

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

Daniel Bossér, 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 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 m to 390 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.

**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

This work has been funded 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.

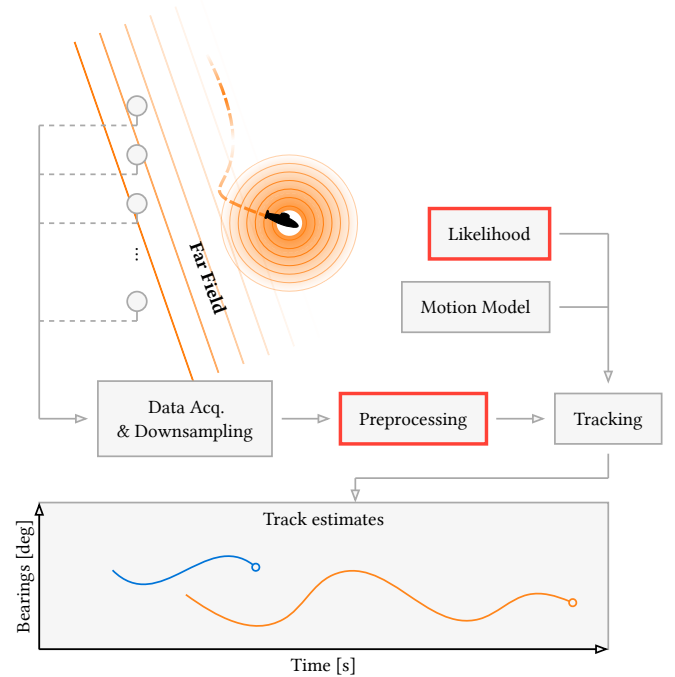


Fig. 1. Illustration of the primary components in a passive sonar target tracking application. Sound signals are recorded using a hydrophone array. The recorded signals are preprocessed and fed into a target-tracking algorithm, which determines whether one or more targets are present and their trajectories. In this work, the two components marked in red, i.e., the preprocessing and measurement model (likelihood function), are studied, and a new preprocessing method and measurement model is proposed.

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]. This creates systems like the one illustrated in Fig. 1. 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 processing, 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 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. 2a, which shows the received signal energy in different bearings over time. From Fig. 2a, 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. 2b. 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 spatio-temporal 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

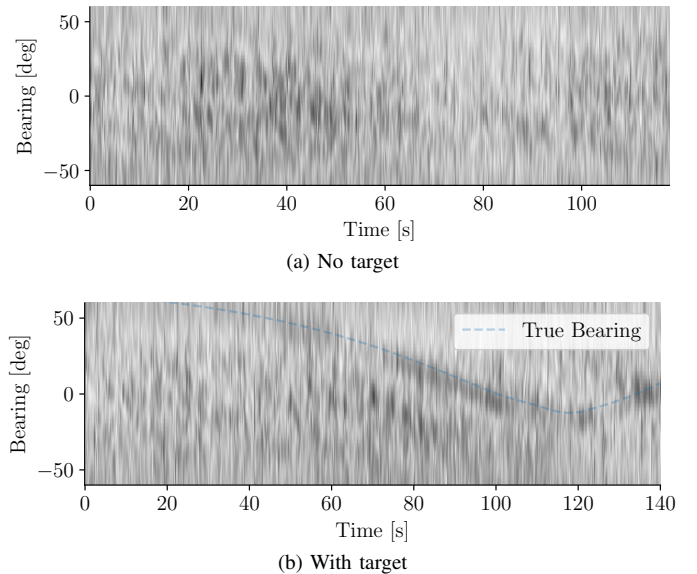


Fig. 2. BTRs from data collected during a sea trial. The BTR values were calculated using a conventional delay-and-sum beamformer. The upper and lower plots show the BTR without and with a target present. From the plots, it is clear that the temporal and spatially correlated ambient noise causes disturbances that are hard to distinguish from the actual target.

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 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](https://gitlab.liu.se/coast/tkbd_using_raw_data).  
s

## II. TARGET TRACKING AND BAYESIAN ESTIMATION

The objective of target detection and tracking is to, given a set of measurements  $z_{1:k} = \{z_1, z_2, \dots, z_k\}$  collected at discrete time instances  $1, 2, \dots, k$ , 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]

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

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

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

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]

$$\begin{aligned} q_{k|k-1} &= p_b(1 - q_{k-1|k-1}) + p_s q_{k-1|k-1} \quad (3a) \\ s_{k|k-1}(x) &= \frac{p_b(1 - q_{k-1|k-1}) b_{k|k-1}(x)}{q_{k|k-1}} \\ &\quad + \frac{p_s q_{k-1|k-1} \int \pi_{k|k-1}(x|x') s_{k-1|k-1}(x') dx'}{q_{k|k-1}} \end{aligned} \quad (3b)$$

and

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

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

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  $x_i$  conditioned on the state  $x'$ , describing the target motion. Finally,  $L(z_k|x)$  is the conditional likelihood ratio of given the target state  $x$ . This article uses the particle filter in [23] to implement and execute these filter recursions.

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

$$\vec{y}_n = [y_n^{(0)} \quad \dots \quad y_n^{(M-1)}]^\top \in \mathbb{R}^M, \quad (5)$$

as the collection of samples from all the  $M$  hydrophones. Further, let

$$y_k = [\vec{y}_{(k-1)N+1}^\top \quad \dots \quad \vec{y}_{kN}^\top]^\top \in \mathbb{R}^{NM}, \quad (6)$$

be a batch of  $N$  hydrophone samples.

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

$$y_k = H(\psi) s_k + e_k, \quad (7)$$

where  $e_k$  is given by

$$e_k = [\vec{e}_{(k-1)N+1}^\top \quad \dots \quad \vec{e}_{kN}^\top]^\top \in \mathbb{R}^{NM}, \quad (8a)$$

$$\vec{e}_n = [e_n^{(0)} \quad \dots \quad e_n^{(M-1)}]^\top \in \mathbb{R}^M, \quad (8b)$$

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

$$H(\psi) = [H_0^\top(\psi) \quad \dots \quad H_{M-1}^\top(\psi)]^\top \in \mathbb{R}^{NM \times N}. \quad (9)$$

where

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

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

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

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)  $\psi$ . 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  $B(\psi, y_k)$  for DOA angle  $\psi$  and hydrophone sample batch  $y_k$  is given by

$$B(\psi, y_k) = \|H^\top(\psi) y_k\|_2^2. \quad (11)$$

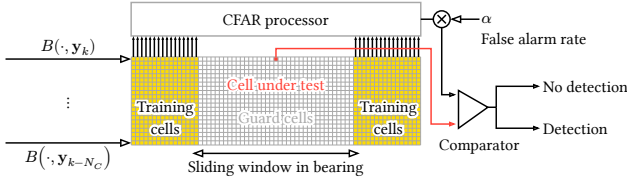


Fig. 3. Illustration of the CFAR detector. The CFAR processor evaluates a cell under test to determine whether it contains a detection or not. This is done by comparing the energy in the cell under test with the energy in training cells.

### 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  $B(\psi, y_k)$ . The resulting measurements are a set of bearings,

$$z_k \equiv \{\psi_0, \dots, \psi_D\}, \quad (12)$$

where the number of detections  $D$  will vary. The cell-averaging CFAR detector evaluates  $B(\psi, y_k)$  over a discrete set of angles  $\psi$ . Each bearing-value pair  $(\psi, B(\psi, y_k))$ , 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. 3. More details on the CFAR detector can be found in [25].

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

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

Here,  $\lambda$  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

$$\ell_1(z|x) = \ell_0(z)(1 - p_d) + p_d \sum_{\psi_d \in z} g(\psi_d|x) \ell_0(z \setminus \psi_d), \quad (14)$$

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  $\psi_d$  due to the target. The measurement uncertainty is assumed to follow a Gaussian distribution with variance  $R$ . That is,

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

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

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

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

The proposed TkBD measurement model assumes that each batch of hydrophone samples follows a multivariate-t distribution

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

where

$$\Sigma(\psi) = H(\psi) \Sigma_{ss} H^T(\psi) + \Sigma_{ee}, \quad (18)$$

$\Sigma_{ss} = \text{cov}(s_k)$ , and  $\Sigma_{ee} = \text{cov}(e_k)$ . Further, the PDF of the multivariate-t distribution is given by

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

where  $\Gamma$  is the gamma function,  $\nu$  is the degrees of freedom,  $\mu$  is the median, and  $S$  is the scale matrix. For  $\nu > 2$  then  $\text{cov}(y) = \frac{\nu}{\nu - 2} S$ . Noteworthy is that when  $\nu \rightarrow \infty$ , 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  $\Sigma(\psi) = \sigma_s^2 H(\psi) H^*(\psi) + \sigma_e^2 I_{NM}$ , then

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

and

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

Hence, under these assumptions, the log of the PDF in (19) can be efficiently calculated from the beamformer  $B(\psi, y)$  and the signal energy  $\|y\|^2$ .

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  $\Sigma_{ee}$  is known the batch data  $y_k$  may be whitened as

$$\tilde{y}_k = \Sigma_{ee}^{-1/2} y_k \sim t_{NM}(\nu, 0, \bar{\Sigma}_{ss}(\psi) + I_{NM}), \quad (21)$$

where

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

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

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

Here,  $\eta$  is the power of the signal originating from the target relative to the ambient noise power after whitening, i.e.,  $\eta$  is the SNR after whitening. The motivation for making the ad-hoc approximation in (23) is that it enables the PDF of  $\tilde{y}_k$  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  $\tilde{y}_k$  of whitened hydrophone samples as inputs, i.e.,

$$z_k \equiv \tilde{y}_k. \quad (24)$$

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

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

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

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

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

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

where

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

For a derivation, see Appendix.

#### D. Data-efficient Learning of $\Sigma_{ee}$ using VAR Models

Typically, the covariance  $\Sigma_{ee}$  of the ambient noise is unknown and must be learned from historical data. Directly estimating  $\Sigma_{ee}$  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  $\Sigma_{ee}$  and, as will be shown, the whitening to be done without factorizing and inverting  $\Sigma_{ee}$ . VAR models have previously been successfully applied in other sonar applications to capture

the spatial and temporal characteristics of underwater sounds [26], though they are more commonly used in economic modeling to capture relationships between variables over time [27].

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

$$\vec{e}_n = A_1 \vec{e}_{n-1} + \dots + A_p \vec{e}_{n-p} + \Sigma_w^{1/2} \vec{w}_n, \quad (29)$$

where  $\vec{w}_n$  is white noise with  $\text{cov}(\vec{w}_n) = I_M$ . The matrices  $A_1, \dots, A_p$ , and  $\Sigma_w$  are model parameters that defines the structure of  $\Sigma_{ee}$ . However, instead of directly constructing  $\Sigma_{ee}$  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  $\vec{e}_n, \dots, \vec{e}_{n-p}$  as follows

$$\vec{w}_n = \Sigma_w^{-1/2} (\vec{e}_n - A_1 \vec{e}_{n-1} - \dots - A_p \vec{e}_{n-p}). \quad (30)$$

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

The model parameters  $A_1, \dots, A_p, \Sigma_w$  can be learned using linear least squares [27, Ch. 3]. That is, the parameters are calculated as

$$\{\hat{A}_1, \dots, \hat{A}_p\} = \arg \min_{A_1, \dots, A_p} \sum_{n=p+1}^{N_t} \epsilon_n^\top \epsilon_n \quad (31a)$$

where

$$\epsilon_n = \vec{y}_n - A_1 \vec{y}_{n-1} - \dots - A_p \vec{y}_{n-p}, \quad (31b)$$

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

$$\hat{\Sigma}_w = \frac{1}{N_t - p - 1} \sum_{n=p+1}^{N_t} \epsilon_n \epsilon_n^\top. \quad (32)$$

The model order  $p$  can be selected using, e.g., the Akaike information criterion [28, p.221]. An alternative method for learning the parameters that take into account the heavy-tailed distribution of the data can be found in [29].

## IV. TARGET DYNAMICS MODEL

Recall from Sec. II that the Bernoulli filter recursions require a mode 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

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

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

where  $\dot{\psi}_k$  is the bearing change rate and  $\eta_k^{(\text{dB})}$  is the SNR  $\eta_k$  in dB, i.e.,  $\eta_k^{(\text{dB})} = 10 \log_{10} \eta_k$ . Changes in target bearing  $\psi$  are modeled according to a constant velocity model, and the SNR  $\eta^{(\text{dB})}$  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 [30]. Hence, conditioned on target state  $x_k$ , the PDF of  $x_{k+1}$  is modeled as

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

where

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

Here  $T$  is the time between instant  $k$  and  $k+1$ . Further,  $q_{\text{CV}}^2$  and  $q_{\text{dBSNR}}^2$  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  $x$  of a new target given the latest measurement  $z_{k-1}$ . Since  $z_{k-1}$  contains information about the bearing  $\psi$  and the SNR  $\eta^{(\text{dB})}$ , they are assumed to be distributed proportionally to the likelihood ratio of the measurement. That is, their joint PDF is modeled as

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

where  $p(\eta^{(\text{dB})})$  denotes an uniform prior distribution assigned to the SNR  $\eta^{(\text{dB})}$ . The prior  $p(\eta^{(\text{dB})})$  should reflect the expected SNR of yet to be detected targets. This ensures the tracker is less prone to lock onto a noisy source that momentarily produces much sound energy.

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

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

## 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. To that end, the following four trackers will be evaluated.

- The *TVAR14* tracker, which uses a VAR model of order  $p = 14$  and a t-distributed signal model.
- The *TVAR0* tracker, which uses a VAR model of order  $p = 0$  (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  $p = 14$  and a Gaussian distributed signal model.
- The *CFAR* tracker, which uses detections from a CFAR detector as its input.

The CFAR detection based tracker is included as a reference. One may argue that the CFAR based tracker should include

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 EVALUATION USING REAL-WORLD DATA.

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

a whitening step similar to the other trackers. However, the CFAR tracker uses spatially and temporally distributed training cells, which should compensate for the spatiotemporal correlated ambient noise. Applying a whitening step would duplicate existing compensatory mechanisms.

### 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 [31]. In the case of single-target tracking, the OSPA metric is given by

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

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

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

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

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

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

### B. Parameter Settings

The uniform prior  $p(\eta^{(\text{dB})})$  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  $\lambda$  in (13) is adjusted similarly. 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 in Tab. I.



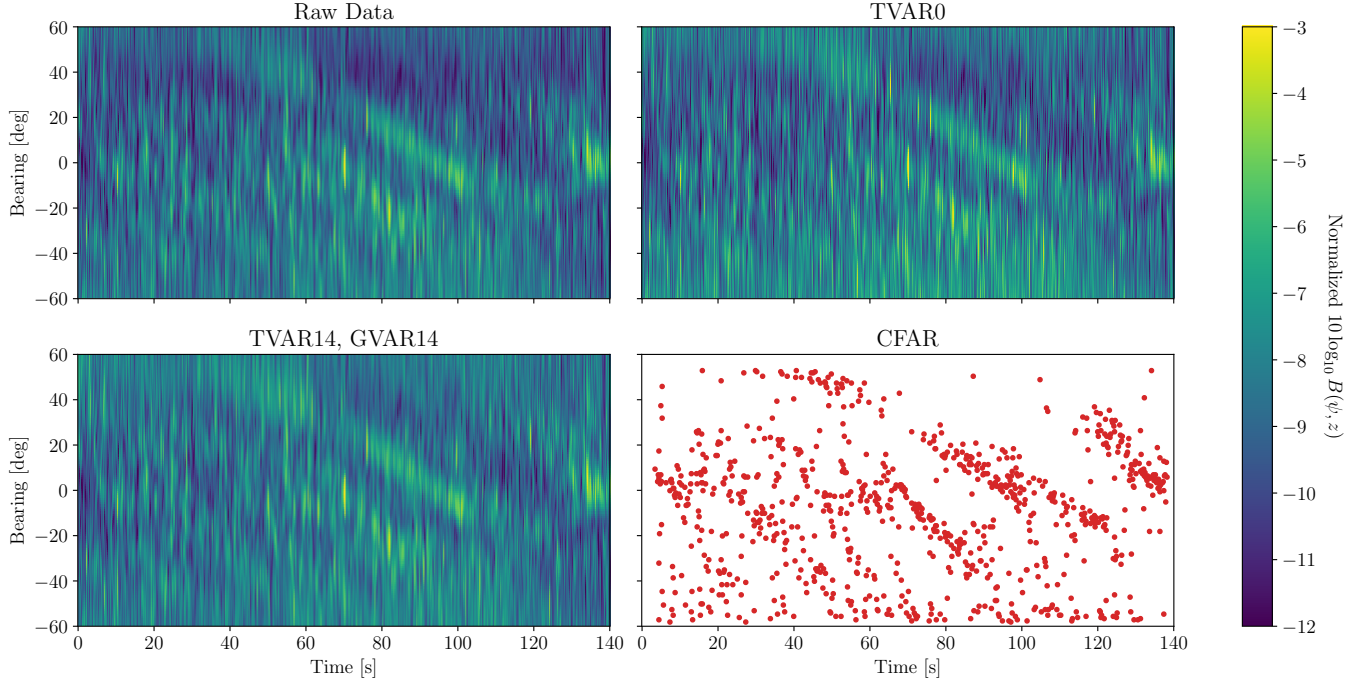
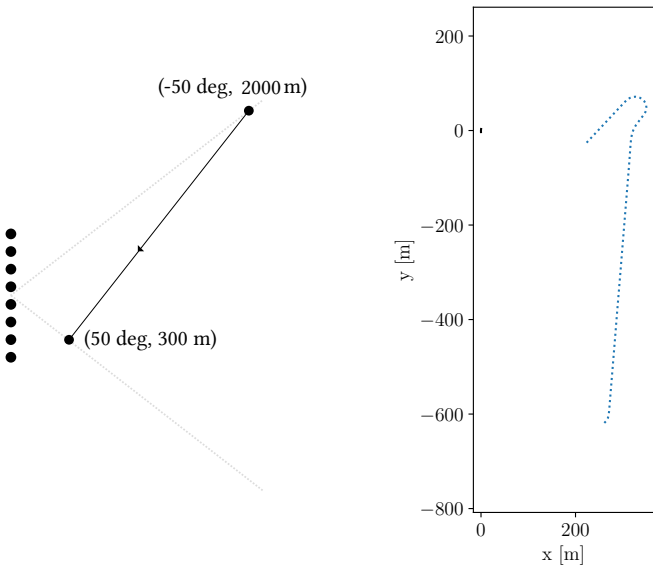


Fig. 4. CFAR detections and BTRs after the data whitening step in the different trackers. All BTRs are normalized by a factor  $b$  so that  $\max_{\psi, k} bB(\psi, z_k) = 1$ . The BTR of the raw samples is shown in the top left. The CFAR detections are shown in the bottom right. The top right plot shows the BTR of the whitened data in the *TVAR0* tracker, which uses a VAR model of order  $p = 0$ . The bottom left plot shows the BTR after the whitening in the *GVAR14* and *TVAR14* trackers, which uses a VAR model of order  $p = 14$ .



(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. 5. Illustration of the target trajectories and array locations in the simulation and real-world scenarios.

### C. Simulated Scenario

The simulated scenario consists of a target that moves towards the array at a constant velocity of 2.5 m/s, starting

at a bearing of  $-50^\circ$  at a distance of 2000 m away from the array and ending at bearing  $50^\circ$  at a distance of 300 m away from the array. The trajectory is shown in Fig. 5a. The range to the target is mapped to an SNR according to

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

where  $r$  is measured in meters. This mimics a propagation loss between cylindrical and spherical spreading [32, p. 39]. 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,  $B(\psi, y)$ , and the estimated tracks can be seen in Fig. 6. The results show that the CFAR tracker detects the target at approximately 700 s when the target is around 300 m from the array. This corresponds to an SNR of  $-5$  dB. 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 the target at a distance of 550 m, which corresponds to an SNR of  $-8$  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

**Algorithm 1:** Generating a simulated dataset.

---

**input :** Target bearing  $\psi_k$  and SNR  $\eta_k$ , VAR parameters  $A_1, \dots, A_p$ ,  $\Sigma_w$ , and degrees of freedom  $\nu$  of the t-distribution.

**output:** Sample  $y_k$

/\* Sample ambient signal \*/

- 1 Sample  $\vec{w}_n \sim \mathcal{N}(0, I_M)$
- 2 Compute  $\vec{e}_n$  using (29) and  $\vec{w}_n$
- 3 Estimate  $\sigma_e^2 = |\mathbb{E}[\vec{e}_n \vec{e}_n^T]|^{1/M}$
- 4 Batch  $\vec{e}_n$  into  $e_k$  as in (8a)
- /\* Sample target signal \*/
- 5 Calculate  $\sigma_s^2 = \eta_k \sigma_e^2$
- 6 Sample  $s_k \sim \mathcal{N}(0, \sigma_s^2 \mathbf{H}(\psi_k) \mathbf{H}^T(\psi_k))$
- /\* Construct the full multivariate-t distributed signal \*/
- 7 Sample  $c_k \sim \chi^2(\nu)$
- 8 Compute  $y_k = \sqrt{\nu/c_k} \cdot (s_k + e_k)$

---

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 m, which corresponds to an SNR of  $-9$  dB. It is noted that the SNR is systematically underestimated by the trackers, which may result from the approximations made in the derivation of the measurement model.

#### 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 [33]. The distance between the hydrophone elements was approximately 0.93 m. Given that the speed of sound in the baltic sea is approximately  $1500 \text{ m s}^{-1}$ , this corresponds to a design frequency of 800 Hz. The SAAB AUV62 followed the trajectory shown in Fig. 5 (b), starting at a distance of 675 m and ending at 220 m from the array. Throughout the trajectory, the AUV maintained a speed of  $6.2 \text{ m s}^{-1}$  (approximately 12 knots) and a constant depth of 25 m. The depth ranges from 30 m to 45 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. The data were bandpass-filtered with cut-off frequencies 750 Hz and 937.5 Hz, then demodulated using a 750 Hz cosine signal and finally down-sampled to a sampling frequency of 375 Hz. In the frequency

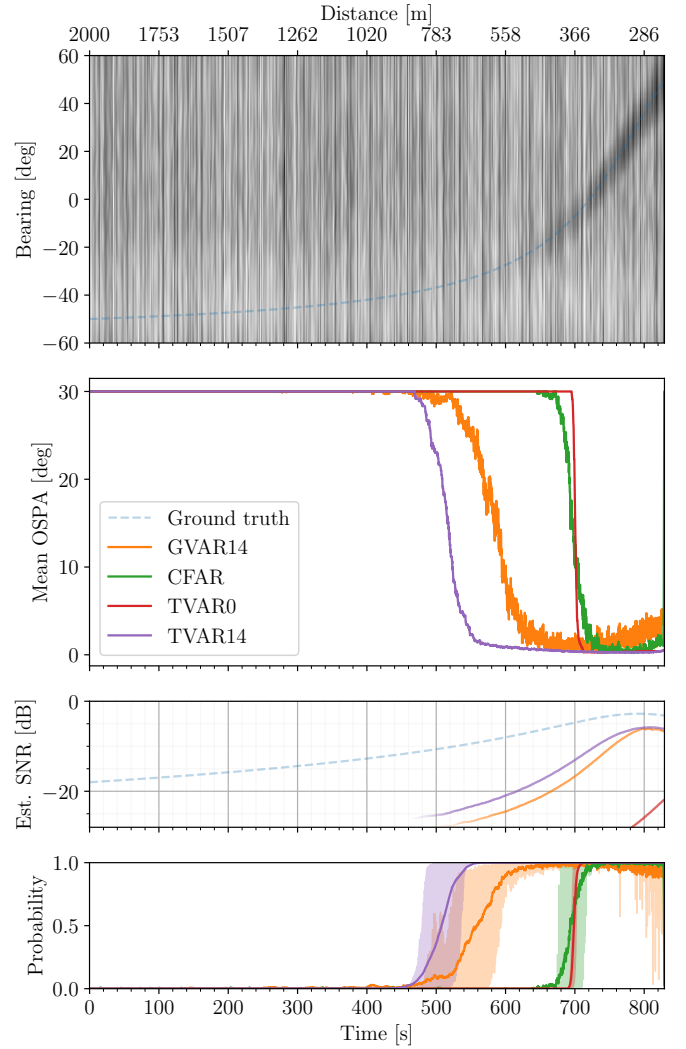


Fig. 6. Results of 100 Monte Carlo runs of the simulated scenario. The top plot shows the BTR of the generated signal for a single run, together with the ground truth bearing in dashed blue. Below the BTR, the OSPA, estimated SNR, and estimated target existence probability are shown. The transparent regions in the estimated target existence probability correspond to the top 90 and bottom 10 quantiles.

band 750 Hz and 937.5 Hz, the received signal can be considered approximately white. This is because the spectrum of the sound generated by the AUV 62 is approximately flat, and the differences in propagation loss across the frequency band are negligible.

*1) Effects of Ambient Noise Modeling:* In Fig. 4, the effect of whitening the hydrophone data in Fig. 2b with a VAR model learned on the data in Fig. 2a, 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 *TVAR14* and *GVAR14* trackers use the same VAR model of order  $p = 14$  for the ambient noise modeling and data whitening, whereas the *TVAR0* tracker uses a  $p = 0$  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



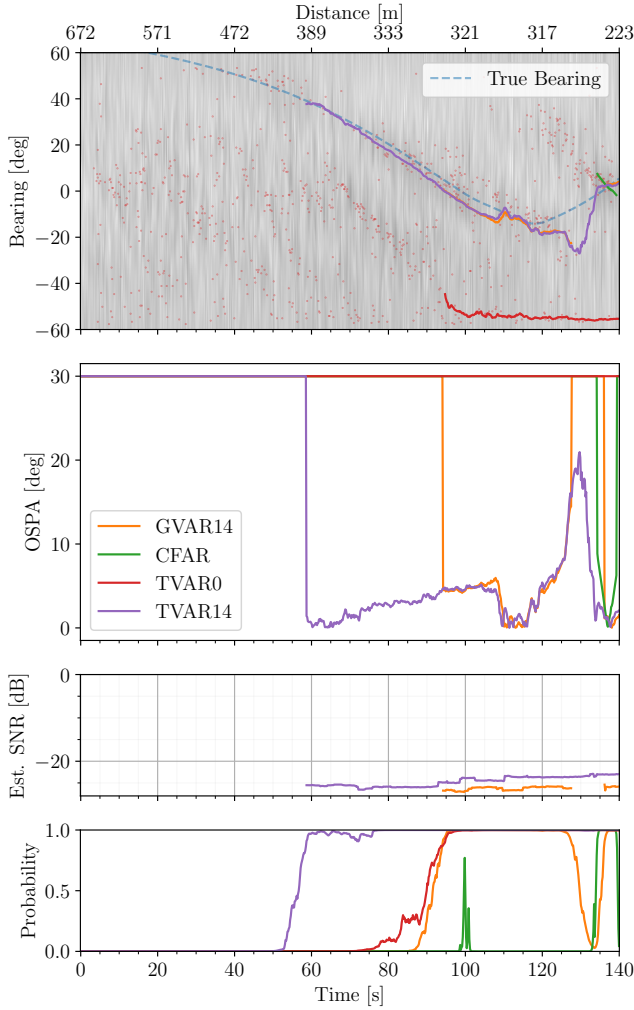


Fig. 7. Results from the real-world scenario. The top plot shows the BTR, along with the truth and estimated bearings. Below the BTR, the OSPA, estimated SNR, and estimated target existence probability are shown.

after the whitening. Specifically, bearings within  $20^\circ$  to  $60^\circ$  exhibit lower signal energy in the BTR of the raw data and in *TVAR0*, compared to *TVAR14* and *GVAR14*. Conversely, the energy in bearings  $-60^\circ$  to  $-20^\circ$  during the timespan 90 s to 110 s is higher in *TVAR0*.

2) *Detection and Tracking Performance*: In Fig. 7, the estimated target track, SNR, target existence probability, and the calculated OSPA for the four trackers are shown. The CFAR tracker detects and starts to track the target at a distance of 250 m. The *TVAR0* tracker detects the target at a distance of 325 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. 4, where the negative bearings exhibited higher beamforming energy due to the absence of temporal whitening. The *GVAR14* tracker detects the target at approximately 325 m and maintains the track for approximately 125 s until the SNR temporarily decreases. However, the method quickly recovers the track.

A common issue across the CFAR, *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 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, significantly 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. Further, 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 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 m to 390 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 increasing the safety of maritime activities.

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 [34] and belief propagation for multi-Bernoulli filters [35] have shown promising results. 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  $\Sigma = \eta H(\psi) H^\top(\psi) + I_{NM}$ . 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

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

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

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

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. Garcia-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. Bossér, 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] 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.
- [27] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Berlin, Germany: Springer, June 2005.
- [28] L. Ljung, *System identification (2nd ed.): theory for the user*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] FMV, *Hydroakustik och sonarteknik för marinen v. 2.2*, 2014.
- [33] 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.
- [34] E. S. Davies and A. F. Garcia-Fernández, “Information Exchange Track-Before-Detect Multi-Bernoulli Filter for Superpositional Sensors,” *IEEE Trans. Signal Process.*, vol. 72, pp. 607–621, 2024.
- [35] M. Liang, T. Kropfreiter, and F. Meyer, “A BP Method for Track-Before-Detect,” *IEEE Signal Process. Lett.*, vol. 30, pp. 1137–1141, 2023.



**Daniel Bossér** received his BSc and MSc degrees in Engineering Mathematics and Computational Science from Chalmers University of Technology, Gothenburg, Sweden in 2018 and 2020, respectively.



**Isaac Skog** (S'09-M'10) received his BSc and MSc degrees in Electrical Engineering from KTH Royal Institute of Technology, Stockholm, Sweden in 2003 and 2005, respectively. In 2010, he received a PhD degree in Signal Processing with a thesis on low-cost navigation systems. In 2009, he spent 5 months with the Mobile Multi-Sensor System research team, University of Calgary, Canada, as a visiting scholar and in 2011 he spent 4 months at the Indian Institute of Science (IISc), Bangalore, India, as a visiting scholar. Between 2010 and 2017 he was a researcher

at KTH Royal Institute of Technology, Stockholm, Sweden. Between 2017 and 2023 he was a assistant professor in automatic control at Linköping University, Sweden. Currently he is an associate professor in communication systems at KTH Royal Institute of Technology, Stockholm, Sweden, and a senior researcher at the Swedish Defence Research Agency (FOI), Stockholm, Sweden.



**Magnus Lundberg Nordenvaad** was born in Luleå, Sweden, in 1973. He received the MSc degree in Computer Engineering from the Luleå University of Technology, Luleå, Sweden, in 1998, and the PhD degree in Signal Processing from the Chalmers University of Technology, Gothenburg, Sweden, in 2003. He has authored around 60 papers in international journals and conferences. His research interests lie in statistical signal processing and how it applies to sonar, digital communication, radar, and process diagnostics.

He has held faculty positions with the Department of Information Technology, Uppsala University, Uppsala, Sweden, and the Department of Computer Science and Electrical Engineering, Luleå University of Technology. He has also held visiting researcher/professor positions at Purdue University, Colorado State University, and the University of Florida. He is currently a Deputy Research Director with the Swedish Defense Research Agency.



**Gustaf Hendeby** (S'04-M'09-SM'17) is Associate Professor and Docent in the division of Automatic Control, Department of Electrical Engineering, Linköping University. He received his MSc in Applied Physics and Electrical Engineering in 2002 and his PhD in Automatic Control in 2008, both from Linköping University. He worked as Senior Researcher at the German Research Center for Artificial Intelligence (DFKI) 2009–2011, and Senior Scientist at Swedish Defense Research Agency (FOI) and held an adjunct Associate Professor position at Linköping University 2011–2015. Dr. Hendeby's main research

interests are stochastic signal processing and sensor fusion with applications to nonlinear problems, target tracking, and simultaneous localization and mapping (SLAM). He has experience of both theoretical analysis as well as implementation aspects.