Broadband Passive Sonar Track-Before-Detect Using Raw Acoustic Data

Daniel Bossér (Student Member IEEE), Magnus Lundberg Nordenvaad, Gustaf Hendeby (Senior Member IEEE), Isaac Skog (Senior Member IEEE)

Abstract

This article concerns the challenge of reliable broadband passive sonar target detection and tracking in complex acoustic environments. Addressing this challenge is becoming increasingly crucial for safeguarding underwater infrastructure, monitoring marine life, and providing defense during seabed warfare. To that end, a solution is proposed based on a vector-autoregressive model for the ambient noise and a heavy-tailed statistical model for the distribution of the raw hydrophone data. These models are integrated into a Bernoulli track-before-detect (TkBD) filter that estimates the probability of target existence, target bearing, and signal-to-noise ratio (SNR). The proposed solution is evaluated on both simulated and real-world data, demonstrating the effectiveness of the proposed ambient noise modeling and the statistical model for the raw hydrophone data samples to obtain early target detection and robust target tracking. The simulations show that the SNR at which the target can be detected is reduced by $4\,\mathrm{dB}$ compared to when using the standard constant false alarm rate detector-based tracker. Further, the test with real-world data shows that the proposed solution increases the target detection distance from $250\,\mathrm{m}$ to $390\,\mathrm{m}$. The presented results illustrate that the TkBD technology, in combination with data-driven ambient noise modeling and heavy-tailed statistical signal models, can enable reliable broadband passive sonar target detection and tracking in complex acoustic environments and lower the SNR required to detect and track targets.

This work has been founded by the research organization Zenith (project id. 20.03) and SecurityLink.

Daniel Bossér and Gustaf Hendeby are with Dept. of Electrical Engineering, Linköping University, (e-mail: daniel.bosser@liu.se; gustaf.hendeby@liu.se).

Magnus Lundberg Nordenvaad is with the Div. of Underwater Technology, Swedish Defence Research Agency (FOI), Kista, Sweden (e-mail: magnus.lundberg@foi.se)

Isaac Skog is with Dept. of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, and the Div. of Underwater Technology, Swedish Defence Research Agency (FOI), Kista, Sweden (e-mail: skog@kth.se)

Manuscript received Month DD, Year; revised Month DD, Year.

Index Terms

Underwater passive surveillance, target tracking, array signal processing

I. Introduction

A significant portion of today's critical infrastructure is located underwater. This includes gas pipelines, power transmission lines, and communication cables, which are essential in modern society. Due to their remote locations and strategic importance, these assets are vulnerable to damage and sabotage. Monitoring these infrastructures is crucial, especially during times of conflict, when targeted attacks on them could have severe consequences [1]. This has led to the emergence of a new domain of conflict on the world's seafloors, known as seabed warfare, which necessitates the development of advanced countermeasures [2], [3].

Passive sonar surveillance is essential for discreetly monitoring underwater infrastructure. This technique enables the detection of submarines and other underwater vehicles without disclosing the location of the sonar system. Moreover, passive sonar does not introduce noise pollution in the ocean, benefiting marine life [4], [5], while allowing for continuous, unobtrusive monitoring of the underwater environment. However, compared to active sonar, passive sonar systems typically operate at a lower signal-to-noise ratio (SNR) and are more susceptible to ambient noise. This requires them to have higher sensitivity and use more complex noise models.

Passive surveillance has historically relied on a combination of signal processing techniques such as low frequency analysis and recording (LOFAR), beamforming, and bearing time record (BTR) analysis. These methods are often manually operated by human sonar operators, who may also listen to the sounds. However, relying on human expertise is costly and resource-limited, making large-scale monitoring challenging. Energy-based detectors, such as the constant false alarm rate (CFAR) detector, have been employed to automate the surveillance process. These detectors output a set of bearings at each time step, corresponding to potential target detections. When a target is present, some detections may over time form a track, which can be identified using target tracking methods such as multiple hypothesis tracking [6] or more recent approaches such as the Poisson multi-Bernoulli mixture (PMBM) filter [7]. Applications of these methods to underwater surveillance have been explored in [8], [9]. However, a prerequisite for successful detection is that the SNR is sufficient to exceed the detection threshold.

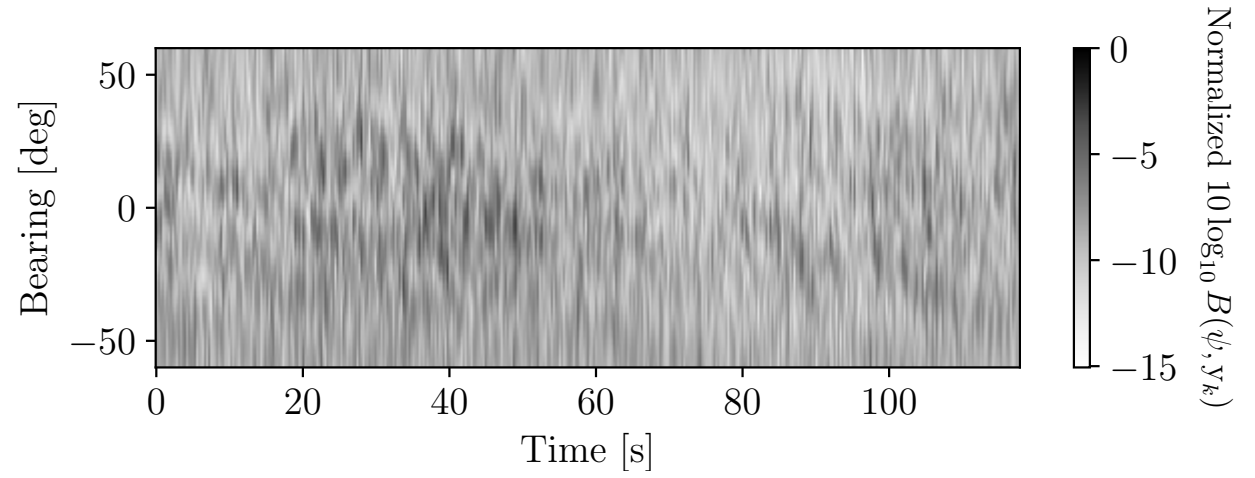
One way to increase the performance in poor SNR is by using the track-before-detect (TkBD) tracking strategy [10, p. 239]. In this approach, the target detection occurs much later in the

signal processing chain, after constructing the potential track, hence the name. Thus, target detection and tracking are done jointly. The major benefit of this method is that no information is discarded in the detection process, allowing for a longer integration time of the raw data before the decision is made, consequently lowering the SNR requirement. Theoretically, a performance gain of approximately $6\,\mathrm{dB}$ is possible [11, p. 318]. However, TkBD makes the tracker more sensitive to signal and noise modeling errors, which is why the application of TkBD to the underwater passive sonar problem has been so challenging. Consider the BTR in Fig. 1a, which shows the received signal energy in different bearings over time. From Fig. 1a, it is evident that the signal energy varies greatly over time and bearing, implying that the received signal is spatiotemporally correlated. These correlations are difficult to discern from the signal components produced by a target, as seen in Fig. 1b. Another challenge is that ambient noise is known to exhibit a heavy-tailed distribution [12, p. 403]. Previous work has attempted to model the acoustic samples using alpha-stable distributions [13] or Gaussian mixture models [14] to address this issue, but this has not been done in the context of target tracking.

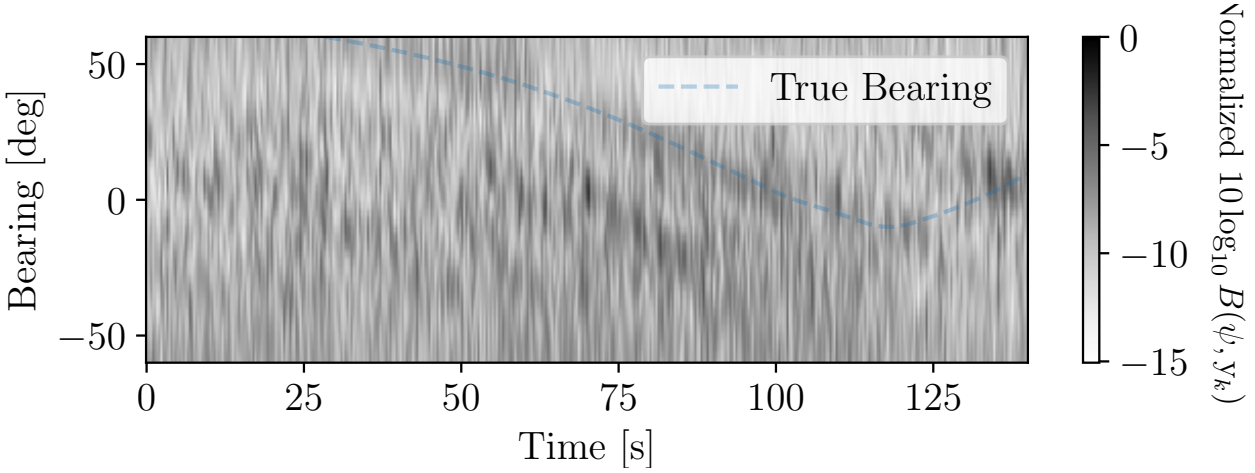
Different approaches utilizing TkBD for passive sonar are explored in [15]–[18]. Many of these methods rely on modeling signal energy post-beamforming as their measurements and only consider narrowband signals. Some approaches have attempted to address the challenges posed by spatiotemporal correlations. For instance, the work in [17] averages the signal energy post-beamforming in different bearing bins, thereby reducing the impact of the temporal energy variations. Similarly, the post-beamforming energy is also used in [18]. By comparing the beamforming energy in the presence and absence of a target, the authors of [18] fit a probability distribution to each bearing bin. Although this approach adapts to spatial variations in ambient noise, the fitted distributions are time-invariant, meaning that temporal energy variations are not accounted for. Likewise, the work in [19] and [20] applies models that assume noise interferers at specific bearings that emit independent, constant power signals. None of these models handle the temporal variation in the ambient noise.

A. Contributions

This article builds upon and extends the work presented in [20], [21], which explored the possibility of sample-level source and ambient sound modeling in a passive sonar TkBD application, circumventing the challenges of developing accurate statistical models for the signal after beamforming. In [20], it was observed that spatiotemporal correlations and other modeling



(a) No target



(b) With target

Fig. 1. BTRs from data collected during a sea trial. The BTR values were calculated using the conventional delay-and-sum beamformer $B(\psi, \mathrm{y}_k)$ in (11). The upper and lower plots show the BTR without and with a target present. From the plots, it is clear that the temporal and spatially correlated ambient noise causes disturbances that are hard to distinguish from the actual target.

errors make the tracker prone to initiating false tracks. This article proposes a solution to this challenge by adding a new data preprocessing step and a new measurement model for the data samples. In summary, the contributions of this work are:

1) a vector auto-regressive (VAR) model for modeling of the spatially and temporally correlated ambient noise;
2) a measurement model for wideband signals with a heavy-tailed distribution; and
3) an experimental evaluation of the proposed models within a bearing-only TkBD framework.

Reproducible research: The code and data used to reproduce the presented results can be

downloaded at `https://gitlab.liu.se/coast/tkbd_using_raw_data`.

II. Target Tracking and Bayesian Estimation

The objective of target detection and tracking is to, given a set of measurements $z_{1:k} = \{z_1, z_2, \ldots, z_k\}$ collected at discrete time instances $1, 2, \ldots, k$, determine if a target is present and, if so, estimate its current state. Commonly, the measurements $z_{1:k}$ are a set of detections, but in TkBD applications, they may instead be intensity measurements, images, beamforming outputs, or, as this article proposes, a set of hydrophone samples.

To simultaneously describe the target state and the probability of its existence, the target is modeled using a Bernoulli random finite set (BRFS). To that end, let q_k denote the probability of existence, and x_k denote the state of the target, given that the target exists. The BRFS X_k jointly describes these components with the finite set statistics (FISST) probability density function (PDF) [22]

$$f_k(X) = \begin{cases} 1 - q_k & \text{if } X = \emptyset, \\ q_k s_k(x) & \text{if } X = \{x\}, \end{cases} \tag{1}$$

where $s_k(x)$ is the PDF of state x_k. The posterior PDF

$$f_{k|k'}(X) \equiv f_k(X|z_{1:k'}) = \begin{cases} 1 - q_{k|k'} & \text{if } X = \emptyset, \\ q_{k|k'} s_{k|k'}(x) & \text{if } X = \{x\}, \end{cases} \tag{2}$$

can be calculated using the Bernoulli filter [23]. Here, $q_{k|k'}$ and $s_{k|k'}(x)$ denote the posterior probability of existence and posterior state distribution at time instant k given measurements up to time instant k'. Given the posterior PDF $f_{k-1|k-1}(X)$, the Bernoulli filter recursions for calculating the posterior distribution $f_{k|k}(X)$ are given by the time and measurement updates [23]

$$q_{k|k-1} = p_\mathrm{b}(1 - q_{k-1|k-1}) + p_\mathrm{s} q_{k-1|k-1}, \tag{3a}$$

$$\begin{aligned} s_{k|k-1}(x) = {} & \frac{p_\mathrm{b}(1 - q_{k-1|k-1} b_{k|k-1}(x))}{q_{k|k-1}} \\ & + \frac{p_\mathrm{s} q_{k-1|k-1} \int \pi_{k|k-1}(x|x') s_{k-1|k-1}(x')\, dx'}{q_{k|k-1}}, \end{aligned} \tag{3b}$$

and

$$q_{k|k} = \frac{q_{k|k-1} \int L(z_k|x) s_{k|k-1}(x)\, dx}{1 - q_{k|k-1} + q_{k|k-1} \int L(z_k|x) s_{k|k-1}(x)\, dx}, \tag{4a}$$

$$s_{k|k}(x) = \frac{L(z_k|x) s_{k|k-1}(x)}{\int L(z_k|x) s_{k|k-1}(x)\, dx}, \tag{4b}$$

respectively. Here, p_{b} and p_{s} are the probability of target birth and survival between time steps, respectively. Further, $b_{k|k-1}(x)$ denotes the PDF for states of newly born targets. Moreover, $\pi_{k|k-1}(x|x')$ is a PDF of the state x conditioned on the state x', describing the target motion. Finally, $L(z_k|x)$ is the conditional likelihood ratio given the target state x. This article uses the particle filter to implement the recursions in (3) and (4). Pseudocode describing the particle filter implementation can be found in [23].

To execute the filter recursion, one must specify models for the target dynamics, target birth and survival probability, and measurement likelihood. This article focuses on how the measurement likelihood should be modeled in a broadband passive sonar system. Next, the pros and cons of the commonly used CFAR likelihood model will be reviewed, and a new likelihood model that addresses some of the cons of existing measurement models will be proposed.

III. Measurement Models

The section describes how the relationship between the measurements z_k and the target state x_k, and the associated likelihood ratio function $L(z_k|x_k)$, is modeled. First, the traditional case when z_k consists of detections is described. Second, a new measurement model for when z_k consists of raw hydrophone samples is presented.

A. Spatial Signal Processing and Beamforming

Consider a hydrophone array consisting of M hydrophones. Let $y_n^{(m)}$ denote sample n from hydrophone m in the array. Define

$$\vec{y}_n = \begin{bmatrix} y_n^{(0)} & \dots & y_n^{(M-1)} \end{bmatrix}^\top \in \mathbb{R}^M, \tag{5}$$

as the collection of samples from all the M hydrophones. Between the two time steps $k-1$ and k, N such samples are collected, denoted

$$\mathrm{y}_k = \begin{bmatrix} \vec{y}_{(k-1)N+1}^\top & \dots & \vec{y}_{kN}^\top \end{bmatrix}^\top \in \mathbb{R}^{NM}, \tag{6}$$

and is referred to as a batch of samples.

If the target is in the far field and emits a broadband signal $\mathrm{s}_k \in \mathbb{R}^N$, the batch of hydrophone samples y_k can be modeled as

$$\mathrm{y}_k = \mathrm{H}(\psi)\mathrm{s}_k + \mathrm{e}_k, \tag{7}$$

where e_k is given by

$$\mathrm{e}_k = \begin{bmatrix} \vec{e}_{(k-1)N+1}^{\top} & \dots & \vec{e}_{kN}^{\top} \end{bmatrix}^{\top} \in \mathbb{R}^{NM}, \tag{8a}$$

$$\vec{e}_n = \begin{bmatrix} e_n^{(0)} & \dots & e_n^{(M-1)} \end{bmatrix}^{\top} \in \mathbb{R}^{M}, \tag{8b}$$

and $e_n^{(m)}$ is the measurement noise of hydrophone m at time instant n. Furthermore, $\mathrm{H}(\psi)$ is a fractional delay filter matrix, given by [24]

$$\mathrm{H}(\psi) = \begin{bmatrix} H_0^{\top}(\psi) & \dots & H_{M-1}^{\top}(\psi) \end{bmatrix}^{\top} \in \mathbb{R}^{NM \times N}, \tag{9}$$

where

$$H_m(\psi) = W^* \Lambda(\tau^{(m)}(\psi))W, \tag{10a}$$

$$\Lambda(\tau) = \mathrm{diag}(\gamma^0(\tau), \dots, \gamma^{N-1}(\tau)), \tag{10b}$$

$$\gamma^n(\tau) = \begin{cases} \exp(-2\pi i n \tau f_s/N) & \text{if } n < \frac{N}{2}, \\ \cos(\tau \pi f_s) & \text{if } n = \frac{N}{2}, \\ \exp(2\pi i (N-n)\tau f_s/N) & \text{if } n > \frac{N}{2}. \end{cases} \tag{10c}$$

Here, * denotes the conjugate transpose operator. Further, $\tau^{(m)}(\psi)$ is the time shift of the signal at hydrophone m caused by the direction of arrival (DOA) ψ. Moreover, W is the unitary discrete Fourier transform matrix, and f_s is the sampling frequency of the hydrophone system.

The delay-and-sum beamformer reverses the delays in the signals, then sums the signals, and finally calculates the energy of the summed signals. That is, the beamformer output $B(\psi, \mathrm{y}_k)$ for DOA angle ψ and hydrophone sample batch y_k is given by

$$B(\psi, \mathrm{y}_k) = \|\mathrm{H}^{\top}(\psi)\mathrm{y}_k\|_2^2. \tag{11}$$

B. Detection-Based Measurement Model

Traditional target detection methods take target detections and associated bearings as inputs. In passive sonar, detections are obtained by applying a peak detector, typically a CFAR detector, to the beamformer output $B(\psi, \mathrm{y}_k)$. The resulting measurements are a set of bearings,

$$z_k \equiv \{\psi_0, \dots, \psi_D\}, \tag{12}$$

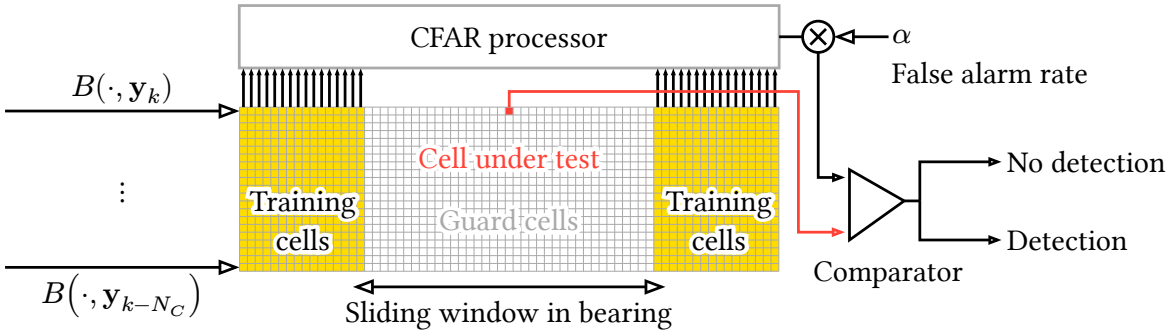


Fig. 2. Illustration of the CFAR detector. The CFAR processor evaluates a cell under test to determine whether it contains a detection or not. This is done by comparing the energy in the cell under test with the energy in training cells.

where the number of detections D will vary. The cell-averaging CFAR detector evaluates $B(\psi, \mathrm{y}_k)$ over a discrete set of angles ψ. Each bearing-value pair $(\psi, B(\psi, \mathrm{y}_k))$, known as the cell under test, is a detection candidate. To account for the ambient noise variations, a Gaussian distribution is fitted to the energy in neighboring cells, including those from previous time steps. A significance threshold calculated from the fitted Gaussian distribution is then used to classify whether the cell under test deviates from the training cells. If so, the bearing of the test cell is added to the set of detections in z_k. This process is illustrated in Fig. 2. More details on the CFAR detector can be found in [25]. Several detections may be directly adjacent to each other due to the beamwidth of the beamformer. Such detections are clustered into a single detection at their mean using the DBScan algorithm [26]. For this application, the algorithm is configured with half a bearing bin width as the neighborhood radius and a minimum of one detection per cluster.

Among the true detections, there will be false detections. The false detections are assumed to follow a Poisson point process, implying that the likelihood of observing z_k when no target is present is given by

$$\ell_0(z) = e^{-\lambda} \prod_{\psi \in z} \lambda \kappa(\psi). \tag{13}$$

Here, λ is the Poisson point process intensity, and $\kappa(\psi)$ is the PDF of each false detection. The false detections are assumed to be independent, identically, and uniformly distributed over the beamforming interval.

The likelihood of z_k given that a target with state x exists is derived in [23, Sec. V] and is given by

$$\ell_1(z|x) = \ell_0(z)(1-p_{\mathrm{d}}) + p_{\mathrm{d}} \sum_{\psi_d \in z} g(\psi_d|x)\ell_0(z \setminus \psi_d), \tag{14}$$

where $(\cdot \setminus \cdot)$ denote the set difference, p_{d} is the probability that the target generates a detection, and $g(\psi_d|x)$ is the likelihood function of detection ψ_d due to the target. The measurement uncertainty is assumed to follow a Gaussian distribution with variance R. That is,

$$g(\psi_d|x) = \mathcal{N}(\psi_d; \psi, R), \tag{15}$$

where ψ is the bearing of the target. In summary, the log-likelihood ratio of the detection measurements is given by

$$\begin{aligned} \ln L(z|x) &= \ln \ell_1(z|x) - \ln \ell_0(z) \\ &= \ln \left(1 - p_{\mathrm{d}} + p_{\mathrm{d}} \sum_{\psi \in z} \mathcal{N}(\psi_d; \psi, R) \frac{\ell_0(z \setminus \psi)}{\ell_0(z)} \right). \end{aligned} \tag{16}$$

A CFAR-based tracker is then constructed by using (16) as the likelihood ratio in the filter recursions in (3) and (4).

A CFAR-based tracker is computationally efficient as it compresses samples into a set of detections. Since a significant amount of data is discarded in the detection process, a high SNR is required for the tracker to function well. Furthermore, the training cells may not accurately represent the ambient noise, leading to a degraded performance. As will be shown next, this issue may be addressed by defining the measurement model in terms of the raw hydrophone samples. The drawback is an increased computational complexity.

C. Proposed TkBD Measurement Model

Recall that underwater noise is known to exhibit heavy-tailed distribution in its samples, see e.g., [12, p. 403]. Thus, the proposed TkBD measurement model assumes that each batch of hydrophone samples follows a multivariate-t distribution

$$\mathrm{y}_k \sim t_{NM}\left(\nu, 0, \frac{\nu-2}{\nu}\Sigma(\psi_k)\right) \quad \nu > 2, \tag{17}$$

where

$$\Sigma(\psi) = \mathrm{H}(\psi)\Sigma_{ss}\mathrm{H}^{\top}(\psi) + \Sigma_{ee}, \tag{18}$$

$\Sigma_{ss} = \mathrm{cov}(\mathrm{s}_k)$, and $\Sigma_{ee} = \mathrm{cov}(\mathrm{e}_k)$. Further, the PDF of the multivariate-t distribution is given by

$$t_{NM}(\mathrm{y};\nu,\mu,S) = \frac{\Gamma([\nu+NM]/2)}{\Gamma(\nu/2)(\nu\pi)^{NM/2}|S|^{1/2}} \cdot \left(1+\frac{1}{\nu}((\mathrm{y}-\mu)^\top S^{-1}(\mathrm{y}-\mu))\right)^{-\frac{\nu+NM}{2}}, \quad (19)$$

where Γ is the gamma function, ν is the degrees of freedom, μ is the mean, and S is the scale matrix. For $\nu > 2$ then $\mathrm{cov}(\mathrm{y}) = \frac{\nu}{\nu-2}S$. Smaller values of ν imply heavier tails in the distribution. Noteworthy is that when $\nu \to \infty$, the frequently used Gaussian distribution is obtained.

To evaluate the PDF in (19), the determinant and the inverse of $\Sigma(\psi_k)$ must be computed, which is computationally expensive due to its size $NM \times NM$. This computational complexity may be unmanageable for particle filter-based trackers that evaluate the measurement likelihood many times at each step of the filter recursions. However, prior work [21] demonstrates that if the signal s_k and noise e_k are temporally and spatially white, i.e., $\Sigma_{ss} = \sigma_s^2 I_N$ and $\Sigma_{ee} = \sigma_e^2 I_{NM}$, so that $\Sigma(\psi) = \sigma_s^2 \mathrm{H}(\psi)\mathrm{H}^\top(\psi) + \sigma_e^2 I_{NM}$, then

$$\log|\Sigma(\psi)| = N\log(M\sigma_s^2+\sigma_e^2) + N(M-1)\log(\sigma_e^2), \quad (20a)$$

and

$$\mathrm{y}^\top\Sigma^{-1}(\psi)\mathrm{y} = \frac{\|\mathrm{y}\|_2^2}{\sigma_e^2} - \frac{\sigma_s^2}{\sigma_e^2}\frac{B(\psi,\mathrm{y})}{\sigma_e^2+M\sigma_s^2}. \quad (20b)$$

Hence, under these assumptions, the log of the PDF in (19) can be efficiently calculated from the beamformer $B(\psi,\mathrm{y})$ and the signal energy $\|\mathrm{y}\|^2$. Note that the use of the delay-and-sum beamformer follows from the measurement model rather than a discretionary choice.

In reality, the ambient noise e_k is seldom white, i.e., $\Sigma_{ee} \neq \sigma_e^2 I_{NM}$. However, if covariance of the ambient noise Σ_{ee} is known the batch data y_k may be whitened as

$$\tilde{\mathrm{y}}_k = \Sigma_{ee}^{-\top/2}\mathrm{y}_k \sim t_{NM}(\nu, 0, \bar{\Sigma}_{ss}(\psi) + I_{NM}), \quad (21)$$

where

$$\bar{\Sigma}_{ss}(\psi) = \Sigma_{ee}^{-\top/2}\mathrm{H}(\psi)\Sigma_{ss}\mathrm{H}^\top(\psi)\Sigma_{ee}^{-1/2}, \quad (22)$$

Next, if one assumes that the source signal s_k is white so that $\Sigma_{ss} = \sigma_s^2 I_N$ and neglecting the effect (except for scaling) of the whitening on the source signal, then

$$\bar{\Sigma}_{ss}(\psi) \approx \eta\mathrm{H}(\psi)\mathrm{H}^\top(\psi). \quad (23)$$

Here, η is the power of the signal originating from the target relative to the ambient noise power after whitening, i.e., η is the SNR after whitening. The motivation for making the ad-hoc approximation in (23) is that it enables the PDF of $\tilde{\mathrm{y}}_k$ to be calculated efficiently via the beamformer. The argument in support of the approximation is that the distortion caused to the signal s_k by the whitening is likely negligible compared to other modeling errors, such that the assumption of s_k being a white signal; this assumption will be further discussed in Sec. V.

In summary, the proposed measurement model uses the batch $\tilde{\mathrm{y}}_k$ of whitened hydrophone samples as inputs, i.e.,

$$z_k \equiv \tilde{\mathrm{y}}_k. \tag{24}$$

The likelihood of z given that target with state x is modeled as

$$\ell_1(z|x) = t_{NM}\left(z; \nu, 0, \frac{\nu-2}{\nu}(\bar{\Sigma}_{ss}(\psi) + I_{NM})\right). \tag{25}$$

When no target is present, i.e., $\eta = 0$, the likelihood simplifies to

$$\ell_0(z) = t_{NM}\left(z; \nu, 0, \frac{\nu-2}{\nu} I_{NM}\right). \tag{26}$$

Thus, the log-likelihood ratio can efficiently be calculated from the beamformer as

$$\begin{aligned}\ln L(z|x) = &-\frac{N}{2}\ln\left(M\eta + 1\right) \\ &-\frac{\nu + NM}{2}\ln(1 - cB(\psi, z)),\end{aligned} \tag{27}$$

where

$$c = \frac{\eta}{(\nu + \|z\|^2)(1 + M\eta)}. \tag{28}$$

For a derivation, see Appendix.

D. Data-efficient Learning of Σ_{ee} using VAR Models

Typically, the covariance Σ_{ee} of the ambient noise is unknown and must be learned from historical data. Directly estimating Σ_{ee} using the sample covariance matrix requires a substantial amount of data. Instead, it is proposed that the correlation structure of the ambient noise is modeled using a VAR model. This allows for more data-efficient learning of Σ_{ee} and, as will be shown, the whitening to be done without factorizing and inverting Σ_{ee}. VAR models have previously been successfully applied in other sonar applications to capture the spatial and temporal characteristics of underwater sounds [27], though they are more commonly used in economic modeling to capture relationships between variables over time [28].

A ϱ:th order VAR model describes the ambient noise $\vec{e}_n$ as

$$\vec{e}_n = A_1\vec{e}_{n-1} + \cdots + A_\varrho\vec{e}_{n-\varrho} + \Sigma_w^{1/2}\vec{w}_n, \tag{29}$$

where $\vec{w}_n$ is white noise with $\mathrm{cov}(\vec{w}_n) = I_M$. The matrices A_1, ..., A_ϱ, and Σ_w are model parameters that defines the structure of Σ_{ee}. However, instead of directly constructing Σ_{ee} from the model parameters and then factorizing and inverting the matrix to do the whitening in (21), the fact that the VAR model is invertible can be used. That is, the white noise $\vec{w}_n$ can be retrieved from $\vec{e}_n, \ldots, \vec{e}_{n-\varrho}$ as follows

$$\vec{w}_n = \Sigma_w^{-1/2}(\vec{e}_n - A_1\vec{e}_{n-1} - \cdots - A_\varrho\vec{e}_{n-\varrho}). \tag{30}$$

Hence, the whitening of y_k in (21) can be efficiently implemented using the inverted VAR model by substituting $\vec{e}_n$ in (30) with the samples $\vec{y}_n$ from the batch y_k.

The model parameters $A_1, \ldots, A_\varrho, \Sigma_w$ can be learned using linear least squares [28, Ch. 3]. That is, the parameters are calculated as

$$\{\hat{A}_1, \ldots, \hat{A}_\varrho\} = \underset{A_1,\ldots,A_\varrho}{\arg\min} \sum_{n=\varrho+1}^{N_t} \epsilon_n^\top \epsilon_n \tag{31a}$$

where

$$\epsilon_n = \vec{y}_n - A_1\vec{y}_{n-1} - \cdots - A_\varrho\vec{y}_{n-\varrho}, \tag{31b}$$

and N_t denotes the number of samples in the dataset used to estimate the parameters. Further, the covariance Σ_w is estimated as

$$\hat{\Sigma}_w = \frac{1}{N_t - \varrho - 1} \sum_{n=\varrho+1}^{N_t} \epsilon_n\epsilon_n^\top. \tag{32}$$

The model order ϱ can be selected using, e.g., the Akaike information criterion [29, p.221]. An alternative method for learning the parameters that take into account the heavy-tailed distribution of the data can be found in [30].

IV. Target Dynamics Model

Recall from Sec. II that the Bernoulli filter recursions require a model of the target state dynamics $\pi(x|x')$ and a birth model $b_{k|k-1}(x)$. These models are defined next.

A. Motion Model

Given a multi-array setup, it would be possible to track the target in a cartesian state space. For simplicity of analysis, a bearing-only target detection and tracking setup with only one hydrophone array is considered. To that end, let the target state at time step k be

$$x_k = \begin{bmatrix} \psi_k & \dot{\psi}_k & \eta_k^{(\mathrm{dB})} \end{bmatrix}^\top, \tag{33}$$

where $\dot{\psi}_k$ is the bearing change rate and $\eta_k^{(\mathrm{dB})}$ is the SNR η_k in dB, i.e., $\eta_k^{(\mathrm{dB})} = 10\log_{10}\eta_k$. Changes in target bearing ψ are modeled according to a constant velocity model, and the SNR $\eta^{(\mathrm{dB})}$ of the target is modeled as a random walk. Details about these models and other motion models commonly used in target tracking can be found in [31]. Hence, conditioned on target state x_k, the PDF of x_{k+1} is modeled as

$$\pi(x_{k+1}|x_k) = \mathcal{N}(x_{k+1}; Fx_k, GQG^\top), \tag{34}$$

where

$$F = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T \end{bmatrix} \quad Q = \begin{bmatrix} q_{\mathrm{CV}}^2 & 0 \\ 0 & q_{\mathrm{dBSNR}}^2 \end{bmatrix}.$$

Here T is the time between instant k and $k+1$. Further, q_{CV}^2 and q_{dBSNR}^2 are the process noise variances for the constant velocity model and the random walk model, respectively.

B. Birth Model

The birth model $b_{k|k-1}(x)$ describes the probability distribution of the state $x = \begin{bmatrix} \psi & \dot{\psi} & \eta^{(\mathrm{dB})} \end{bmatrix}^\top$ of a new target given the latest measurement z_{k-1}. Since z_{k-1} contains information about the bearing ψ and the SNR $\eta^{(\mathrm{dB})}$, they are assumed to be distributed proportionally to the likelihood ratio of the measurement. That is, their joint PDF is modeled as

$$p(\psi, \eta^{(\mathrm{dB})}|z_{k-1}) \propto L(z_{k-1}|x)p(\psi)p(\eta^{(\mathrm{dB})}), \tag{35}$$

where $p(\psi)$ and $p(\eta^{(\mathrm{dB})})$ denote uniform prior distributions assigned to the bearing ψ and SNR $\eta^{(\mathrm{dB})}$, respectively. Recall that the likelihood $L(z_{k-1}|x)$ in (27) is a function of the target bearing ψ and the SNR $\eta = 10^{\eta^{(\mathrm{dB})}/10}$ which are part of the target state x. The prior $p(\eta^{(\mathrm{dB})})$ should reflect the expected SNR of yet-to-be-detected targets, that is, newly sampled targets have a low SNR. This ensures the tracker is less prone to lock onto a noisy source that momentarily

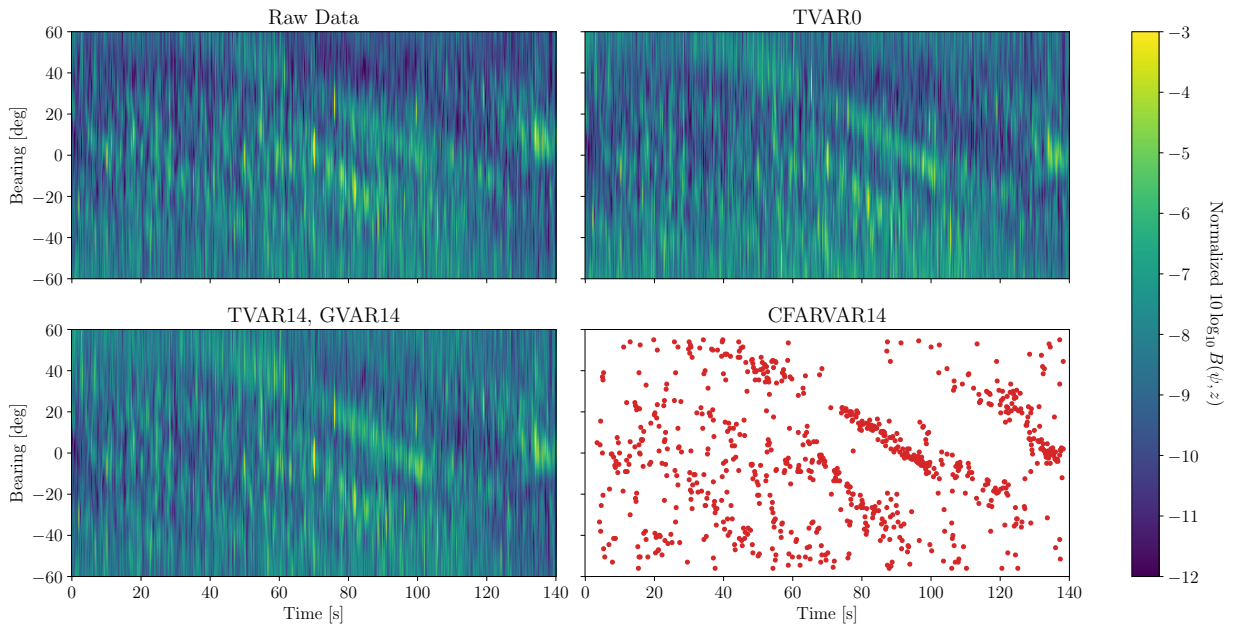


Fig. 3. CFAR detections and BTRs after the data whitening step in the different trackers. All BTRs are normalized by a factor b so that $\max_{\psi,k} bB(\psi, z_k) = 1$. The BTR of the raw samples is shown in the top left. The top right plot shows the BTR of the whitened data in the *TVAR0* tracker, which uses a VAR model of order $\varrho = 0$. The bottom left plot shows the BTR after the whitening in the *GVAR14* and *TVAR14* trackers, which uses a VAR model of order $\varrho = 14$. The bottom right plot shows the CFAR detections for the *CFARVAR14* tracker, obtained by running the detector on the whitened BTR from the GVAR14 and TVAR14 models.

produces much sound energy. Moreover, the prior $p(\psi)$ ensures that the newborn targets are born within the considered field of view of the array, and not in the end-fire directions.

The measurement z_{k-1} contains no information about the bearing change rate $\dot{\psi}$. Consequently, $\dot{\psi}$ is assumed to be normal distributed as $\mathcal{N}(\dot{\psi}; 0, P_{\dot{\psi}})$. Bringing it all together, the probability distribution of the state x of a new target is modeled as

$$b_{k|k-1}(x) = p(\psi, \eta^{(\mathrm{dB})}|z_{k-1})\mathcal{N}(\dot{\psi}; 0, P_{\dot{\psi}}). \tag{36}$$

V. Evaluation

The proposed tracking method is evaluated through experiments on both real and simulated datasets. The objective of the evaluation is to examine the impact of the ambient noise modeling and the associated data whitening, as well as the use of the t-distribution for modeling the raw acoustic data. These factors are analyzed in isolation by comparing the proposed tracker with variants that either do not include the ambient noise modeling and data whitening step or

assume the data to be Gaussian distributed. All trackers use the same underlying particle filter implementation from [23], where the only difference is the used likelihood ratio $L(z|x)$ and the used VAR model. To that end, the following five trackers will be evaluated.

- The *TVAR14* tracker, which uses a VAR model of order $\varrho = 14$ and a t-distributed signal model.
- The *TVAR0* tracker, which uses a VAR model of order $\varrho = 0$ (equivalent to only modeling the spatial correlations in the noise) and a t-distributed signal model.
- The *GVAR14* tracker, which uses a VAR model of order $\varrho = 14$ and a Gaussian distributed signal model.
- The *CFAR* tracker, which uses detections from a CFAR detector as its input.
- The *CFARVAR14* tracker, which uses detections from a CFAR detector as its input where the data fed to the detector has been whitened using a VAR model of order $\varrho = 14$.

The CFAR detection-based trackers are included for reference. Comparison between the CFAR and CFARVAR14 tracker highlights the potential effects of the whitening filter on detection-based trackers.

A. Performance Metrics

The performance of each tracker is assessed by analyzing the estimated probability that a target exists and the optimal subpattern assignment (OSPA) metric [32]. In the case of single-target tracking, the OSPA metric is given by

$$\bar{d}_f^{(\rho)}(Y_{k|k}, Y_k) = \sum_{\hat{x} \in Y_{k|k}} d^{(\rho)}(\hat{x}, x_k) + \rho^f (1 - |Y_{k|k}|), \tag{37}$$

where x_k is the ground truth target state, f is the norm order, and $d^{(\rho)}$ is the distance measure with cut-off ρ. Here, targets are compared based on their bearing differences, i.e.,

$$d^{(\rho)}(\hat{x}, x_k) = \min(\|\hat{\psi} - \psi_k\|_f, \rho). \tag{38}$$

Furthermore, $Y_{k|k}$ is the set of confirmed targets

$$Y_{k|k} = \{x \,|\, x \in X_{k|k} \text{ and } q_{k|k} > \gamma\}, \tag{39}$$

where $\gamma = 90\,\%$. The cut-off is set to be $\rho = 30^\circ$ and the norm order $f = 1$ is used.

TABLE I

PARAMETER VALUES USED IN THE EVALUATION. VALUES IN PARENTHESES ARE VALUES USED IN THE SIMULATION IF THEY DIFFER FROM THE PARAMETER VALUES USED IN THE REAL-WORLD DATA EVALUATION.

(a) TRACKER AND FILTER PARAMETERS.

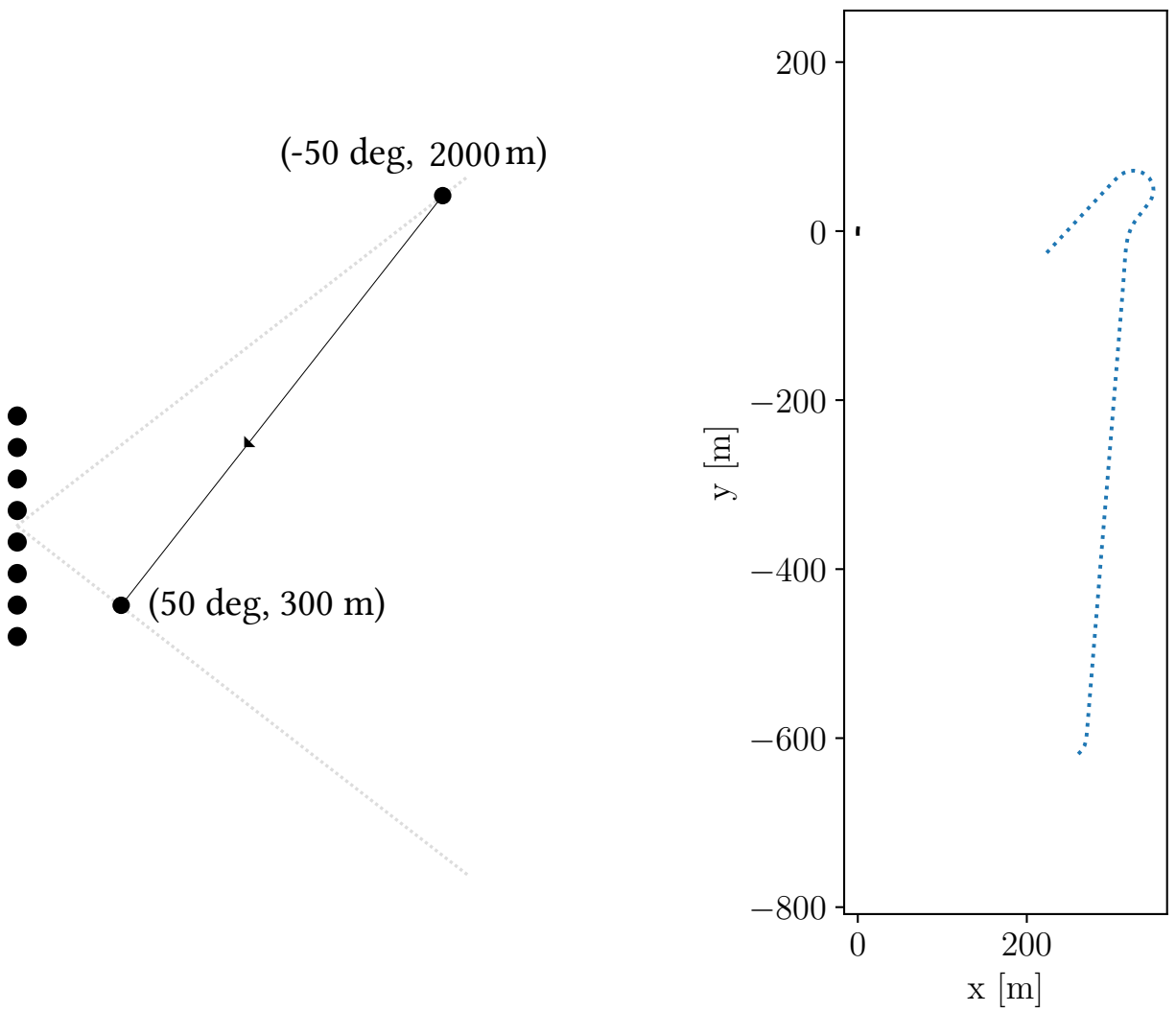
Sym.	Value	Description
p_s	$1-10^{-6}$ (0.99347)	Prob. of survival.
p_b	$2\cdot10^{-10}$ ($4.56\cdot10^{-8}$)	Prob. of target birth.
q_CV	$0.13\,^\circ/\mathrm{s}^2$	Motion process noise.
q_dBSNR	$0.05\,\mathrm{dB/s}$	SNR process noise.
$P_{\dot\psi}$	$0.001\,^{\circ 2}/\mathrm{s}^2$	Initial uncertainty in target bearing change rate
ν	3 (12)	Deg. of freedom in mv-t dist.
N	64	Number of samples per batch.
T	0.17 s	Time between batches.
ϱ	14	Model order of the VAR model.

(b) CFAR PARAMETERS. "COLUMNS" AND "ROWS" REFER TO FIG. 2.

Value	Description
25%	CFAR false alarm rate.
36	Width of CFAR guard band in bearing cells. Number of white columns.
20	Length CFAR guard band in time steps. Number of rows.
32	Total width of CFAR training cells in bearing cells. Number of yellow columns.

B. Parameter Settings

The uniform prior $p(\eta^{(\mathrm{dB})})$ on the SNR for newborn states x in (35) is calibrated individually for each tracker. The lower and upper bounds of the prior are increased simultaneously until false tracks start to appear in a target-free dataset. This ensures that each tracker is as sensitive as possible without generating false detections when applied to the target-free dataset. For the CFAR-based tracker, the false detection intensity λ in (13) is adjusted similarly. The support of the uniform prior $p(\psi)$ is -60° to 60°. All other parameters are kept the same to the greatest extent possible to ensure a fair comparison of the trackers. The parameter values used are listed



(a) Simulated scenario (not to scale). The target begins its journey in the upper part of the figure, traveling in the negative y direction. The eight dots indicate the array.

(b) Real-world scenario. The array is located at $(0, 0)$, and the target begins its journey in the lower part of the figure, traveling in the positive y direction.

Fig. 4. Illustration of the target trajectories and array locations in the simulation and real-world scenarios.

in Tab. I.

C. Simulated Scenario

The simulated scenario consists of a target that moves towards an 8-element array tuned for $800\,\mathrm{Hz}$ signals. The frequency content of the considered signals is between $750\,\mathrm{Hz}$ and $937.5\,\mathrm{Hz}$, demodulated using a $750\,\mathrm{Hz}$ cosine signal and sampled at a sampling frequency of $375\,\mathrm{Hz}$. The

Algorithm 1: Generating a simulated dataset.

input : Target bearing ψ_k and SNR η_k, VAR parameters $A_1, \ldots, A_\varrho$, Σ_w, and degrees of freedom ν of the t-distribution.

output: Sample y_k

`/* Sample ambient signal */`

Sample $\vec{w}_n \sim \mathcal{N}(0, I_M)$

Compute $\vec{e}_n$ using (29) and $\vec{w}_n$

Estimate $\sigma_e^2 = |\mathsf{E}[\vec{e}_n \vec{e}_n^{\top}]|^{1/M}$

Batch $\vec{e}_n$ into e_k as in (8a)

`/* Sample target signal */`

Calculate $\sigma_s^2 = \eta_k \sigma_e^2$

Sample $\mathrm{s}_k \sim \mathcal{N}(0, \sigma_s^2 \mathrm{H}(\psi_k)\mathrm{H}^{\top}(\psi_k))$

`/* Construct the full multivariate-t distributed signal */`

Sample $c_k \sim \chi^2(\nu)$

Compute $\mathrm{y}_k = \sqrt{\nu / c_k} \cdot (\mathrm{s}_k + \mathrm{e}_k)$

target moves at a constant velocity of $2.5\,\mathrm{m\,s^{-1}}$, starting at a bearing of $-50°$ at a distance of $2000\,\mathrm{m}$ away from the array and ending at bearing $50°$ at a distance of $300\,\mathrm{m}$ away from the array. The trajectory is shown in Fig. 4a. The range to the target is mapped to an SNR according to

$$\eta^{(\mathrm{dB})} = 10 \log_{10} \left(\frac{r}{200} \right)^{1.8}, \tag{40}$$

where r is measured in meters. This mimics a propagation loss between cylindrical and spherical spreading [33, p. 39]. As the likelihood ratio functions of the trackers are independent of the absolute noise level σ_e^2, the covariance of the innovation noise $\vec{w}_k$ can be set arbitrarily as long as it is a scaled identity matrix. For simplicity, $\mathrm{cov}(\vec{w}_k) = I_M$ is used. Generating the simulated dataset is a multistep process, as detailed in Alg. 1. Notably, the target signal is added to the correlated noise, which means that the inaccuracies of the approximation in (23) should affect the trackers similarly as it will in the later analyzed real data scenario. The same VAR model parameters are used both in the data generation and the data whitening in the tracker.

The true bearing, true SNR, $B(\psi, \mathrm{y})$, and the estimated tracks can be seen in Fig. 5. The results show that the CFAR tracker detects the target at approximately $700\,\mathrm{s}$ when the target is around

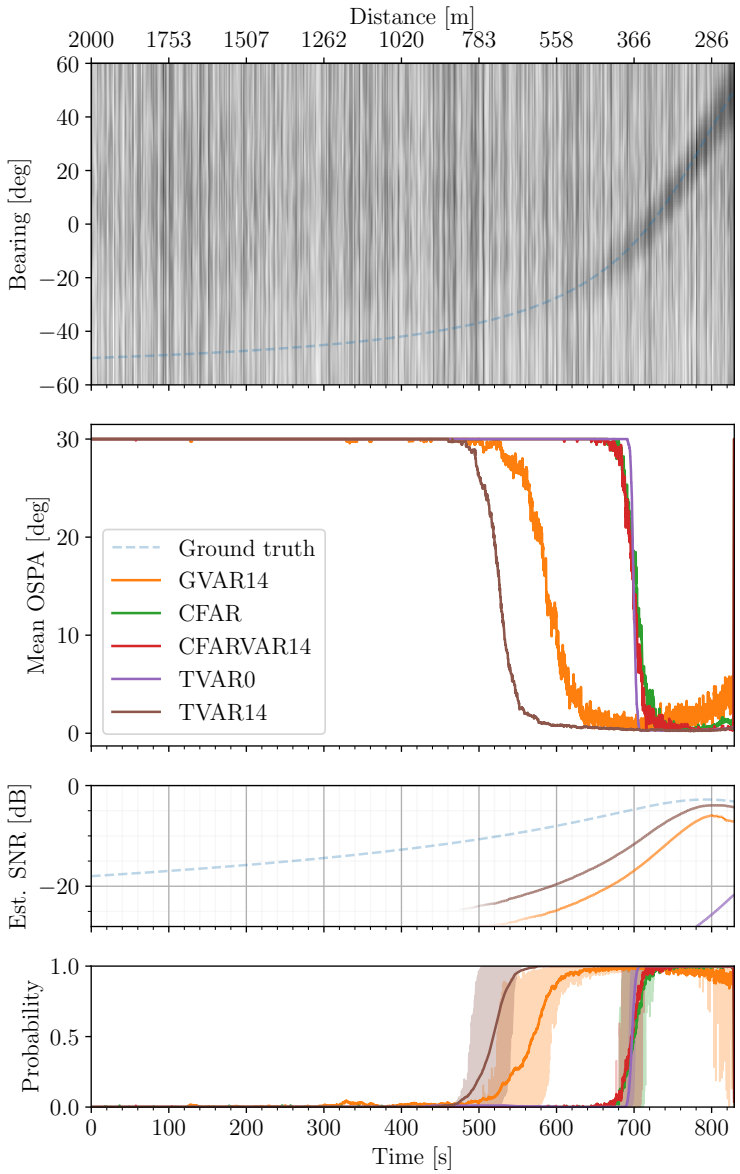


Fig. 5. Results of 100 Monte Carlo runs of the simulated scenario. The top plot shows the BTR of the generated signal for a single run, together with the ground truth bearing in dashed blue. Below the BTR, the OSPA, estimated SNR, and estimated target existence probability are shown. The transparent regions in the estimated target existence probability correspond to the top 90 and bottom 10 quantiles.

300 m from the array. This corresponds to an SNR of -5 dB. Prewhitening the measurements before constructing the detections does not contribute to any significant performance gain, as evident when one compares the CFAR tracker to the CFARVAR14 tracker. The *TVAR0* tracker shows similar performance. Modeling the ambient noise using VAR model of order 14 improves the detection performance, as demonstrated by the *GVAR14* tracker. The *GVAR14* tracker detects

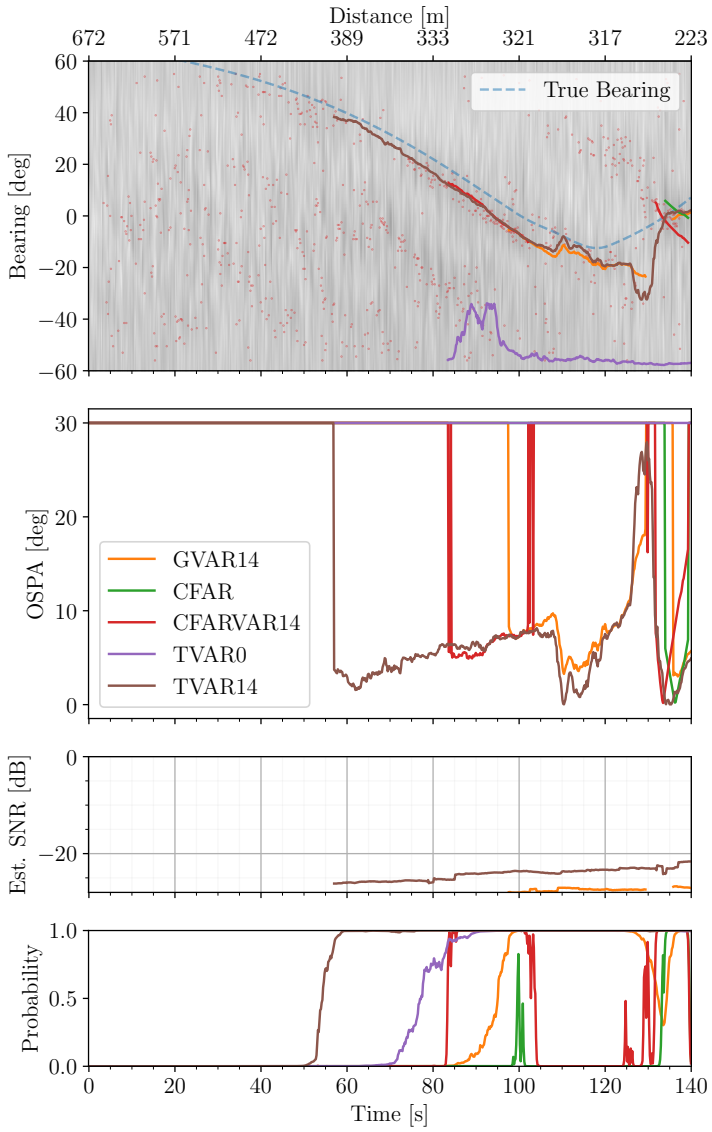


Fig. 6. Results from the real-world scenario. The top plot shows the BTR, along with the truth and estimated bearings. Below the BTR, the OSPA, estimated SNR, and estimated target existence probability are shown.

the target at a distance of $550\,\mathrm{m}$, which corresponds to an SNR of $-8\,\mathrm{dB}$. However, the estimated target existence probability fluctuates rapidly above and below the detection threshold, causing the OSPA metric to vary correspondingly. Further improvement is observed when the Gaussian distribution is replaced by the proposed multivariate-t distribution, as seen from the performance of the *TVAR14* tracker. The multivariate-t distribution stabilizes the estimates of target existence probability and leads to a smoother OSPA measure. It also enhances the tracker's detection

capability, enabling even earlier target detection. The target is detected at a distance of $560\,\mathrm{m}$, which corresponds to an SNR of $-9\,\mathrm{dB}$. It is noted that the SNR is systematically underestimated by the trackers. There are two likely causes behind this. First, it may be due to the approximations made in the derivation of the measurement model. Secondly, newborn targets are assumed to have a low SNR to increase the robustness of the tracker against false tracks, resulting in a bias.

D. Real-world Scenario

The real-world dataset was collected during a sea trial in the Stockholm archipelago using an 8-element horizontal hydrophone array and a SAAB AUV62 autonomous underwater vehicle acting as the target. The array's shape, orientation, and location were calibrated using the method described in [34]. The distance between the hydrophone elements was approximately $0.93\,\mathrm{m}$. Given that the speed of sound in the baltic sea is approximately $1500\,\mathrm{m\,s^{-1}}$, this corresponds to a design frequency of $800\,\mathrm{Hz}$. The SAAB AUV62 followed the trajectory shown in Fig. 4 (b), starting at a distance of $675\,\mathrm{m}$ and ending at $220\,\mathrm{m}$ from the array. Throughout the trajectory, the AUV maintained a speed of $6.2\,\mathrm{m\,s^{-1}}$ (approximately $12\,\mathrm{knots}$) and a constant depth of $25\,\mathrm{m}$. The depth ranges from $30\,\mathrm{m}$ to $45\,\mathrm{m}$ at the test site, with a few islands located to the left of the array in the plot. Due to the shallow water at the site, the hydroacoustic environment is expected to be complex, with many reflections and other unmodeled properties. As a result, the t-distribution model in the *TVAR14* and *TVAR0* trackers were set to have a degree-of-freedom $\nu = 3$, the lowest value for which the covariance matrix of the t-distribution remains defined.

Before the trackers processed the recorded hydrophone, the data was preprocessed as follows. Similarly to the simulated scenario, the data were bandpass-filtered with cut-off frequencies $750\,\mathrm{Hz}$ and $937.5\,\mathrm{Hz}$, then demodulated using a $750\,\mathrm{Hz}$ cosine signal and finally downsampled to a sampling frequency of $375\,\mathrm{Hz}$. In the frequency band $750\,\mathrm{Hz}$ and $937.5\,\mathrm{Hz}$, the received signal can be considered approximately white. This is because the spectrum of the sound generated by the AUV62 is approximately flat, and the differences in propagation loss across the frequency band are negligible.

1) Effects of Ambient Noise Modeling: In Fig. 3, the effect of whitening the hydrophone data in Fig. 1b with a VAR model learned on the data in Fig. 1a, is shown in terms of BTR after the whitening process. Also shown are the detections found by applying the CFAR detector to the same dataset. Note that the *CFARVAR14*, *TVAR14* and *GVAR14* trackers use the same

VAR model of order $\varrho = 14$ for the ambient noise modeling and data whitening, whereas the *TVAR0* tracker uses a $\varrho = 0$ order model. The effects of the whitening processes are apparent by comparing the BTR after the data whitening in the *TVAR14* and *TVAR0* trackers to the BTR calculated on the raw data. The energy distribution across the bearings is more uniform after the whitening. Specifically, bearings within $20°$ to $60°$ exhibit lower signal energy in the BTR of the raw data and in *TVAR0*, compared to *TVAR14* and *GVAR14*. Conversely, the energy in bearings $-60°$ to $-20°$ during the timespan $90\,\mathrm{s}$ to $110\,\mathrm{s}$ is higher in *TVAR0*

2) Detection and Tracking Performance: In Fig. 6, the estimated target track, SNR, target existence probability, and the calculated OSPA for the five trackers are shown. The detections and tracks are slightly offset compared to the true bearing, which may be due to errors in the calibration of the array geometry and orientation. The *CFAR* tracker detects and starts to track the target at a distance of $250\,\mathrm{m}$. By prewhitening the data before constructing the detections, the the *CFARVAR14* tracker is able to detect the target at a distance of $325\,\mathrm{m}$ and maintains the track for $20\,\mathrm{s}$ before it is dropped. The *TVAR0* tracker detects the target at a distance of $325\,\mathrm{m}$, but the estimated bearing is too far from the ground truth to be considered a valid track. This is in line with the observations in Fig. 3, where the negative bearings exhibited higher beamforming energy due to the absence of temporal whitening. The *GVAR14* tracker detects the target at approximately $325\,\mathrm{m}$ and maintains the track for approximately $125\,\mathrm{s}$ until the SNR temporarily decreases. However, the method quickly recovers the track.

A common issue across the *CFAR*, *CFARVAR14*, *GVAR14*, and *TVAR0* trackers is that they either lose track or initiate false tracks. In contrast, the *TVAR14* tracker detects the target the earliest, when the target is $390\,\mathrm{m}$ from the array, and maintains a stable track thereafter.

VI. Discussion and Conclusions

The challenge of reliable broadband passive sonar target detection and tracking in complex acoustic environments has been addressed. A solution has been proposed based on a vector-autoregressive model for the ambient noise and a heavy-tailed statistical model for the distribution of the raw hydrophone data. These models have been integrated into a Bernoulli track-before-detect (TkBD) filter to realize a bearing-only tracker. To facilitate a computationally less expensive evaluation of the proposed statistical model, approximations have been introduced to compensate for the effects of the ambient noise via a recursive preprocessing step where the data is whitened. The whitening of the data facilitates the statistical model to be expressed as

a function of the conventional beamformer, significanly reducing the computational complexity of the statistical model.

The proposed solution has been evaluated on both simulated and real-world data. Results showed that the proposed vector-autoregressive model can learn and compensate for a lot of the spatiotemporal correlations in ambient noise, which was also seen to be useful for conventional detection-based trackers. Moreover, the results showed that the proposed heavy-tailed multivariate-t distribution model made the trackers more robust than the case when the data was modeled as Gaussian distributed. The simulations show that the SNR at which the target can be detected is reduced by $4\,\mathrm{dB}$ compared to when using the standard constant false alarm rate detector based tracker. Further, the test with real-world data shows that using the vector-autoregressive model improves the detection distance of the conventional CFAR tracker from $250\,\mathrm{m}$ to $325\,\mathrm{m}$. The proposed solution increases the target detection distance further to $390\,\mathrm{m}$.

The presented results illustrate that the TkBD technology, in combination with data-driven ambient noise modeling and heavy-tailed statistical signal models, can enable reliable broadband passive sonar target detection and tracking in complex acoustic environments and lower the SNR required to detect and track targets. Hence, the technology can contribute to more effective monitoring of critical underwater infrastructure and potentially increase the safety of maritime activities. Additionally, the technology's capability to whiten stochastic disturbances makes it well-suited for use in the archipelago or other shallow water environments with complex acoustic interference.

While the proposed solution shows promise, it has limitations. If the motion model does not align with the true target motion, tracking effectiveness might be compromised. Additionally, assuming that the target is stationary during each batch could lead to "smearing" of energy if the target moves quickly. Although this has not been investigated, the proposed solution might already be able to adapt to this by using a more spread-out particle cloud.

Future research should explore the possibility to integrate the proposed ambient noise model and heavy-tailed signal model in multi-target tracking. While multiple target TkBD is still in development, methods such as the information exchange filter [35] and belief propagation for multi-Bernoulli filters [36] have shown promising results. Multi-target TkBD tracking using superpositional maritime radar measurements has been explored in [37]. The reason behind the observed bias in the SNR estimates should also be investigated.

APPENDIX

Given the PDF in (19), it holds that

$$\begin{aligned}\ell_0(z) &= t_{NM}\left(z;\nu,0,\frac{\nu-2}{\nu}I_{NM}\right)\\ &= \frac{C}{|I_{NM}|^{1/2}}\left(1+\frac{\|z\|^2}{\nu-2}\right)^{-(\nu+NM)/2},\end{aligned}$$

for some constant C and

$$\begin{aligned}\ell_1(z|x) &= t_{NM}\left(z;\nu,0,\frac{\nu-2}{\nu}\Sigma\right)\\ &= \frac{C}{|\Sigma|^{1/2}}\left[1+\frac{z^\top\Sigma^{-1}z}{\nu-2}\right]^{-(\nu+NM)/2},\end{aligned}$$

where $\Sigma = \eta \mathrm{H}(\psi)\mathrm{H}^\top(\psi) + I_{NM}$. According to [20], [21], it holds that

$$z^\top\Sigma^{-1}z \approx \|z\|^2 - \frac{\eta B(\psi,z)}{(1+M\eta)},$$

and

$$\ln|\Sigma| \approx N\ln(M\eta+1).$$

This gives the likelihood ratio

$$\begin{aligned}\ln L(z|x) =& \ln\ell_1(z|x) - \ln\ell_0(z)\\ \approx& -\frac{N}{2}\ln\left(M\eta+1\right)\\ &+\frac{\nu+NM}{2}\left(\ln\left(1+\frac{\|z\|^2}{\nu}\right)\right.\\ &\left.-\ln\left(1+\frac{\|z\|^2}{\nu}-\frac{\eta B(\psi,z)}{\nu(1+M\eta)}\right)\right),\end{aligned}$$

By using that $\ln(a)-\ln(a-b) = -\ln(1-b/a)$, the likelihood may be rewritten as

$$\begin{aligned}\ln L(z|x) \approx& -\frac{N}{2}\ln\left(M\eta+1\right)\\ &-\frac{\nu+NM}{2}\ln\left(1-cB(\psi,z)\right),\end{aligned}$$

where

$$c = \frac{\eta}{(\nu+\|z\|^2)(1+M\eta)}.$$

REFERENCES

[1] G. Soldi *et al.*, "Monitoring of Critical Undersea Infrastructures: The Nord Stream and Other Recent Case Studies," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 38, no. 10, pp. 4–24, Oct. 2023.

[2] S. Bashfield, "Defending seabed lines of communication," *Australian Journal of Maritime & Ocean Affairs*, pp. 1–13, Jul. 2024.

[3] L. Alleslev, "NATO Anti-submarine warfare: rebuilding capability, preparing for the future," *Science and Technology Committee (STC)*, 2019.

[4] S. J. Dolman, P. G. Evans, G. Notarbartolo-di Sciara, and H. Frisch, "Active sonar, beaked whales and European regional policy," *Mar. Pollut. Bull.*, vol. 63, no. 1–4, pp. 27–34, 2011.

[5] E. C. M. Parsons, "Impacts of Navy Sonar on Whales and Dolphins: Now beyond a Smoking Gun?" *Front. Mar. Sci.*, vol. 4, Sep. 2017.

[6] S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 5–18, Jan. 2004.

[7] A. F. Garcia-Fernandez, J. L. Williams, K. Granstrom, and L. Svensson, "Poisson Multi-Bernoulli Mixture Filter: Direct Derivation and Implementation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1883–1901, Aug. 2018.

[8] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE J. Ocean. Eng.*, vol. 8, no. 3, pp. 173–184, Jul. 1983.

[9] X. Li, B. Lu, W. Ali, J. Su, and H. Jin, "Passive Sonar Multiple-Target Tracking with Nonlinear Doppler and Bearing Measurements Using Multiple Sensors," *Int. J. Aerosp. Eng.*, vol. 2021, pp. 1–11, Oct. 2021.

[10] B. Ristic, S. Arulampalam, and N. J. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Norwood, MA, USA: Artech House, 2003.

[11] M. Mallic, V. Krishnamurthy, and B. Vo, *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*. Wiley-IEEE Press, 2012, ch. 8, pp. 315–318.

[12] D. A. Abraham, *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation*. Cham, Switzerland: Springer, 2019.

[13] G. Song, X. Guo, and L. Ma, "The Alpha stable distribution in ocean ambient noise modelling," *MATEC Web Conf.*, vol. 283, p. 08002, 2019.

[14] Y. Zhang, K. Yang, and Q. Yang, "Probability density function of ocean noise based on a variational Bayesian Gaussian mixture model," *The Journal of the Acoustical Society of America*, vol. 147, no. 4, pp. 2087–2097, Apr. 2020.

[15] L. Xu, C. Liu, W. Yi, G. Li, and L. Kong, "A particle filter based track-before-detect procedure for towed passive array sonar system," in *Proc. IEEE Radar Conf.*, Seattle, WA, May 2017.

[16] Q. Peng, W. Li, and L. Kong, "Multi-frame Track-before-detect Algorithm for Passive Sonar System," in *International Conference on Control, Automation and Information Sciences (ICCAIS)*, Chengdu, China, Oct. 2019, pp. 1–6.

[17] T. Northardt and S. C. Nardone, "Track-Before-Detect Bearings-Only Localization Performance in Complex Passive Sonar Scenarios: A Case Study," *IEEE J. Ocean. Eng.*, vol. 44, no. 2, pp. 482–491, Apr. 2019.

[18] W. Yi, L. Fu, A. F. García-Fernández, L. Xu, and L. Kong, "Particle filtering based track-before-detect method for passive array sonar systems," *Signal Process.*, vol. 165, pp. 303–314, Dec. 2019.

[19] B. A. Yocom, B. R. La Cour, and T. W. Yudichak, "A Bayesian Approach to Passive Sonar Detection and Tracking in the Presence of Interferers," *IEEE J. Ocean. Eng.*, vol. 36, no. 3, pp. 386–405, Jul. 2011.

[20] D. Bossér, R. Forsling, I. Skog, G. Hendeby, and M. L. Nordenvaad, "Underwater Environment Modeling for Passive Sonar Track-Before-Detect," in *OCEANS 2023 - Limerick*, Limerick, Ireland, Jun. 2023, pp. 1–6.

[21] D. Bosser, G. Hendeby, M. L. Nordenvaad, and I. Skog, "A Statistically Motivated Likelihood for Track-Before-Detect," in *2022 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, vol. 5204, Bedford, UK, Sep. 2022, pp. 1–6.

[22] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, Inc., 2007.

[23] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina, "A Tutorial on Bernoulli Filters: Theory, Implementation and Applications," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3406–3430, Jul. 2013.

[24] S.-C. Pei and Y.-C. Lai, "Closed Form Variable Fractional Time Delay Using FFT," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 299–302, May 2012.

[25] M. A. Richards, *Fundamentals of Radar Signal Processing, Third Edition*, third edition. ed. New York, NY, USA: McGraw-Hill, 2022.

[26] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining*, ser. KDD'96. AAAI Press, 1996, p. 226–231.

[27] A. Barthelemy and P. Willett, "An algorithm for prewhitening a large parallel line array," in *1995 International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, Detroit, MI, May, pp. 3543–3546.

[28] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Berlin, Germany: Springer, June 2005.

[29] L. Ljung, *System identification (2nd ed.): theory for the user*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[30] J. Liu, S. Kumar, and D. P. Palomar, "Parameter Estimation of Heavy-Tailed AR Model with Missing Data via Stochastic EM," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2159–2172, Apr. 2019.

[31] X. Rong Li and V. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.

[32] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A Consistent Metric for Performance Evaluation of Multi-Object Filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[33] FMV, *Hydroakustik och sonarteknik för marinen v. 2.2*, 2014.

[34] I. Skog, M. L. Nordenvaad, and G. Hendeby, "Signals-of-Opportunity-Based Hydrophone Array Shape and Orientation Estimation," *IEEE J. Ocean. Eng.*, vol. 49, no. 3, pp. 679–691, Jul. 2024.

[35] E. S. Davies and A. F. García-Fernández, "Information Exchange Track-Before-Detect Multi-Bernoulli Filter for Super-positional Sensors," *IEEE Trans. Signal Process.*, vol. 72, pp. 607–621, 2024.

[36] M. Liang, T. Kropfreiter, and F. Meyer, "A BP Method for Track-Before-Detect," *IEEE Signal Process. Lett.*, vol. 30, pp. 1137–1141, 2023.

[37] D. Y. Kim, B. Ristic, R. Guan, and L. Rosenberg, "A Bernoulli Track-Before-Detect Filter for Interacting Targets in Maritime Radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 3, pp. 1981–1991, Jan. 2021.