**Authors' reply**

| | |
|---|---|
| Title: | Broadband Passive Sonar Track-Before-Detect Using Raw Acoustic Data |
| Authors: | Daniel Bossér, Magnus Lundberg Nordenvaad, Gustaf Hendeby, and Isaac Skog |
| Manuscript #: | 2024JOE004926 |

# General comments

Thanks for all the valuable comments. We have addressed all comments to the best of our ability.

# Reviewer 1

## General comments

The paper presents a Bernoulli TBD for detection and tracking (in angle) of an underwater target using a hydrophone array. The paper is well written, and it demonstrates the difficulties in practical application of TBD: the sensitivity to mathematical modelling of received signals. The key points of this paper: (1) t-distribution for modeling the received signal amplitude (in the presence and in the absence of a target); (2) learning the covariance of ambient noise using VAR model. The paper demonstrates the performance improvement both in simulations and using a set of real data.

**Reply:** Thanks for carefully reviewing the manuscript and providing detailed and constructive feedback. It is greatly appreciated! We have done our best to address all your comments.

## Specific comments

1. A color-bar in Fig.1 (a) and 1(b) would be useful

   **Reply:** Fixed.

2. Sentence "Finally $L(z_k|x)$ ... " is unclear (page 3, left-hand column)

   **Reply:** Fixed.

The following questions, question 3 and 4, are answered jointly:

3. Equation (6), definition of $y_k$ does not make sent to me. In particular, I don't follow the subscripts ((k-1)N+1 ...

and

4. Similarly, equation (8a) does not make sense.

   **Reply:** Thank you! Samples are collected at a different rate compared to the tracker update rate. That means $N$ acoustic samples are collected between two time steps $k-1$ and $k$ of the tracker. We have clarified this in Sec. III.A.

5. The choice of letter p for the order of VAR model is unfortunate: it confuses with notation for the PDF.

   **Reply:** Fixed. We replaced $p$ with $\varrho$.

6. Equation (35) is unclear. Where is x on the right-hand side of proportional sign?

   **Reply:** Thank you for raising this unclarity. The target state $x$ consists of $x = \begin{bmatrix}\psi & \dot{\psi} & \eta^{(\mathrm{dB})}\end{bmatrix}$, that is, it is the bearing, bearing change rate, and the SNR in dB. $L(z_{k-1}|x)$ is a function of the $\psi$ and $\eta$.

   We have added an additional sentence in IV.B that to clarify this.

7. Figure 4: why are the SNR estimates so poor? Please comment.

   **Reply:** We believe that there are two potential causes:

   (a) The approximation in the likelihood function.

   (b) The fact that new targets are born at a low SNR to make the tracker less prone to initiate false tracks, causing a bias in the estimate.

   This is now discussed in the manuscript; see Sec. V.C.

8. Did you implement the algorithm in Section II using a particle filter, or some other method?

   **Reply:** We have clarified in Sec. II and in Sec. V that the particle filter implementation from [23] is used to estimate the BRFS state.

9. Multi-target track-before-detect algorithms (2nd last sentence in the paper): in addition to [34], and [35], there is also the following reference: Kim at al. "A Bernoulli Track-Before-Detect Filter for Interacting Targets in Maritime Radar," in IEEE Transactions on Aerospace and Electronic Systems, June 2021.

   **Reply:** We have added the reference.

# Reviewer 2

## General comments

The submitted work presents the application of a Bernoulli track-before-detect filter in passive sonar tracking, utilising different noise densities with data-driven background noise estimation. The results are shown for both simulated and real-world data with a publicly available code and data set.

**Reply:** Thanks for carefully reviewing the manuscript and providing detailed and constructive feedback. It is greatly appreciated! We have done our best to address all your comments.

## Specific comments

1. Which tracker (reference?) was used together with the CFAR detections in the non-TBD case?

   **Reply:** Thank you for highlighting this unclarity. The CFAR tracker uses the same underlying Bernoulli filter as the TkBD-based trackers. The only difference is the measurement model (likelihood function) used. We have clarified this in Sec. II.B and in Sec. V.

Question 2 and 3 are answered jointly:

2. Generally, further explanations/justifications for determining the thresholds are needed. Especially regarding the effect on false tracks.

and

3. Please list the used CFAR parameters and the reasoning behind them. Further, how did changing them affect the results? It seems like the settings of the probability of false alarm were rather low, yet the majority of the detection points are located around the true target trajectory. This should be analysed and clarified for a fair comparison.

   **Reply:** We have done our best to tune the models and thresholds for a fair comparison. All models have been tuned following the same scheme, that is, the trackers' sensitivities have been tuned until there are no longer any apparent false tracks in a dataset where no target target is present. There are several parameters in all trackers that can be tuned for the same effect. Searching for an optimal set of parameters is infeasible. Instead, we supplied the source code for the evaluations that have been done. While we believe we have done a good job at tuning the filters, we are open to the possibility that a better set of parameters can be found, and welcome anyone who wishes to take a look at the source code find these.

   We also want to make it clear that all trackers are scrutinized in the same manner as they all are evaluated on based on the same threshold $\gamma$ on $q_{k|k}$ for false alarms. Naturally, decreasing the threshold $\gamma$ will imply more false tracks if other parameters in the trackers are not adjusted accordingly. We opted to use a threshold of 90 % that could likely have been used by an operator and tuned our filters thereafter.

   Additional parameters that have been used in CFAR, such as the number of guard cells, training cells, and false alarm rate, have been added to Table II.

4. Looking at Fig. 5a, the red dots are not aligned with the grid of the underlying intensity values. Is there another processing e.g., clustering? If so, this must be explained.

   **Reply:** We are clustering directly adjacent cells using the DBScan algorithm. This has been clarified in Sec. III.B. Furthermore, the offset may possibly be explained by errors in the calibration of the array's orientation and geometry. We added a note on this in Sec. V.D.

5. The definition of the SNR with respect to the distance is given in (40). It is however not clear how the noise power relates to that and how it was used to create the data for the simulation. (That might explain the offset in your SNR estimates)

   **Reply:** The trackers are only dependent on the signal-to-noise ratio $\eta$. Thus, upon generating the dataset, the variance of the driving process noise $\vec{w}_k$, and thereby also the variance of background noise $\vec{e}_k$ can be set arbitrarily. For simplicity we used $\mathrm{cov}(\vec{w}_k) = I_M$, i.e., it has a unit variance. The spatiotemporal VAR filter changes the amplitude as it filters $\vec{w}_k$ to construct $\vec{e}_k$. That is why there is an additional estimation step in Alg. 1, line 3, to get the background noise power. Given the background power and SNR $\eta$, we can calculate the power of the target signal as done on line 5 in the algorithm. The background noise $e_k$ could have equivalently been normalized instead with $\vec{e}_k \leftarrow \vec{e}_k/\det(\mathrm{cov}(\vec{e}_k))^{M/2}$, without a change to the results.

   We have added a note to clarify this in Sec. V.C.

6. The argument that whitening should not be applied before using a CFAR detector seems questionable and should be verified.

   **Reply:** This question prompted us to investigate the statement. We did not find a significant improvement in the tracking performance on the simulated dataset but did find that the CFAR tracker was able to detect the target a bit earlier on the sea-trail dataset. We have thus updated Sec. V and Fig. 4 and Fig. 5 in the manuscript to reflect this result, and adjusted the discussion accordingly.

7. In Section IV-B, there should be a conditioning in the equations as well.

   **Reply:** Sorry, but we do not understand what you mean. Please clarify.

8. Line 179, $\mu$ should be mean rather than median.

   **Reply:** Fixed.

9. Line 224/225, check notation. Arrow notation and subscript k vs. n.

   **Reply:** The notation has been corrected.

10. Units are sometimes written as m/s and sometimes as $\mathrm{ms}^{-1}$.

    **Reply:** Fixed.

11. (11) should end with a full stop.

    **Reply:** Fixed.

# Reviewer 3

## General comments

Thank you for the opportunity to review your work. I think the approach is technically sound. I'd just like to see a few modifications made:

**Reply:** Thanks for carefully reviewing the manuscript and providing detailed and constructive feedback. It is greatly appreciated! We have done our best to address all your comments.

## Specific comments

1. could you address why the t distribution was selected? motivated the choice?

   **Reply:** As we tried to use our previous work on real data, we encountered issues using the Gaussian distribution. There were "bursts" in energy in seemingly random bearings that caused the tracker to lose track of the target. We concluded that there must be random energy shifts in the data that are neglected by the model. Thus began our work on heavy-tailed distributions. According to Abraham [12], ocean noise is known to be heavy-tailed in its distribution.

   We have added some additional motivation before eq. (17) on why the t-distribution is used.

2. could you address why only a bearing/bearing-rate state was chosen? the particles could exist in a cartesian space but observability would not be possible w/out the contact maneuver or multipath present. bearing/bearing-rate is not that interesting from a classification perspective. range-rate is, however.

   **Reply:** We limit the analysis to bearings-only for simplicity. As you state, a cartesian state would not be observable given the far-field approximation and the single array that is used. There is nothing that hinders the extension of the method to cartesian space though.

   We have added a note on this in Sec. IV.A.

3. could you specify where you think the approach is most applicable? My guess is harbor or port surveillance.

   **Reply:**

   The method is likely most applicable in shallow water areas with non-isotropic background noise, like an archipelago. We have included this comment in Sec. VI.

4. could you discuss any limitations in performance? For example, high bearing-rate. If this were applied in a non-harbor scenario, high bearing-rate would be experienced as contacts cross over the top of the sensor.

   **Reply:**

- If the motion model does not match the target, the tracker's performance will be compromised.
- The target is assumed to be stationary during one batch. If the target moves fast, its directional energy will likely be "smeared" over several batches. Given that each batch is short in time, it would imply a bearing change rate greater than $5\,^\circ\,\mathrm{s}^{-1}$. We have not investigated this, but we believe that the method may be able to handle this, with the hypothesis that the particle cloud will be more spread out as a result.

A discussion about these limitations is now included in Sec. VI.

5. could you comment on why a conventional beamformer was chosen? In a practical application I'd imagine using a different adaptive beamformer. Was this to isolate the performance of the algorithm?

   **Reply:** Using the conventional beamformer was not by choice, but a result of the modeling of the raw samples. In the single target case with Gaussian or t-distributed noise that is spatially and temporally white noise, the data likelihood can be expressed in terms of the conventional beamformer.

   We have clarified this in Sec. III.C.

6. Could you provide a Table for the algorithm similar to the one for the data generation?

   **Reply:** Given that tabulating the algorithm would require a large set of additional mathematical constants, that may be confused with the notation currently in place in the manuscript, we have found it more suitable to refer to the reference text in [23]. The work by Ristic et.al. in [23] provides a table with pseudocode describing the algorithm implementation.

   We have updated Sec. II to refer to the pseudocode.

7. could you discuss how the particle filter was initialized? What was the prior? I am assuming uniform initialization in bearing over 0-360 but perhaps bearing-rate was limited which would relate to my prior question on how the algorithm handles high bearing-rate

   **Reply:**

   The prior is given by the birth density $b_{k|k-1}(x)$ in Sec. IV.B. During the evaluation, we limit the bearings of newborn targets to avoid the end-fire directions of the array. We have added this condition to the birth model in eq. (35), and state the allowed range of bearings in Sec. V.B.