering methods are developed to deal with the problem of joint passive detection and tracking of an underwater acoustic target with multiple arrays. Three different likelihood calculation methods based on local beamforming results are proposed for the Bernoulli filter updating. Firstly, multiple peaks, including both mainlobe and sidelobe peaks, are selected to form the direction-of-arrival (DOA) measurement set, and then the Bernoulli filter is used to extract the target track. Secondly, to make full use of the informations in the beamforming output, not only the DOAs but also their intensities, the beam powers are used as the input measurement sets of the filter, and an approach based on Pearson correlation coefficient (PCC) is developed for distinguishing between signal and noise. Lastly, a hybrid method of the former two is proposed in the case of fewer then three arrays. The tracking performances of the three methods are compared in simulations and experiment. The simulations with three distributed arrays show that, compared with the DOA-based method, the beam-based method and the hybrid method can both improve the target tracking accuracy. The processing results of the shallow water experimental data collected by two arrays show that the hybrid method can achieve a better tracking performance.

Keywords: Bernoulli filter; target tracking; multiple array processing; beamforming; random finite set

1. Introduction

In many target detection and tracking applications, detection and estimation are usually carried out separately under the Bayes framework [1]. Hypothesis tests based on the Neyman–Pearson or Bayes criterion [2] are applied to detect the existence of the target. Sequential probability ratio test (SPRT) which makes full use of all collected data based on the continuity of the target state [3] is also a detector. The estimators start working after the target has been declared to be detected. Examples of this type of sequential estimators include the Kalman filter (KF) [4], the extended Kalman filter (EKF) [5], the unscented Kalman filter (UKF) [6], and the particle filter (PF) [7]. Especially, the PF is a sequential Monte Carlo (SMC) method that uses a random particle system (states and weights) to approximate the relevant probability distribution. PF provides a generalized approximate solution for non-linear, non-Gaussian, and unconventional measurement problems where analytical solutions are difficult to obtain [8]. In PF applications, importance sampling is critical. To better apply the particle filter, researchers proposed appropriate resampling methods to reduce the degeneracy of the particle system [9,10]. Gaussian sum particle filter is a special kind of particle filter easier to implement where the relevant probability distribution is assumed as the sum of Gaussian distributions [11]. For low signal-to-noise (SNR) scenarios, track-before-detect (TBD) algorithms arise. In the TBD

algorithm, detection is not declared for every frame. Instead, detection is declared when the estimated track is returned after processing of several data frames [12]. In the multi-target tracking problems, multiple hypothesis on measure-to-track association can be filtered to achieve the tracking result. A comparison between multiple hypothesis Kalman filter and the multi-target particle filter can be found in Reference [13].

In passive acoustic target surveillance applications, however, a non-cooperative target may not always radiate a signal that is observable for detection and tracking. The noisy environment increases the uncertainty of obtaining the exact target existence information from available observations. A filter that can distinguish the existence of a target by noisy observations and can be able to resume track after losing the target for a while is needed.

Recently, target tracking based on random finite set (RFS) theory provides new opportunities for state-space processing which has been widely used in many tracking examples, such as video image recognition [14], infrared target tracking [15], and vehicle tracking detection [16]. Mahler and Vo systematically extended RFS theory to target tracking and made great progress in both theoretical research and applications [17]. Bernoulli filter is one of the RFS target tracking methods where the target existence and location state is assumed to be a Bernoulli RFS. It can be used to handle a variety of problems such as multiple on/off switching systems, detection imperfections, and inaccurate estimates [18]. Target tracking with bearings-only measurements by RFS-based filters has been studied from various aspects [19–22]. In Reference [23], Bernoulli filter tracking along with fusion is applied in an active multi-static acoustic sensor network. Gune considered the problem of detecting and tracking a high-frequency acoustic target in the frequency-azimuth plane obtained by an acoustic vector using Bernoulli filter [24]. However, to the best of our knowledge, modeling and filtering were conducted directly on the direction-of-arrival (DOA) estimations, and the DOA estimation process was usually neglected in the previous research.

In active radar monitoring, the measurement set of Bernoulli filter is assumed to be composed of a DOA estimation of the true target, whose error is a zero mean Gaussian variable, and several DOAs generated by clutters, which form a Poisson RFS [25]. However, the physical basis is different in the study of passive detection and tracking of an underwater target. There are no clutters, but, in estimating target DOA by passive beamforming in a noisy environment, outputs at the sidelobes may be higher than that at the mainlobe corresponding to the target position, resulting in significant errors. To deal with this problem, we jointly use beamforming and Bernoulli filters. A more general assumption is made for the underwater acoustic signal [26], where Gaussian noise is added directly to the received sound pressure. We conduct conventional beamforming (CBF) on the target signal band to process frequency snapshots at distributed arrays separately, and then employ the Bernoulli filter to track the target.

The main contribution of our work is that three different likelihood calculation methods based on local beamforming results are proposed for the Bernoulli filter. Firstly, we construct a DOA measurement set from the passive beamforming results (the DOA-based method). Similar to the method used in Reference [27], to increase the probability of selecting the target DOA, we select the angles corresponding to multiple peaks of the beamforming results as the measurement set. The set contains the DOA estimation corresponding to the real target (mainlobe) and several spurious peaks caused by sidelobes. Secondly, different from the method where only DOA peaks are conserved, we propose to use the beam power as the measurement set instead, because the beam power has not only DOA peak position information but also peak intensity information. Hence, we can use beam pattern to refine the target position (or DOA, or azimuth) from the data beam power. However, the tracking results can be trapped in the end-fire direction of one array where the beam pattern is flat and ambiguous. Therefore, at last, we propose a hybrid method, aiming to combine the advantages of the DOA-based method and the beam-based method in different situations. We use the practical acoustic model to generate sound pressures in the simulations to compare the proposed methods. Filters are also used to process experimental data collected in the real world.

Throughout this paper, we use uppercase calligraphic letters to denote the RFS variable, and $\varnothing$ denotes the emptyset. For any RFS $\mathcal{S}$, the cardinality of $\mathcal{S}$ is denoted by $|\mathcal{S}|$. Symbol $\setminus$ denotes the set-difference operation. Bold lowercase letters are used to denote column vectors, and bold uppercase letters denote matrices. Spaces of variables are denoted by uppercase blackboard bold letters, where $\mathbb{N}_0$ is the space of non-negative integers. We use blkdiag($\cdot$) to denote block diagonal matrix, where matrices in the brackets are diagonal blocks. $\mathcal{N}(\mathbf{m_x}, \mathbf{C_x})$ is denoted as a Gaussian distribution with mean $\mathbf{m_x}$ and covariance matrix $\mathbf{C_x}$, and the corresponding probability density function (PDF) with respect to $\mathbf{x}$ is denoted as $\mathcal{N}(\mathbf{x}; \mathbf{m_x}, \mathbf{C_x})$. We use f as a generic symbol for the finite set statistics (FISST) PDF of the RFS variables, while p is denoted as a generic symbol for the traditional PDF of the vector variables. For any variable or function ζ iterated in the prediction and update process of the Bayesian filtering framework, $\zeta_{k|k-1}$ is used to represent the prediction of ζ from time $k-1$ to time k, and $\zeta_{k|k}$ represents the update of ζ with all previous states and measurements from the beginning to time k.

The rest of the paper is organized as follows. The Bernoulli filter are briefly reviewed in Section 2, and the prediction step is provided. In Section 3, we establish two kinds of measurement sets derived from beamforming in a multiple-array system, which leads to different update process refers to DOA-based, beam-based, and hybrid Bernoulli filter, correspondingly. Section 4 presents the numerical implementation of the filters for practical usage. The proposed filters are implemented to process data from both simulations and real experiments. The performance results are shown in Section 5, and the paper is finally concluded in Section 6.

2. Bernoulli RFS Target State Model and Prediction Process

The concept of RFS was introduced by Mahler to describe the state of a randomly present target or multiple targets [17]. RFS-based Bayes filter aims to solve problems such as detection, estimation, and data association of target tracking within a unified paradigm [17,28].

In the RFS-based Bayes filter, target states and sensor measurements are modeled as RFSs instead of traditional vector formulation. Suppose that n_k targets present in the surveillance area at time k, and their state vectors are $\mathbf{x}_{k,1}, \ldots, \mathbf{x}_{k,n_k}$, respectively. The state of all the targets at time k is presented by an RFS on the state vector space, $\mathcal{X}_k = \{\mathbf{x}_{k,1}, \ldots, \mathbf{x}_{k,n_k}\} \in \mathcal{F}(\mathbb{X})$, where $\mathbb{X}$ is the state vector space. All the important information about the target, including the target number ($|\mathcal{X}_k| = n_k$) and the dynamic states of each target, has been contained. Likewise, the RFSs of sensor measurements are defined as $\mathcal{Z}_k = \{\mathbf{z}_{k,1}, \ldots, \mathbf{z}_{k,m_k}\} \in \mathcal{F}(\mathbb{Z})$, where $\mathbb{Z}$ is the measurement vector space.

The posterior probability density of the state variable is propagated in the predicting and updating procedure of the Bayes filter. By using the RFS state and measurement sets, we can extend the traditional Bayes filter from vector space to RFS framework. The predicting and updating process of the RFS posterior densities $f_{k|k}(\mathcal{X}_k|\mathcal{Z}_{1:k})$ can be expressed as follows [18]:

$$f_{k|k-1}(\mathcal{X}_k|\mathcal{Z}_{1:k-1}) = \int \phi_{k|k-1}(\mathcal{X}_k|\mathcal{X}_{k-1})f_{k-1|k-1}(\mathcal{X}_{k-1}|\mathcal{Z}_{1:k-1})\delta\mathcal{X}_{k-1}, \tag{1}$$

$$f_{k|k}(\mathcal{X}_k|\mathcal{Z}_{1:k}) = \frac{\varphi_k(\mathcal{Z}_k|\mathcal{X}_k)f_{k|k-1}(\mathcal{X}_k|\mathcal{Z}_{1:k-1})}{\int \varphi_k(\mathcal{Z}_k|\mathcal{X}_k)f_{k|k-1}(\mathcal{X}_k|\mathcal{Z}_{1:k-1})\delta\mathcal{X}_k}, \tag{2}$$

where $\mathcal{Z}_{1:k} = \mathcal{Z}_1, \ldots, \mathcal{Z}_k$ is the collection of all the previous measurements and $\varphi_k(\mathcal{Z}_k|\mathcal{X}_k)$ is the likelihood function of measurement set $\mathcal{Z}_k$, given the state $\mathcal{X}_k$. The RFS state is assumed to be Markovian with transitional density $\phi_{k|k-1}(\mathcal{X}_k|\mathcal{X}_{k-1})$. With the updated posterior densities $f_{k|k}(\mathcal{X}_k|\mathcal{Z}_{1:k})$ at every time step k, RFS state $\mathcal{X}_k$ can be estimated by criteria such as maximum a posteriori probability (MAP). Although the set integrals in the RFS iteration are usually intractable, solutions can be found for joint detection and tracking of a single target.

For the case of tracking a target which may randomly be present or disappeared in a surveillance area, Bernoulli RFS can be adopted to model the target state, whose cardinality distribution is

Bernoulli. Such RFS is either an empty set (with probability $1 - q$) or a singleton (with probability q), whose element $\mathbf{x}$ is the state vector of the target distributed following the PDF $p(\mathbf{x})$. Thus, the FISST PDF of a Bernoulli RFS is completely determined by the target existence probability q as

$$f(\mathcal{X}) = \begin{cases} 1 - q, & \text{if } \mathcal{X} = \varnothing \\ q \cdot p(\mathbf{x}), & \text{if } \mathcal{X} = \{\mathbf{x}\} \\ 0, & \text{if } |\mathcal{X}| \geq 2 \end{cases}. \tag{3}$$

Apparently, the Bernoulli RFS $\mathcal{X}$ can be completely determined by the corresponding existence probability q and vector PDF $q(\mathbf{x})$. Therefore, predicting and updating the RFS posterior densities $f_{k|k}(\mathcal{X}_k | \mathcal{Z}_{1:k})$ are equivalent to predict and update the corresponding existence probability and vector PDF. The predicted existence probability and vector PDF are denoted by $q_{k|k-1}$ and $p_{k|k-1}(\mathbf{x})$, respectively. $q_{k|k}$ and $p_{k|k}(\mathbf{x})$ are denoted as the posterior existence probability and vector PDF after update.

The kinetic characterization of the target is used in the prediction stage. This dynamic is described by the Markov transitional density (from time $k - 1$ to k):

$$\phi_{k|k-1}(\mathcal{X}|\varnothing) = \begin{cases} 1 - p_b, & \text{if } \mathcal{X} = \varnothing \\ p_b b_{k|k-1}(\mathbf{x}), & \text{if } \mathcal{X} = \{\mathbf{x}\} \\ 0, & \text{if } |\mathcal{X}| \geq 2 \end{cases}, \; \phi_{k|k-1}(\mathcal{X}|\{\mathbf{x}'\}) = \begin{cases} 1 - p_s, & \text{if } \mathcal{X} = \varnothing \\ p_s \pi_{k|k-1}(\mathbf{x}|\mathbf{x}'), & \text{if } \mathcal{X} = \{\mathbf{x}\} \\ 0, & \text{if } |\mathcal{X}| \geq 2 \end{cases}, \quad (4)$$

where $\phi_{k|k-1}(\mathcal{X}|\varnothing)$ describes the birth probability of target with state $\mathbf{x}$, $p_b = \Pr(|\mathcal{X}| = 1 || \mathcal{X}'| = 0)$ is the probability of event "target birth", and $b_{k|k-1}(\mathbf{x})$ denotes the distribution of the state of the born target on the vector space, which is defined in advance. When prior knowledge of the target is not available, the distribution of the born target is usually assumed to be a uniform distribution over the vector space. $\phi_{k|k-1}(\mathcal{X}|\{\mathbf{x}'\})$ is the survive probability, $p_s = \Pr(|\mathcal{X}| = 1 || \mathcal{X}'| = 1)$ is the probability of "target survive", and $\pi_{k|k-1}(\mathbf{x}|\mathbf{x}')$ is the transitional density on the vector space, which is related to the motion characteristics of the target. A typical target motion model is the constant velocity (CV) model. In a two-dimensional region, the target state vector is expressed by $\mathbf{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k]^{\mathrm{T}}$, where (x_k, y_k) is the Cartesian coordinate of the target and $(\dot{x}_k, \dot{y}_k)$ is the velocity component, correspondingly. The vector state transition equation of CV model is given by:

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{v}_k, \tag{5}$$

where $\mathbf{F}_k = \text{blkdiag}(\mathbf{G}_k, \mathbf{G}_k)$ is the state transition matrix, and $\mathbf{v}_k$ is the zero-mean Gaussian process noise with a covariance matrix $\mathbf{Q}_k = \text{blkdiag}(\Xi_k, \Xi_k)$. The diagonal blocks are given by

$$\mathbf{G}_k = \begin{bmatrix} 1 & T_k \\ 0 & 1 \end{bmatrix}, \; \Xi_k = \begin{bmatrix} T_k^3/3 & T_k^2/2 \\ T_k^2/2 & T_k \end{bmatrix} \cdot \varpi_1, \tag{6}$$

where T_k is the sampling interval from time $k - 1$ to time k, and ϖ_1 is the potential acceleration disturbance of the moving target. Therefore, the transitional density is Gaussian, $\pi_{k|k-1}(\mathbf{x}|\mathbf{x}') = \mathcal{N}(\mathbf{x}; \mathbf{F}_k \mathbf{x}', \mathbf{Q}_k)$.

With the definition of the Bernoulli RFS and its Markov transitional density, the prediction of RFS state PDF can be achieved by predicting the target existence probability and vector PDF as follows [18]:

$$q_{k|k-1} = p_b(1 - q_{k-1|k-1}) + p_s q_{k-1|k-1}, \tag{7}$$

$$p_{k|k-1}(\mathbf{x}) = \frac{p_b(1 - q_{k-1|k-1})b_{k|k-1}(\mathbf{x})}{q_{k|k-1}} + \frac{p_s q_{k-1|k-1} \int \pi_{k|k-1}(\mathbf{x}|\mathbf{x}')p_{k-1|k-1}(\mathbf{x}')d\mathbf{x}'}{q_{k|k-1}}. \tag{8}$$

Equations (7) and (8) completely specify the prediction stage of the Bernoulli filter.

In the next section, we present the update stage for the multi-array passive beamforming system. Different kinds of measurement sets are used, which lead to different forms of the update procedure.

3. Measurement Model for Multiple Arrays and Update Process

Using a multi-array system for joint passive detection and tracking of the underwater acoustic target, we first perform beamforming on each local array to obtain the relative DOA information of the target to the local array. Subsequently, in the Bernoulli filter, with known array locations, the tracking of the target state (position and velocity in two-dimensional Cartesian coordinates) is achieved using the beamforming results of all the arrays via triangulation.

3.1. Passive Local Array Beamforming

For the convenience of practical implementation, CBF is adopted to process the collected data. We divide the signal band of interest into multiple narrow subbands and perform narrow-band CBF on every subband, and then sum up the subband results to obtain the final beam power.

By fast Fourier transform (FFT), the recorded time domain signal of a given S-sensor array is converted to frequency-domain snapshots on M narrow subbands. The center frequency for the mth subband is denoted by f_m, $m = 1, \ldots, M$. For every subband, L snapshots are collected to form one data frame matrix, $\mathbf{R}_m = [\mathbf{r}_{m1}, \ldots, \mathbf{r}_{mL}]$. The lth column vector in matrix $\mathbf{R}_m$ is an $S \times 1$ vector. In the plane wave assumption, for signal oriented from a target with azimuth θ, elevation ψ, and range r, the propagation delay difference on the array is described by the steering (or replica) vector $\mathbf{a}(\theta) = \frac{1}{\sqrt{S}} \left[e^{-j\mathbf{k}^T\mathbf{p}_1}, \ldots, e^{-j\mathbf{k}^T\mathbf{p}_S} \right]^T$, where $\mathbf{p}_s$, $s = 1, \ldots, S$, are the Cartesian positions of the array elements, and $\mathbf{k} = -\frac{2\pi c}{f_m} [\sin(\psi)\cos(\theta), \sin(\psi)\sin(\theta), \cos(\psi)]^T$ is the wave number of the mth frequency bin with c being the sound speed. In the shallow water environment, depth is negligible compared to the range between target and array. That is, $\psi \approx \pi/2$, which leads to $\cos(\psi) \approx 0$ and $\sin(\psi) \approx 1$. Thus, in the local array beamforming step, we only need to estimate the azimuth (DOA) of the target. Defining the data sample covariance matrix as $\mathbf{C}_m = \frac{1}{L}\mathbf{R}_m\mathbf{R}_m^H$, where the superscript H indicates the conjugate transpose, the beam power of narrow band CBF is then $\mathbf{w}_m^H(\theta)\mathbf{C}\mathbf{w}_m(\theta)$. The weight vector of CBF is exactly the driving vector, $\mathbf{w}_m(\theta) = \mathbf{a}_m(\theta)$. Summation of the beam powers of all the M subbands leads to the basic broadband beamforming of the form

$$P(\theta) = \sum_{m=1}^{M} \mathbf{w}_m^H(\theta)\mathbf{C}_m\mathbf{w}_m(\theta). \tag{9}$$

When the featured frequencies of the target are included in the summed subbands, the target DOA will be enhanced in the beam power. The estimation of DOA is then given by $\hat{\theta} = \arg\max_{\theta \in (-\pi, \pi]} P(\theta)$.

In some scenarios, for example, when SNR is low or the array sensor spacing does not meet half-wavelength requirement, strong noise at the sidelobes and grating lobes will degrade the performance of DOA estimation. The correct peak still exists in the beam power but is usually lower than the spurious peaks. In other words, the beam pattern corresponding to the correct azimuth is corrupted but remains in the presence of noises and interferences. For the purpose of making full use of the beamforming results, we propose two kinds of derivative measurements of $P(\theta)$ in the following context and use them in the update procedure of Bernoulli filter.

3.2. Update Process in Multiple Array Systems

Suppose that a system equipped with N distributed arrays is deployed in the surveillance region. Every array in the system provides a measurement set at frame k. Denote the measurement set of the nth array at frame k by $\mathcal{Z}_k^{(n)}$, $n = 1, \ldots, N$. As the arrays are usually spatially separated, it is reasonable to make the assumption that measurements obtained at different arrays are statistically

independent. Thus, the entire likelihood can be expressed as the product of local likelihood functions of individual arrays:

$$
\varphi_k(\mathcal{Z}_k|\mathcal{X}_k) = \begin{cases} \prod_{n=1}^{N} \eta_{*,1}^{(n)}\left(\mathcal{Z}_k^{(n)}\Big|\mathbf{x}\right), & \text{if } \mathcal{X}_k = \{\mathbf{x}\} \\ \prod_{n=1}^{N} \eta_{*,0}^{(n)}\left(\mathcal{Z}_k^{(n)}\right), & \text{if } \mathcal{X}_k = \varnothing \end{cases},
\tag{10}
$$

where $\eta_{*,1}^{(n)}(\cdot)$ and $\eta_{*,0}^{(n)}(\cdot)$ denote the likelihood of the nth array. $\eta_{*,1}^{(n)}$ characterizes the similarity between measurement sets to a target-oriented measurement with vector state $\mathbf{x}$. On the contrary, $\eta_{*,0}^{(n)}$ is the likelihood to noise. Symbol $*$ will be replaced by specific methods in the following context. With the definition of the Bernoulli RFS and the likelihood function of the measurement RFS in Equation (10), the update step can be expressed by updating the target existence probability $q_{k|k}$ and the posterior PDF $p_{k|k}(\mathbf{x})$ on the vector space, respectively [18],

$$
q_{k|k} = \frac{q_{k|k-1} \int \ell_k(\mathcal{Z}_k|\mathbf{x}) p_{k|k-1}(\mathbf{x}) d\mathbf{x}}{1 - q_{k|k-1} + q_{k|k-1} \int \ell_k(\mathcal{Z}_k|\mathbf{x}) p_{k|k-1}(\mathbf{x}) d\mathbf{x}},
\tag{11}
$$

$$
p_{k|k}(\mathbf{x}) = \frac{p_{k|k-1}(\mathbf{x}) \prod_{n=1}^{N} \eta_1^{(n)}\left(\mathcal{Z}_k^{(n)}\Big|\mathbf{x}\right)}{\int p_{k|k-1}(\mathbf{x}) \prod_{n=1}^{N} \eta_1^{(n)}\left(\mathcal{Z}_k^{(n)}\Big|\mathbf{x}\right) d\mathbf{x}},
\tag{12}
$$

where

$$
\ell_k(\mathcal{Z}_k|\mathbf{x}) = \prod_{n=1}^{N} \frac{\eta_1^{(n)}\left(\mathcal{Z}_k^{(n)}\Big|\mathbf{x}\right)}{\eta_0^{(n)}\left(\mathcal{Z}_k^{(n)}\right)}
\tag{13}
$$

is the measurement likelihood ratio function.

The likelihood functions $\eta_{*,1}^{(n)}$ and $\eta_{*,0}^{(n)}$ are important in the update iteration, and will be introduced in the rest part of this section.

3.3. DOA-Based Method

In this section, we discuss the construction method of a DOA-based measurement set. In the case of low SNR, the sidelobe outputs may be higher than the mainlobe in the beam power. However, our simulations showed that, if we conserve certain number of peaks, the correct DOA information of the target has a high probability to be retained, which can thus be transferred to the subsequent filter. Therefore, for each local array, we search for λ largest peaks on the beam power curve $P(\theta)$. For an individual array, the azimuths corresponding to these peaks compose the measurement set $\mathcal{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_\lambda\}$. The measurement set can be divided into two parts $\mathcal{Z} = \mathcal{T} \cup \mathcal{F}$, where $\mathcal{T}$ is the RFS for the true target DOA, and $\mathcal{F}$ is the RFS for the spurious peaks. As mentioned before, true set $\mathcal{T}$ can be modeled as a Bernoulli RFS. When the true azimuth is not included in the selected peaks, $\mathcal{T}$ is an empty set. At each frame, let peak number λ be randomly generated according to the Poisson distribution with the mean of ν, so that the set $\mathcal{F}$ of spurious peaks is approximately a Poisson RFS whose PDF is defined as

$$
\kappa(\mathcal{F}) = e^{-\nu} \prod_{\mathbf{z} \in \mathcal{F}} \nu u(\mathbf{z}),
\tag{14}
$$

where $u(\mathbf{z})$ is the density of the false DOA peak $\mathbf{z}$. For the target generated true azimuth set, the conditional PDF is given by

$$
f(\mathcal{T}|\{\mathbf{x}\}) = \begin{cases} 1 - p_d(\mathbf{x}), & \text{if } \mathcal{T} = \varnothing \\ p_d(\mathbf{x}) p(\mathbf{z}|\mathbf{x}), & \text{if } \mathcal{T} = \{\mathbf{z}\} \end{cases},
\tag{15}
$$

where $p_d(\mathbf{x})$ is the probability of detecting the object in state $\mathbf{x}$; that is, for a given state $\mathbf{x}$, the probability of the corresponding azimuth peak is selected. Moreover, $p(\mathbf{z}|\mathbf{x})$ is the conventional likelihood function of true azimuth measurement due to the object $\mathbf{x}$. Similar to other Bernoulli filters, $p(\mathbf{z}|\mathbf{x})$ is assumed to follow a Gaussian distribution $\mathcal{N}(\mathbf{z}; \theta(\mathbf{x}), \sigma_B^2(\mathbf{x}))$. The mean value $\theta(\mathbf{x})$ is the correct DOA of the target $\mathbf{x}$, and the standard deviation $\sigma_B(\mathbf{x})$ equals to half of the 3-dB mainlobe width of the corresponding beam pattern [29].

In an N-array system, let $\mathcal{Z}^{(n)} = \mathcal{T}^{(n)} \cup \mathcal{F}^{(n)}$ be the azimuth peak set of the nth array. When the target is absent, $\mathcal{X} = \varnothing$ and the measured peak set is exactly the set of spurious peaks. Therefore, the likelihood is given by

$$\eta_{\text{doa},0}^{(n)} = \kappa(\mathcal{Z}^{(n)}). \tag{16}$$

When a target is present, the likelihood is given by

$$\eta_{\text{doa},1}^{(n)} = f(\varnothing|\{\mathbf{x}\}) \cdot \kappa(\mathcal{Z}^{(n)}) + \sum_{\mathbf{z} \in \mathcal{Z}^{(n)}} f(\{\mathbf{z}\}|\{\mathbf{x}\}) \cdot \kappa(\mathcal{Z}^{(n)} \setminus \{\mathbf{z}\})$$

$$= \kappa\left(\mathcal{Z}^{(n)}\right) \left[1 - p_d(\mathbf{x}) + p_d(\mathbf{x}) \sum_{\mathbf{z} \in \mathcal{Z}^{(n)}} \frac{p(\mathbf{z}|\mathbf{x})}{vu(\mathbf{z})} \right]. \tag{17}$$

The update process of the DOA-based method is completed by substituting the likelihood functions in Equations (16) and (17) into Equations (11)–(13).

3.4. Beam-Based Method

In the DOA-based method, we only retain the DOA peak positions from the beamforming outputs for filtering, but the intensity information in the beam power can also be useful. Therefore, we proposed a beam-based method in this section, where the entire beam power for $\theta \in (-\pi, \pi]$ of the arrays are used as the measurements, that is, $\mathcal{Z}_k^{(n)} \equiv P_k^{(n)}(\theta)$, $n = 1, \ldots, N$, where $P_k^{(n)}(\theta)$ denotes the beam power of the signal received by the nth array at time k. According to the definition of likelihood functions, $\eta_{*,1}^{(n)}$ should take large value when the beam power is close to those oriented from the true direction, while $\eta_{*,0}^{(n)}$ gets larger when the beam power is as random as the result of the omnidirectional noise. To get a likelihood function of these features, we propose to use the concept of correlation between the data beam power and the beam pattern for some assumed signal arrival direction.

For a fixed array, the element positions in the array are known a priori. Thus, we have known beam pattern (in power) for a given DOA $\theta_{\mathbf{x}}$ of some target state $\mathbf{x}$:

$$B(\theta; \theta_{\mathbf{x}}) = \frac{1}{M} \sum_{m=1}^{M} \|\mathbf{w}_m^{\mathrm{H}}(\theta)\mathbf{a}_m(\theta_{\mathbf{x}})\|_2^2, \tag{18}$$

where $\|\cdot\|_2$ is to take the ℓ_2-norm of a vector. The beam pattern of the nth array is denoted by $B^{(n)}(\theta; \theta_{\mathbf{x}})$. The correlation between the beam pattern and the noisy beam power can tell the likelihood of measurements to target state. In statistics, the Pearson correlation coefficient (PCC) is a measure of the linear correlation between two variables. The PCC between the measured beam power $P_k^{(n)}(\theta)$ and beam pattern $B^{(n)}(\theta; \theta_{\mathbf{x}})$ is given by

$$\rho_k^{(n)}(\mathbf{x}) = \frac{\sum_\theta \left(P_k^{(n)}(\theta) - \bar{P}_{\theta,k} \right) \left(B^{(n)}(\theta; \theta_{\mathbf{x}}) - \bar{B}_\theta(\theta_{\mathbf{x}}) \right)}{\sqrt{\sum_\theta \left(P_k^{(n)}(\theta) - \bar{P}_{\theta,k} \right)^2} \sqrt{\sum_\theta \left(B^{(n)}(\theta; \theta_{\mathbf{x}}) - \bar{B}_\theta(\theta_{\mathbf{x}}) \right)^2}}, \tag{19}$$

where $\bar{P}_{\theta,k}$ and $\bar{B}_\theta(\theta_{\mathbf{x}})$ are the means of $P_k^{(n)}(\theta)$ and $B^{(n)}(\theta; \theta_{\mathbf{x}})$ with respect to θ, respectively.

The PCC has a value between -1 and $+1$, where $+1$ is the perfect positive linear correlation, 0 is the no-linear correlation, and -1 is the perfect negative linear correlation. While the likelihood

functions take values in $[0, 1]$, mapping rules from PCC to likelihood functions are required. In our applications, when the local beamforming result does not match the beam pattern from the assumed direction, the PCC is negative or close to 0, indicating that there is only noise in this direction, so that the likelihood function to noise should be larger. On contrast, when the local beamforming result completely matches the beam pattern, PCC is equal to 1, indicating that there is a target in the assumed direction. The closer the PCC is to 1, the larger the target likelihood function should be set.

Note that negative PCCs are considered as mismatch here. Readers may be confused because negative correlations are usually considered as strong correlations in applications such as image recognition. That is, if the direction of a measured image are opposite to the original one, the PCC between them equals to -1, but they are still the same object (highly correlated). However, in our application, the PCC between the local beamforming result and the beam pattern being negative means that peaks appear at the DOA positions that should have been valleys, while false peaks appear at where they should not have been existed. This is not a match in our expectation. If the negative PCC was still considered as strong correlation, we would get wrong DOA estimation.

The desired mapping rules should satisfy the requirements listed above. For this purpose, we first define a non-negative PCC $\rho_{k+}^{(n)}(\mathbf{x})$, where $\rho_{k+}^{(n)}(\mathbf{x}) = \rho_k^{(n)}(\mathbf{x})$ if $\rho_k^{(n)}(\mathbf{x}) > 0$ and $\rho_{k+}^{(n)}(\mathbf{x}) = 0$ if $\rho_k^{(n)}(\mathbf{x}) \leq 0$. Then, we generate likelihood functions from $\rho_{k+}^{(n)}(\mathbf{x})$ as follows:

$$\eta_{\text{beam},1}^{(n)} = \frac{1}{\sqrt{2\pi}\sigma_\rho} \exp\left\{ -\frac{\left(\rho_{k+}^{(n)}(\mathbf{x}) - 1\right)^2}{2\sigma_\rho^2} \right\}, \quad \eta_{\text{beam},0}^{(n)} = \frac{1}{\sqrt{2\pi}\sigma_\rho} \exp\left\{ -\frac{\left(\rho_{k+}^{(n)}(\mathbf{x}) - 0\right)^2}{2\sigma_\rho^2} \right\}, \quad (20)$$

where σ_ρ is an adjustable probability mapping scale parameter. The update process of the beam-based method is completed by substituting the likelihood functions in Equation (20) into Equations (11)–(13).

3.5. Hybrid Method

The beam-based method is beneficial to improve the accuracy when the beamforming result highly matches the beam pattern of the target DOA. However, since there are wider peaks near the end-fire direction in the beam pattern of an individual local array, the tracking results can sometimes be trapped in the end-fire regions. If there are more than three arrays, the effect of such error can be eliminated thanks to the geometric relations between the arrays. In the case of fewer arrays, this problem leads to localization error. Although the DOA-based method may also have ambiguity problems in the end-fire direction, since only one or two DOA peaks are retained for filtering, it is less affected than the beam-based method where the operation of calculating the PCC will aggravate the performance degradation caused by the ambiguity in the end-fire direction. Therefore, the method using DOA measurements only is advantageous for protecting the true value from the effects of the end-fire direction in the filtering procedure, especially when the number of arrays is not enough. Both methods have their own advantages and disadvantages. To complement each other, we propose a hybrid approach combining these two choices. We define the new likelihood functions as the geometric mean of likelihood functions of the DOA-based and beam-based methods as follows:

$$\eta_{\text{hybrid},0}^{(n)} = \left[\eta_{\text{beam},0}^{(n)}\right]^\alpha \left[\eta_{\text{doa},0}^{(n)}\right]^{1-\alpha}, \quad \eta_{\text{hybrid},1}^{(n)} = \left[\eta_{\text{beam},1}^{(n)}\right]^\alpha \left[\eta_{\text{doa},1}^{(n)}\right]^{1-\alpha}, \quad (21)$$

where $\alpha \in [0, 1]$ is the portion coefficient to adjust the weights of beam-based and DOA-based likelihood in the hybrid method. When $\alpha = 1$, the function is the same as the beam-based likelihood, while $\alpha = 0$, it is exactly the same as that of the DOA-based method.

4. SMC Implementation of the Bernoulli Filter

The PF/SMC method provides a general framework for the implementation of Bernoulli filter [18]. The PDF $p_{k|k}(\mathbf{x})$ is approximated by a particle system $\{w_k^{(i)}, \mathbf{x}_k^{(i)}\}_{i=1}^{N_p}$, where $\mathbf{x}_k^{(i)}$ is the state of particle i and $w_k^{(i)}$ is the normalized particle weight, that is, $\sum_{i=1}^{N_p} w_k^{(i)} = 1$. The PDF is approximated by

$$\hat{p}_{k|k}(\mathbf{x}) = \sum_{i=1}^{N_p} w_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}), \tag{22}$$

where $\delta_{\mathbf{b}}(\cdot)$ is the Dirac function at point $\mathbf{b}$. In the prediction stage, the prediction of PDF in Equation (8) is completed by predicting the particle system

$$\hat{p}_{k|k-1}(\mathbf{x}) = \sum_{i=1}^{N_p+N_b} w_{k|k-1}^{(i)} \delta_{\mathbf{x}_{k|k-1}^{(i)}}(\mathbf{x}). \tag{23}$$

The particles are divided into two parts: N_b newly born particles and N_p persistent particles. The states of the birth particles are drawn from a pre-defined birth distribution, for example, uniform distribution. The birth particles are equally weighted by $w_{k|k-1}^{(i)} \propto p_b(1 - q_{k-1|k-1}) \cdot \frac{1}{N_b}$, $i = 1, \ldots, N_b$. The persistent particles are the extension of those particles at the previous time. Their states are drawn from the state transitional density $\mathbf{x}_{k|k-1}^{(i)} \sim \mathcal{N}(\mathbf{F}_{k-1}\mathbf{x}_{k-1}^{(i)}, \mathbf{Q}_{k-1})$ for $i = 1, \ldots, N_p$. The relationship between the weights of the persistent particles and the previous particle weights is given by $w_{k|k-1}^{(i)} \propto p_s q_{k-1|k-1} \cdot w_{k-1}^{(i)}$, $i = 1, \ldots, N_p$. Note that we use proportion rather than equality, because the weights will also be normalized in the prediction stage.

In the update equations of target existence probability, the integrals can be approximated using predicted particles as

$$\int \ell_k(\mathcal{Z}_k|\mathbf{x}) p_{k|k-1}(\mathbf{x}) d\mathbf{x} \approx \sum_{i=1}^{N_p+N_b} \ell_k(\mathcal{Z}_k|\mathbf{x}_{k|k-1}^{(i)}) w_{k|k-1}^{(i)}. \tag{24}$$

To update PDF $p_{k|k}(\mathbf{x})$, the weights of particles are updated by

$$w_{k|k}^{(i)} \propto \varphi(\mathcal{Z}_k|\{\mathbf{x}_{k|k-1}^{(i)}\}) w_{k|k-1}^{(i)}. \tag{25}$$

At the end of the update procedure, sampling importance resampling (N_p particles persist) is adopted to regularize the particle system [9]. With all these particles, the MAP estimation of the target state is approximately the weighted summation of the updated particles at time k, that is, $\sum_{i}^{N_p} w_k^{(i)} \mathbf{x}_k^{(i)}$.

5. Performance Tests

In this section, performance of the three Bernoulli filters is evaluated using simulational and experimental data.

5.1. Simulation Results

In the simulations, we test the performance of the DOA-based, beam-based and hybrid methods in a shallow water environment, as depicted in Figure 1. As shown in Figure 1a, in a 10 km × 10 km surveillance region, three horizontal line arrays (HLA) of length 256 m each with 24 uniformly spaced hydrophones are deployed at the seabed (of depth about 200 m). The centers of the HLAs are (6.0, 7.0) km (HLA1), (2.0, 2.0) km (HLA2) and (7.5, 3.5) km (HLA3), respectively. The HLAs are drawn as black dots in the figure. An acoustic target at depth 50 m moves with a constant velocity

$\dot{x}_t = 1.5 \text{ m/s}$ and $\dot{y}_t = 1.6 \text{ m/s}$. The source trace used is also presented by the red line. The source we test starts to appear from frame $k = 10$ at (3.0, 2.0) km and finally disappears at frame $k = 185$.
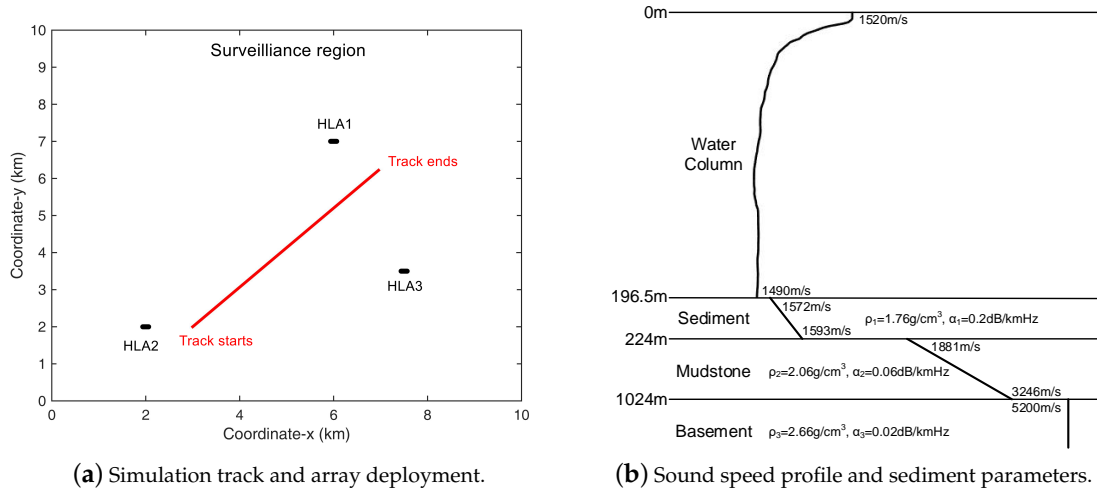

(**a**) Simulation track and array deployment.　(**b**) Sound speed profile and sediment parameters.

Figure 1. Simulation environment: (**a**) The surveillance region with three HLAs shown by black dots, and the red line is the target track. (**b**) The shallow water environment used in the simulation. The sound speed profile is chosen from the SWellEx96 experiment scenarios. The bottom is modeled as a sediment overlying a mudstone layer. The layer parameters are shown in the figure.

The featured frequency of the target signal is 50, 60, 70 and 80 Hz. The frequency snapshot of the received signal of the nth array at time k is given by $\mathbf{r}_k^{(n)} = A_k \mathbf{g}_k^{(n)} + \mathbf{n}_k^{(n)}$, where A_k is the complex Gaussian random amplitude of the source signal following the distribution $\mathcal{CN}(0, \sigma_s^2)$. The Greens' function vector $\mathbf{g}_k^{(n)}$ represents the transmission loss from target location $\mathbf{x}_k$ to the sensors in array n. Additive complex Gaussian noise is denoted by $\mathbf{n}_k^{(n)}$ which follows the distribution $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$, where $\mathbf{0}$ is a zero mean vector and $\mathbf{I}$ is the identity matrix of compatible dimension. Noises at different arrays are assumed to be independent and identically distributed. The source level and noise level are given by $\text{SL} = 10 \log \sigma_s^2$ and $\text{NL} = 10 \log \sigma_n^2$, respectively. The received data are generated on the frequency band from 40 Hz to 100 Hz with interval $\Delta f = 1$ Hz. For the featured frequencies of the target, the simulated frequency snapshots are generated containing the signal component, and for simplicity the level of the source signal of each featured frequency is assumed to be identical, i.e., $\sigma_s^2(f_m) = \sigma_s^2$. To get data that better fit the realistic shallow water environment, normal mode model KRAKEN is used to generate the transmission loss $\mathbf{g}_k^{(n)}$ at signal frequencies for all the arrays with environment parameters shown in Figure 1b. For the other frequency bins in the received bandwidth, noise-only data snapshots are generated. In processing, sliding window is used to pick the snapshots. One single frame k with $L = 11$ snapshots are used for beamforming at one time with five snapshots overlapping between the adjacent frames. The sample time interval between snapshots are set to be 2.5 s and thus the time interval in between frames equals to 15 s.

There are two kinds of criteria here to evaluate the filter performance. Optimal subpattern assignment (OSPA) is a measure that is used to analyze the estimation error of the target state and target number at the same time. The OSPA distance between two RFS $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{n_x}\}$ and $\mathcal{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_{n_y}\}$, $n_x, n_y \in \mathbb{N}_0$, is given by

$$\bar{d}_p^{(c)}(\mathcal{X}, \mathcal{Y}) = \left(\frac{1}{n_y} \left(\min_{\pi} \sum_{i=1}^{n_x} d^{(c)}(\mathbf{x}_i, \mathbf{y}_{\pi(i)})^p + c^p(n_y - n_x) \right) \right)^{1/p}, \tag{26}$$

if $n_x \leq n_y$, where $d^{(c)}(\mathbf{x}, \mathbf{y}) = \min(c, d(\mathbf{x}, \mathbf{y}))$ is the distance between $\mathbf{x}$ and $\mathbf{y}$, and $d(\mathbf{x}, \mathbf{y})$ is the traditional distance between vectors. If $n_y \leq n_x$, then $\bar{d}_p^{(c)}(\mathcal{X}, \mathcal{Y}) = \bar{d}_p^{(c)}(\mathcal{Y}, \mathcal{X})$. The summation is taken over all permutations π on $\{1, 2, \ldots, n_y\}$. $\bar{d}_p^{(c)}$ is the OSPA metric of order p with cut-off c. Here, we use the Euclidean distance $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$. The cut-off distance c represents the upper bound of the estimation error of the state vector.

Another metric used is to measure the probability of target lost in the tracking problem, called circular position error probability (CPEP), defined as

$$\mathrm{CPEP}(r) = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{X}} \Pr\left\{\|\mathbf{H}\hat{\mathbf{x}} - \mathbf{H}\mathbf{x}\| > r, \forall \hat{\mathbf{x}} \in \hat{\mathcal{X}}\right\}, \tag{27}$$

where matrix $\mathbf{H}$ is used to extract components of state vector for performance analysis, and r is the pre-defined position error radius, outside which the estimation is regarded as target missing.

Given the target track in Figure 1a, we set NL $= 70$ dB, and Bernoulli SMC filters are applied for SL $= 130$ dB and SL $= 120$ dB, respectively. Based on experience, the Bernoulli filter parameters are set to $N_p = 5000$, $N_b = 5000$, $p_s = 0.99$ and $p_b = 0.01$. When calculating the particle weights for DOA-based method, we set $\nu = 6$ and $u(\mathbf{z}) = 1/2\pi$ (uniform distribution of azimuth). Probability of detection is assumed to be constant $p_d = 0.2$. Probability of target existence, CPEP, and OSPA of both location and velocity tracking are compared for the DOA-based, beam-based and hybrid methods. To eliminate the variation caused by SMC implementation, results are averaged over 500 trails with the same filter parameters. Probability of target existence is the updated probability $q_{k|k}$ in the filter iteration. Target is declared when $q_{k|k} > 0.5$ and the state vector estimation is accepted. For $q_{k|k} <= 0.5$, target is not detected and the RFS state is an empty set. When calculating CPEP, we use $\mathbf{H} = \mathrm{diag}(1, 0, 1, 0)$ and $r = 2$ km. That is, when the positioning error exceeds 2 km, the target is regarded as lost. At last, we extract the position and velocity components of the state estimate to calculate the OSPA separately so that the location and the speed estimation error can be compared individually. The error cut-off is $c = 4$ km for location and $c = 0.1$ m/s for velocity.

Results are shown in Figures 2 and 3. Both beam-based and hybrid methods can better capture the appearance and disappearance of the target signal. The DOA-based method fails to detect the disappearance of the target after $k = 185$. When the target appears, it takes several frames for the beam-based and hybrid filters to detect the target. After the target is detected, the tracking precision of the beam-based and hybrid methods is better than that of the DOA-based method in the simulations. The target track after appearance can be roughly divided into three parts. In the first part, the target is close to HLA2 and is far away from the other two HLAs. As the target moving away from HLA1, the tracking performance degrades (OSPA increases and CPEP decreases). In the middle of the track where target is not close to any of the three arrays, the tracking performance is relatively poor. When the target comes into the area near HLA1 and HLA3, performance improvement can be observed. The performance change is more pronounced when SL $= 120$ dB. In Figure 3, a possible reason for the sudden change around $k = 150$ is that the target enters a relatively high SNR region of HLA3 (the relationship between performance and SNR is not linear), so the target location is corrected. However, while the target location is corrected, the filter actually detects a large change in the received signal, so that the target existence probability decreases in a short time. The sudden correction of the target position increases the estimation error of target velocity. These errors are corrected after a short period of time. When SL $= 130$ dB, similar but insignificant performance changes can be observed in the same period. Both beam-based and hybrid methods, where the amplitude information of beamforming is used, are more sensitive to degradation of signal quality. However, in most frames, both beam-based and hybrid methods still outperform the DOA-based method.
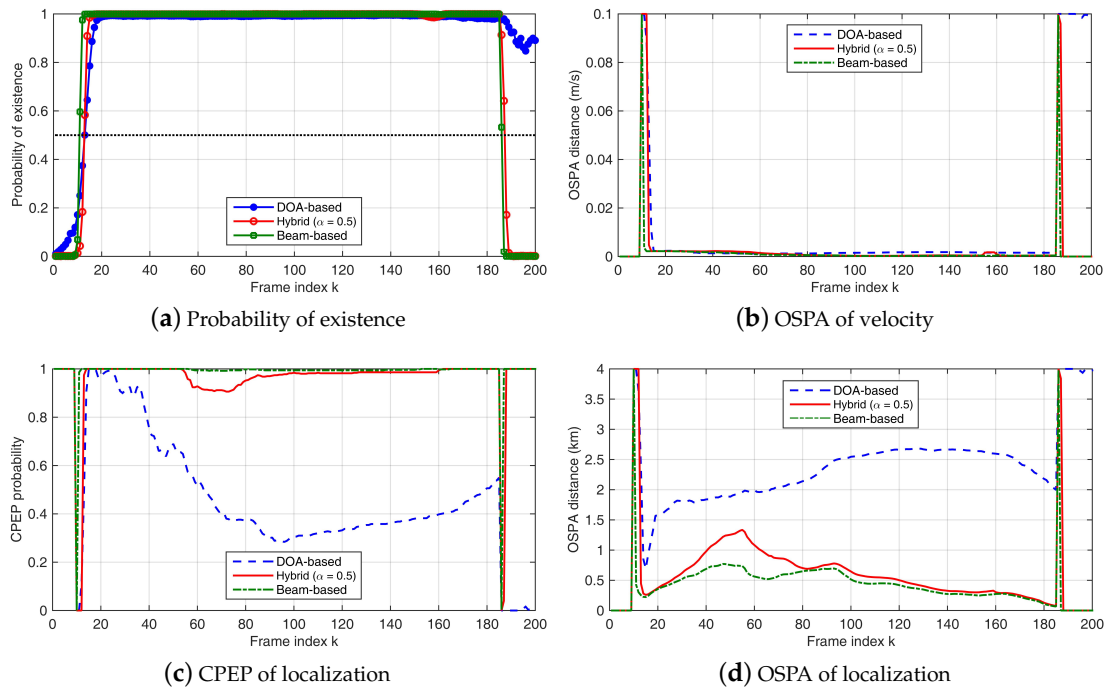
(**a**) Probability of existence

(**b**) OSPA of velocity

(**c**) CPEP of localization

(**d**) OSPA of localization

Figure 2. Simulation results with SL = 130 dB and NL = 70 dB: (**a**) probability of existence; (**b**) OSPA error (m/s) of target velocity tracking; (**c**) CPEP of target location tracking; and (**d**) OSPA error (m) of target location tracking. Results of the DOA-based Bernoulli filter are shown in blue lines, results of the beam-based method are presented by green lines, and results of the hybrid method with $\alpha = 0.5$ are shown in red lines.
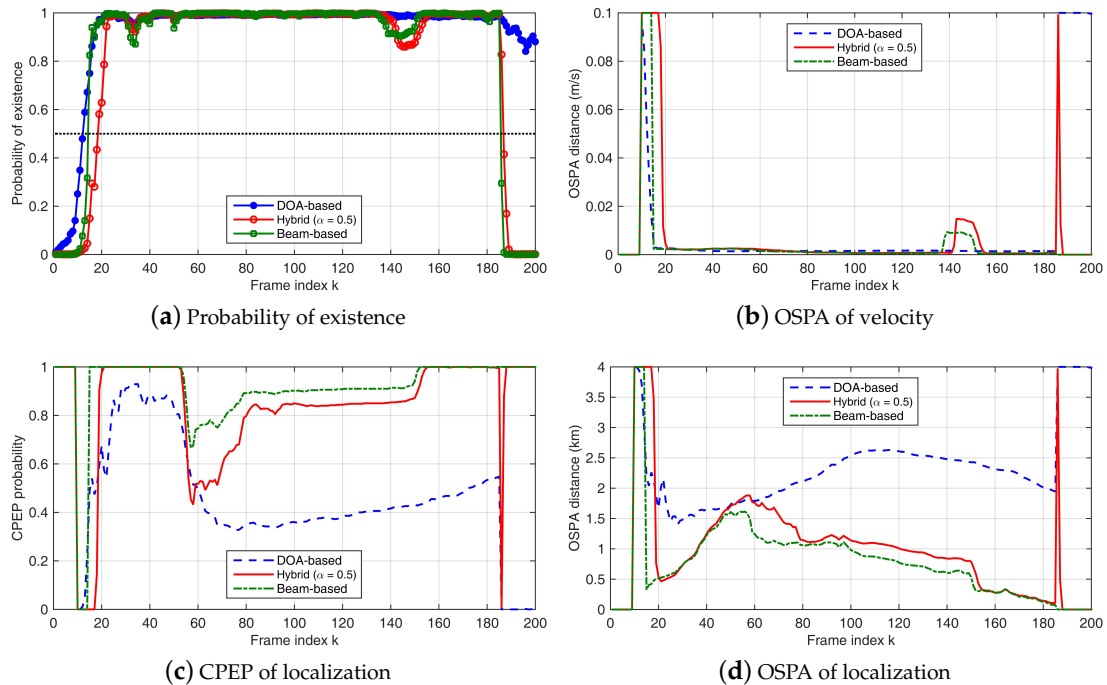

(**a**) Probability of existence

(**b**) OSPA of velocity

(**c**) CPEP of localization

(**d**) OSPA of localization

Figure 3. Simulation results when SL = 120 dB and NL = 70 dB: (**a**) probability of existence; (**b**) OSPA error (m/s) of target velocity tracking; (**c**) CPEP of target location tracking; and (**d**) OSPA error (m) of target location tracking. Results of the DOA-based Bernoulli filter are shown in blue lines, results of the beam-based method are presented by green lines, and results of the hybrid method with $\alpha = 0.5$ are shown in red lines.

5.2. Experiment Data Processing Results

The experiment dataset used to evaluate the performance was collected in the experiment SWellEx-96 Event S5 [30]. In this event, sound pressure was recorded by multiple arrays, a vertical line array (VLA), a tilted line array (TLA) and two HLAs. The experiment area along with the array locations and ship-recorded target trace are shown in Figure 4. The locations of arrays and target trace with recorded data (in the red box) in latitude and longitude are converted to Cartesian coordinates and a surveillance region of size 2 km × 8 km is formed. The corresponding latitude and longitude of the Cartesian origin are 32°37.25′ N and 117°21.82′ W. For the purpose of tracking the target in the XY-plane, only the data collected by 2 HLAs, referred to as HLA north (HLAn) and HLA south (HLAs), are used. Those arrays are deployed at the seabed with the water depth of 213 m (HLAn) and 198 m (HLAs), respectively. The number of effective array elements for processing is 27 (HLAn) and 28 (HLAs), respectively. Two multi-tone sources travel in the region at speed around 5 knots. The respective source depths are 54 m and 9 m, and their transmitted tone frequencies are listed in Table 1. Data from frequency band of 40–150 Hz are used to test our Bernoulli tracking algorithms. Array element locations as provided on the dataset website are used to calculate the beam pattern in advance.

Table 1. Frequencies of the Multi-tone sources.

| Deep Source (54 m) Frequencies (Hz)	
49 64 79 94 112 130 148 166 201 235 283 338 388	High Signal Level
52 67 82 97 115 133 151 169 204 238 286 341 391	2nd Set of Tonals
55 70 85 100 118 136 154 172 207 241 289 344 394	3rd Set of Tonals
58 73 88 103 121 139 157 175 210 244 292 347 397	4th Set of Tonals
61 76 91 106 124 142 160 178 213 247 295 350 400	5th Set of Tonals
Shallow Source (54 m) Frequencies (Hz)	
109 127 145 163 198 232 280 335 385	
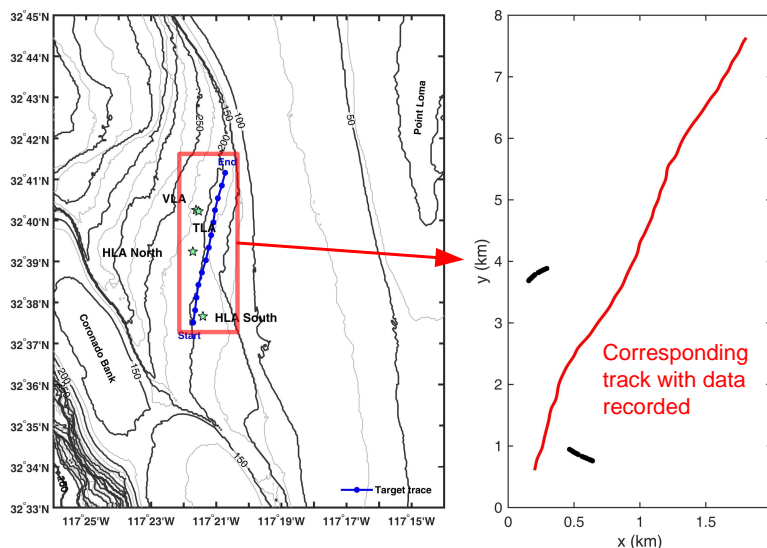
Figure 4. The experiment setup of SWellEx-96 Event S5 is shown in the left figure. The deployed arrays are represented by the stars and the ship-recorded target trace is the blue line. The area in the red rectangle is zoomed in with converted Cartesian coordinates. The data recorded by two HLAs are used for analysis and the HLAs are black lines in the right figure. The red line is the target track corresponding to the data for analysis.

The recorded sound pressure is sampled at 3276.8 Hz. Data flow is transformed to frequency-domain snapshots with 8192-point FFT. The snapshots length is thus 2.5 s (0.4 Hz frequency bin widths). Similar to the simulations in the previous section, in beamforming, one single frame k contains L = 11 snapshots with 5 snapshots overlapping between the adjacent frames. A total of 199 frames were processed. During the experiment, for frames 48∼54, 120∼123 and 131∼134, the deep source stopped projecting continuous wave (CW) tones and started projected chirps that are not on the processed frequency bands. Target is treated as absent during those frames for that only noise is contained in the data for processing.

In the experiment data processing, the Bernoulli filter parameters are set to $N_p = 10,000$, $N_b = 8000$, $p_s = 0.9$ and $p_b = 0.1$. In DOA-based method, $\nu = 6$ and $u(\mathbf{z}) = 1/2\pi$ (uniform distribution of azimuth). Probability of detection is $p_d = 0.9$. Probability of target existence, CPEP, OSPA of both location and velocity tracking are presented in Figure 5. Results are the average of 500 trails with the same filter parameters. Probability of target existence is the updated probability $q_{k|k}$ in the filter iteration. Target is declared when $q_{k|k} > 0.5$ and the state vector estimation is accepted. For $q_{k|k} <= 0.5$, target is not detected and the RFS state is an empty set. $\mathbf{H} = \text{diag}(1, 0, 1, 0)$ and location error radius $r = 1$ km in CPEP metric. Position and velocity OSPA cut-off are set to $c = 2$ km for location and $c = 0.1$ m/s for velocity.
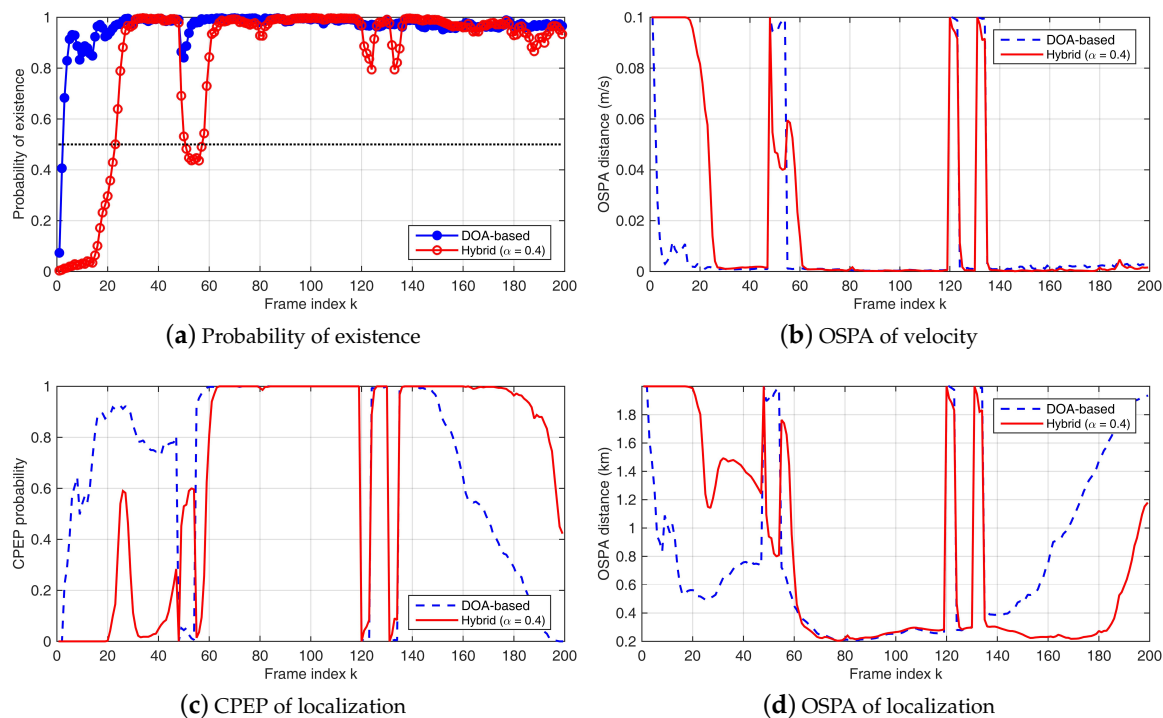

Figure 5. Processing results of data from 2 HLAs in SWellEx-96 Event S5: (**a**) probability of existence; (**b**) OSPA error (m/s) of target velocity tracking; (**c**) CPEP of target location tracking; and (**d**) OSPA error (m) of target location tracking. Results of the DOA-based Bernoulli filter are shown in blue lines, and results of the hybrid method with $\alpha = 0.4$ are shown in red lines.

Only two arrays in the experiment led to wrong estimation when the correlation deviated for any of the arrays. This deviation should have been eliminated if we had one more array. Besides, for many frames, the target is in the end-fire direction of HLAs, where the performance of the beam-based method is not satisfactory. In Figure 5, the performance results of the DOA-based method and the hybrid method with $\alpha = 0.4$ are presented. The performance of the beam-based method decreases dramatically, because of the lack of arrays and the effect of the end-fire direction, and is thus not shown. Similar to the simulation results, compared to the DOA-based algorithm, the hybrid method is better

at detection, i.e., capturing the appearance and disappearance of the target signal. This can be seen between frames 48∼54, 120∼123 and 131∼134 when the target signal is not included in the processed data. Correspondingly, when using the hybrid method, the convergence rate for the relatively more precise estimation is slower than that of the DOA-based method (for frames 1∼26 at the beginning). From frames 62∼118, both tracking algorithms achieve good performance. Near the end of the track (frames 135∼199), the hybrid method outperforms the DOA-based method.

6. Conclusions

We studied the passive joint detection and tracking problem of underwater targets using multiple arrays. The beamforming results obtained by the spatially distributed arrays were used as inputs to the Bernoulli filter to track the target presence, target position, and velocity. Three different likelihood calculation methods, the DOA-based, beam-based and hybrid methods are proposed for the Bernoulli filter. In the DOA-based Bernoulli filter, measurement sets were composed of DOA peak positions of beam power. The beam-based Bernoulli filter adopted the entire beam power as the input measurement set. The hybrid method was proposed to complement the advantages of the DOA-based and the beam-based method in the case of fewer arrays. We used a realistic underwater acoustic waveguide model to generate simulation data and evaluate the performance of the three methods. We also compared their performance using the low-frequency portion of the real data collected in the SWell-Ex96 experiment. The simulation results with three arrays showed that the proposed beam-based method and hybrid method can improve the tracking performance compared with the DOA-based filter. In the experimental data processing with two arrays, results also indicate that the proposed hybrid method has a better ability of detection, as well as a better tracking precision after convergence. It is worth noting that the RFS-based Bernoulli filter has great potential in multi-target tracking. We wish to extend the proposed methods to track multiple acoustic targets in future efforts.

Author Contributions: Z.C. proposed the improved method, carried out simulation and data analysis and wrote the manuscript. W.X. assisted with the integrity of the research and revised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Candy, J.V. *Bayesian Signal Processing: Classical, Modern, and Particle Filtering Methods*; Wiley InterScience: Hoboken, NJ, USA, 2009.
2. Kay, S.M. Fundamentals of statistical signal processing, Vol. II: Detection Theory. In *Signal Processing*; Prentice Hall: Upper Saddle River, NJ, USA, 1998.
3. Wald, A. Sequential tests of statistical hypothesis. *Ann. Math. Stat.* **1945**, *16*, 117–186. [CrossRef]
4. Kalman, R.E.; Bucy, R.S. New results in linear filtering and prediction theory. *J. Basic Eng.* **1961**, *83*, 109. [CrossRef]
5. Ljung, L. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Trans. Autom. Control* **2013**, *24*, 36–50. [CrossRef]
6. Julier, S.J.; Uhlmann, J.K. New extension of the Kalman filter to nonlinear systems. In Proceedings of the SPIE-The International Society for Optical Engineering, Orlando, FL, USA, 28 July 1997; Volume 3068, pp. 182–193.
7. Arulampalam, M.S.; Maskell, S.; Gordon, N.; Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 174–188. [CrossRef]
8. Djuric, P.M.; Kotecha, J.H.; Zhang, J.; Huang, Y.; Ghirmai, T.; Bugallo, M.F.; Miguez, J. Particle filtering. *IEEE Signal Process. Mag.* **2003**, *20*, 19–38. [CrossRef]

9.   Li, T.; Bolic, M.; Djuric, P.M. Resampling methods for particle filtering: Classification, implementation, and strategies. *IEEE Signal Process. Mag.* **2015**, *32*, 70–86. [CrossRef]

10.   Martino, L.; Elvira, V.; Camps-Valls, G. Group importance sampling for particle filtering and MCMC. *Digit. Signal Process.* **2018**, *82*, 133–151. [CrossRef]

11.   Kotecha, J.H.; Djuric, P.M. Gaussian sum particle filtering. *IEEE Trans. Signal Process.* **2003**, *51*, 2602–2612. [CrossRef]

12.   Tonissen, S.M.; Evans, R.J. Peformance of dynamic programming techniques for track-before-detect. *IEEE Trans. Aerosp. Electron. Syst.* **1996**, *32*, 1440–1451. [CrossRef]

13.   Bazzani, L.; Bloisi, D.; Murino, V. A comparison of multi hypothesis Kalman filter and particle filter for multi-target tracking. In Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Miami, FL, USA, 25 June 2009; pp. 47–54.

14.   Eryildirim, A.; Guldogan, M.B. A Bernoulli filter for extended target tracking using random matrices in a UWB sensor network. *IEEE Sens. J.* **2016**, *16*, 4362–4373. [CrossRef]

15.   Li, M.; Li, J.; Zhou, Y. Labeled RFS-based track-before-detect for multiple maneuvering targets in the infrared focal plane array. *Sensors* **2015**, *15*, 30839–30855. [CrossRef] [PubMed]

16.   Vo, B.T.; See, C.M.; Ma, N.; Ng, W.T. Multi-sensor joint detection and tracking with the Bernoulli filter. *IEEE Trans. Aerosp. Electron. Syst.* **2012**, *48*, 1385–1402. [CrossRef]

17.   Mahler, R. *Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2007.

18.   Ristic, B.; Vo, B.T.; Vo, B.N.; Farina, A. A tutorial on Bernoulli filters: Theory, implementation and applications. *IEEE Trans. Signal Process.* **2013**, *61*, 3406–3430. [CrossRef]

19.   Zhang, Q.; Song, T.L. Improved bearings-only multi-target tracking with GM-PHD filtering. *Sensors* **2016**, *16*, 1469. [CrossRef] [PubMed]

20.   Wong, S.J.; Ba, T.V. Square root Gaussian mixture PHD filter for multi-target bearings only tracking. In Proceedings of the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Adelaide, Australia, 6–9 December 2011; pp. 520–525.

21.   Chen, J.; Ma, H.; Liang, C.; Zhang, Y. OTHR multipath tracking using the bernoulli filter. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 1974–1990. [CrossRef]

22.   Evers, C.; Moore, A.H.; Naylor, P.A.; Sheaffer, J. Bearing-only acoustic tracking of moving speakers for robot audition. In Proceedings of the IEEE International Conference on Digital Signal Processing, Singapore, 21–24 July 2015; pp. 1206–1210.

23.   Braca, P.; Goldhahn, R.; Ferri, G.; Lepage, K.D. Distributed information fusion in multistatic sensor networks for underwater surveillance. *IEEE Sens. J.* **2016**, *16*, 4003–4014. [CrossRef]

24.   Gunes, A.; Guldogan, M.B. Joint underwater target detection and tracking with the Bernoulli filter using an acoustic vector sensor. *Digit. Signal Process.* **2016**, *48*, 246–258. [CrossRef]

25.   Yuan, C.; Wang, J.; Peng, L.; Bi, Y.; Sun, Z. Multi-target tracking based on multi-Bernoulli filter with amplitude for unknown clutter rate. *Sensors* **2015**, *15*, 30385–30402. [CrossRef] [PubMed]

26.   Xu, W.; Li, J. Study of statistical signal models in low-frequency underwater acoustic applications. In Proceedings of the Oceans11 MTS/IEEE Kona, Waikoloa, HI, USA, 19–22 September 2011; Volume 8, pp. 1–5.

27.   Wu, Q.; Xu, W. Matched field source localization as a multiple hypothesis tracking problem. In Proceedings of the ACM International Conference on Underwater Networks & Systems, Halifax, NS, Canada, 6–8 November 2017; p. 25.

28.   Mahler, R. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1152–1178. [CrossRef]

29.   Ferri, G.; Munafo, A.; Goldhahn, R.; Lepage, K. A non-myopic, receding horizon control strategy for an AUV to track an underwater target in a bistatic sonar scenario. In Proceedings of the 53rd IEEE Conference on Decision and Control, Los Angeles, CA, USA, 15–17 December 2014; pp. 5352–5358.

30.   Murray, J.; Ensberg, D. The SWellEx-96 Experiment. 1996. Available online: http://swellex96.ucsd.edu (accessed on 9 November 2018).