Hierarchical Multi-Target Tracking in Underwater Environments

Joshua J. Wakefield[1], Stewart G. Haslinger[2], and Jason F. Ralph[1]

[1]Department of Electrical Engineering & Electronics, University of Liverpool
[2]Department of Mathematical Sciences, University of Liverpool

Joshua J. Wakefield, University of Liverpool, Liverpool, L69 7ZF, sgjwakef@liverpool.ac.uk

Abstract: *This paper presents a comparative analysis of tracking filters for passive sonar applications, focusing on scalability under varying signal-to-noise ratio (SNR) conditions. Unlike active sonar, passive systems rely solely on received signals, making tracking particularly challenging due to high noise levels, multipath propagation, and limited measurement information. Existing approaches often apply a single filtering method tailored to specific scenarios; however, no single filter performs optimally across all conditions. While more flexible methods may offer robustness, they typically incur greater computational costs, limiting their practicality for real-time use. Simulation results show that linear Gaussian filters, such as the Kalman filter, are highly effective for high-SNR sources where measurements are reliable and ambiguity is minimal. In these cases, more complex methods offer minimal benefit and can be computationally inefficient. Conversely, sequential Monte Carlo methods, such as particle filters, provide superior performance in low-SNR environments, where non-linearities and sparse detections increase tracking ambiguity at the cost of higher computational load. These findings support a scalable, adaptive filtering strategy in which different filters are used for different sources based on their observed SNR. This approach balances accuracy and efficiency by aligning algorithmic complexity with the difficulty of the tracking task.*

Keywords: *Target tracking, passive sonar, particle filter, kalman filter, state estimation*

1. INTRODUCTION

Passive sonar is a widely used tool for underwater surveillance to detect, track, and monitor the movement of sea vessels, both on the surface and underwater, as well as to observe marine animals and environmental conditions.

Traditional tracking systems often rely on a single filtering method to estimate source trajectories from noisy measurements. Kalman filter variants are commonly employed in underwater tracking, where motion models are approximately linear and noise is Gaussian. However, their performance degrades in the presence of clutter, ambiguous data associations, non-linear measurements, non-Gaussian noise, or multi-modal posteriors [8]. In such conditions, particle filters are often preferred due to their ability to better approximate complex dynamics and distributions [16]. However, this robustness comes at the cost of increased computational complexity, which limits real-time applicability, particularly in systems with power or hardware constraints.

This work supports a hierarchical tracking strategy in which different filters are selected to operate in parallel, with selection informed by each source's observed signal-to-noise ratio (SNR). Rather than proposing a new filter, we analyse the performance of Kalman and particle filters across a range of SNR conditions in two scenarios: *crossing contacts*, where source trajectories intersect, and *close contacts*, where source trajectories remain in close angular proximity for extended periods. Simulation results show that no single method performs optimally across all cases: Kalman filters are well-suited to high-SNR sources with reliable measurements, while particle filters maintain performance in low-SNR settings. By matching filter complexity to tracking difficulty, this approach offers a scalable and resource-efficient solution for multi-target tracking.

2. SIGNAL PROPAGATION AND PROCESSING

We model the signal emitted by a source as a complex sinusoid comprising multiple tonal components, which reflect the mechanical gearing of onboard machinery and form a distinctive narrowband acoustic signature [5]. The source also generates broadband noise, modelled as pink noise (also known as $\frac{1}{f}$ noise), which captures irregular and diverse sounds generated onboard, such as those from human activity, mechanical operations, and other vibrations [3]. Ambient ocean noise, arising from both natural and anthropogenic sources, is modelled as additive white Gaussian noise, consistent with conventional ocean noise characterisations [15].

Acoustic propagation is simulated using Bellhop, a ray-tracing model [12]. For simplicity, the environment is modelled with a flat seabed and a constant ocean depth of 5 km. The sound speed follows the Munk profile, with depth normalised around 1300 m and a surface sound speed of 1500 m/s, consistent with standard deep-ocean conditions [11]. Bellhop provides both transmission loss, which quantifies the reduction in acoustic intensity with range, and ray arrival times. Transmission loss is used to scale the received signal amplitude, while arrival times determine propagation delays and corresponding phase shifts at the hydrophones, which are configured in a static horizontal linear array placed at a depth of 1 km.

The complete signal received at the i-th hydrophone in the array can be expressed as

$$s_i(t) = \left[\sum_{n=1}^{N_t} a_n \exp\left(j2\pi f_n(t + \tau_p + \tau_i) + j\phi_n\right) \right] + \eta_p(t, \tau_p, \tau_i) + \eta_w(t), \tag{1}$$

where N_t is the number of tonal components; a_n, f_n, and ϕ_n are the amplitude, frequency, and

phase of the n-th tonal, respectively; τ_p is the propagation delay from the source to a reference sensor; and τ_i is the relative time delay at the i-th hydrophone due to the array geometry. The terms $\eta_p(\cdot)$ and $\eta_w(\cdot)$ represent the pink and white noise processes, respectively.

To estimate the direction of arrival (DoA) of acoustic signals, we apply a frequency-domain delay-and-sum (DAS) beamformer to the hydrophone data [13]. This method phase-aligns the frequency components of the signals across the array to enhance energy from specific directions while suppressing noise and interference.

Detection is performed in two stages. First, the SNR is computed across discrete bearing angles, obtained by partitioning the angular field of view. A fixed threshold is applied to identify candidate detections. Due to the finite resolution of the beamformer and the presence of sidelobes, strong signals may exceed the threshold in several adjacent bearing angles. To consolidate these into single detections, a peak-finding algorithm identifies local SNR maxima at each time step. The resulting bearing measurements are then associated with tracks over time, enabling the estimation of each source's trajectory and dynamics.

3. TARGET TRACKING

Bayesian target tracking provides a probabilistic framework for recursively estimating the state of a dynamic system given noisy and possibly incomplete measurements. The belief about the system's state is represented by a probability density function (PDF), which is updated as new observations become available through two key steps: prediction and update.

In the prediction step, the prior PDF is propagated forward using the system's motion model, described by the Chapman-Kolmogorov equation [14],

$$p(\mathrm{x}_k \mid \mathrm{z}_{1:k-1}) = \int p(\mathrm{x}_k \mid \mathrm{x}_{k-1}) p(\mathrm{x}_{k-1} \mid \mathrm{z}_{1:k-1}) \, d\mathrm{x}_{k-1}, \tag{2}$$

where x_k and z_k denote the system state and measurement at time step k, respectively. The transition model $p(\mathrm{x}_k \mid \mathrm{x}_{k-1})$ reflects the system dynamics and process noise.

In the update step, the prior is refined using Bayes' rule,

$$p(\mathrm{x}_k \mid \mathrm{z}_{1:k}) = \frac{p(\mathrm{z}_k \mid \mathrm{x}_k) p(\mathrm{x}_k \mid \mathrm{z}_{1:k-1})}{p(\mathrm{z}_k \mid \mathrm{z}_{1:k-1})}, \tag{3}$$

where $p(\mathrm{z}_k \mid \mathrm{x}_k)$ is the likelihood of the measurement given the predicted state, and the denominator normalises the posterior distribution.

In this work, Bayesian target tracking is applied to a passive sonar system. The problem is treated as two-dimensional since all vessels operate on or near the surface relative to the range between the sources and the sensor array, and tracking is performed in the angular domain. The true state of target j at time step k is represented by

$$\mathrm{x}_k^{(j)} = \begin{bmatrix} \theta_k^{(j)} & \dot{\theta}_k^{(j)} \end{bmatrix}^T, \tag{4}$$

where $\theta_k^{(j)}$ denotes the bearing and $\dot{\theta}_k^{(j)}$ the bearing rate of the j-th target. The estimated state $\tilde{\mathrm{x}}_k^{(j)}$ is a filtered approximation of $\mathrm{x}_k^{(j)}$, computed recursively as part of the tracking algorithm.

The target is assumed to follow a nearly constant angular velocity model, where process noise allows small deviations due to unmodelled dynamics such as acceleration,

$$\mathrm{x}_k = \mathbf{f}(\mathrm{x}_{k-1}, \mathrm{v}_{k-1}) = F_k \mathrm{x}_{k-1} + \mathrm{v}_{k-1}, \quad \mathrm{v}_{k-1} \sim \mathcal{N}(0, Q_{k-1}), \tag{5}$$

where F_k is the state transition matrix representing deterministic motion and v_{k-1} is zero-mean Gaussian noise with covariance Q_{k-1}.

In this setup, the measurement provides a noisy observation of the true bearing. Although beamforming is inherently non-linear, it is approximated here by a linear measurement model,

$$\mathrm{z}_k = \mathbf{h}(\mathrm{x}_k, \mathrm{n}_k) = H_k \mathrm{x}_k + \mathrm{n}_k, \quad \mathrm{n}_k \sim \mathcal{N}(0, R_k), \tag{6}$$

where $H_k = \begin{bmatrix} 1 & 0 \end{bmatrix}$ extracts the bearing component from the state, and n_k is zero-mean Gaussian noise with variance R_k, representing sensor and environmental uncertainties. All angular operations are performed with appropriate wrapping to the interval $(-\pi, \pi]$.

The Kalman filter estimates the target's state $\tilde{\mathrm{x}}_k^{(j)}$ and its associated covariance $\tilde{P}_k^{(j)}$ through recursive prediction and update steps [14]. In the prediction step, the motion model $\mathbf{f}(\cdot)$ is applied without process noise to propagate the state estimate, while the effect of process noise is incorporated in the evolution of the covariance:

$$\tilde{\mathrm{x}}_k^{(j)} = \mathbf{f}(\tilde{\mathrm{x}}_{k-1}^{(j)}, 0), \quad \tilde{P}_k^{(j)} = F_k \tilde{P}_{k-1}^{(j)} F_k^T + Q_{k-1}. \tag{7}$$

Upon receiving a new measurement z_k, the filter updates the state and covariance as

$$\tilde{\mathrm{x}}_k^{(j)} = \tilde{\mathrm{x}}_k^{(j)} + K_k(\mathrm{z}_k - \mathbf{h}(\tilde{\mathrm{x}}_k^{(j)}, 0)), \quad \tilde{P}_k^{(j)} = (I - K_k H_k)\tilde{P}_k^{(j)}, \tag{8}$$

where the innovation $\mathrm{z}_k - \mathbf{h}(\tilde{\mathrm{x}}_k^{(j)}, 0)$ quantifies the difference between the actual and predicted measurements, and I is the identity matrix. The Kalman gain,

$$K_k = \tilde{P}_k^{(j)} H_k^T (H_k \tilde{P}_k^{(j)} H_k^T + R_k)^{-1}, \tag{9}$$

determines the relative weight of the new measurement compared to the predicted state.

The particle filter approximates the posterior distribution for each target j using a weighted set of samples, or particles, $\{\tilde{\mathrm{x}}_k^{(i,j)}, w_k^{(i,j)}\}_{i=1}^{N_s}$ [2].

Each particle is propagated independently according to the motion model,

$$\tilde{\mathrm{x}}_k^{(i,j)} = \mathbf{f}(\tilde{\mathrm{x}}_{k-1}^{(i,j)}, \mathrm{v}_{k-1}^{(i,j)}), \quad \text{for } i = 1, \ldots, N_s, \tag{10}$$

where N_s is the number of particles. The particle filter represents uncertainty non-parametrically through the distribution of particles in the target's state space, explicitly incorporating process noise during propagation.

Particle weights are updated based on the likelihood of the new measurement given the predicted state,

$$w_k^{(i,j)} = w_{k-1}^{(i,j)} \cdot p(\mathrm{z}_k \mid \mathbf{h}(\tilde{\mathrm{x}}_k^{(i,j)}, 0)), \quad \text{for } i = 1, \ldots, N_s, \tag{11}$$

where $p(\cdot)$ denotes the likelihood function. Particles that better match the measurement receive higher weights. The weights are then normalised to satisfy $\sum_{i=1}^{N_s} w_k^{(i,j)} = 1$.

Over time, many particles may receive negligible weights, a problem known as *particle degeneracy* [6]. To mitigate this, resampling is performed: low-weight particles are discarded, while those with high weights are replicated proportionally to their weights. This step reallocates computational effort toward more probable regions of the state space and helps maintain a representative and diverse particle set.

The ocean environment is inherently noisy and often produces spurious detections, referred to as *clutter*, which introduces additional uncertainty and complexity. To address this, detections are associated with tracks probabilistically using a well-established data association method,

Joint Probabilistic Data Association (JPDA), which jointly considers all valid associations and assigns likelihood-based weights [4].

A valid association assigns each detection to at most one track and requires that the detection lies within a specified confidence region of the predicted measurement, thereby defining a hypothesis. Missed detections, cases where a track is not associated with any measurement, are also accounted for. Such missed detection hypotheses are assigned likelihoods corresponding to the probability of missed detection and are always considered valid.

Let $\mathcal{H}_k$ denote the set of all joint association hypotheses at time step k, and let $\mathcal{H}_k^{(j)}$ be the set of single-target hypotheses induced for target j. For each $h \in \mathcal{H}_k^{(j)}$, the marginal association likelihood is

$$p(h) = \begin{cases} 1 - (p_d^{(j)} \cdot p_g), & \text{if } h \text{ corresponds to a missed detection,} \\ \dfrac{p(\mathrm{z}_k \mid \mathbf{h}(\tilde{\mathrm{x}}_k^{(j)}, 0)) \cdot p_d^{(j)}}{\kappa}, & \text{otherwise,} \end{cases} \tag{12}$$

where $p_d^{(j)}$ is the detection probability for target j, p_g the gating probability, and κ the spatial clutter density [4]. The missed detection case captures both failure to detect and detections falling outside the gate. The valid association case combines the measurement likelihood and detection probability, normalised by clutter density.

The detection probability p_d depends on factors such as sensor characteristics and target observability. For simplicity, we model it as a function of the target's SNR, using a relationship from the receiver operating characteristic curve for Rayleigh-distributed measurements [7]:

$$p_d = p_{\text{fa}}^{\frac{1}{1+\gamma(\mathrm{x}_k)}}, \quad \text{where} \quad \gamma(\mathrm{x}_k) = 10^{\text{SNR}(\mathrm{x}_k)/10}, \tag{13}$$

with p_{fa} denoting the false alarm probability, and $\text{SNR}(\mathrm{x}_k)$ the target's SNR in decibels at reception [1]. This formulation captures the intuition that stronger signals are more likely to be detected than weaker ones.

The false alarm probability p_{fa} depends on the SNR detection threshold [1]. Assuming a Rayleigh distribution, typical for zero-mean Gaussian noise amplitude detection, it is given by

$$p_{\text{fa}} = \exp\left(-\tfrac{1}{2}\gamma_{\text{thresh}}\right), \quad \text{where} \quad \gamma_{\text{thresh}} = 10^{\text{SNR}_{\text{thresh}}/10}. \tag{14}$$

By considering all hypotheses, weighted by their likelihoods, the filter computes an *expected* state estimate that captures both measurement uncertainty and ambiguity in data association.

For the Kalman filter, each hypothesis contributes a Gaussian component to a mixture model. The final posterior is approximated by reducing the Gaussian mixture to a single Gaussian,

$$\tilde{\mathrm{x}}_k^{(j)} = \sum_{h \in \mathcal{H}_k^{(j)}} p(h) \cdot \tilde{\mathrm{x}}_k^{(h,j)}, \quad \tilde{P}_k^{(j)} = \sum_{h \in \mathcal{H}_k^{(j)}} p(h) \cdot \left(\tilde{P}_k^{(h,j)} + (\tilde{\mathrm{x}}_k^{(h,j)} - \tilde{\mathrm{x}}_k^{(j)})(\tilde{\mathrm{x}}_k^{(h,j)} - \tilde{\mathrm{x}}_k^{(j)})^T \right), \tag{15}$$

where $\tilde{\mathrm{x}}_k^{(h,j)}$ and $\tilde{P}_k^{(h,j)}$ denote state and covariance estimates for target j under hypothesis h.

In the case of the particle filter, the expected estimate is computed by updating each particle's weight through a weighted sum over hypotheses,

$$w_k^{(i,j)} = \sum_{h \in \mathcal{H}_k^{(j)}} p(h) \cdot w_k^{(i,h,j)}, \quad \text{for } i = 1, \ldots, N_s. \tag{16}$$

where $w_k^{(i,h,j)}$ is the weight of particle i under hypothesis h for target j. Resampling is performed only after computing the combined weights, not for each hypothesis. This produces the final particle set $\{\tilde{\mathrm{x}}_k^{(i,j)}, w_k^{(i,j)}\}_{i=1}^{N_s}$ and implements the *expected likelihood particle filter* (ELPF) [9].

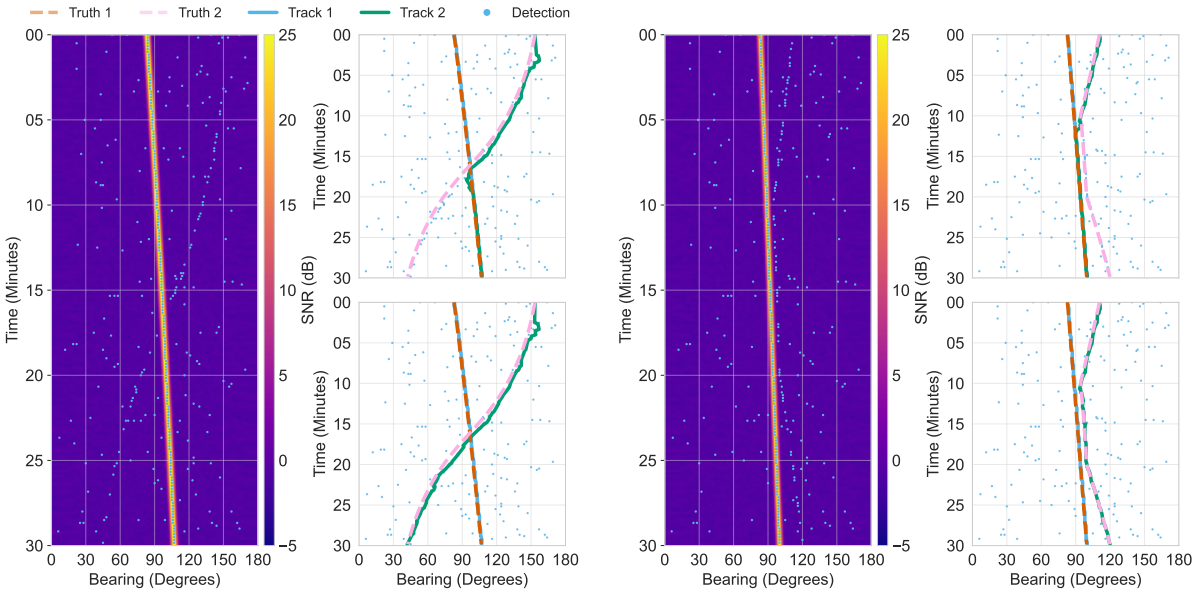
Although both filters use the same Bayesian principles, their computational demands differ. The Kalman filter efficiently updates the state and covariance once per track each time step. The particle filter propagates and updates many particles, then resamples to prevent degeneracy. These steps increase computational load, especially as particle count grows, limiting scalability.

4. RESULTS

Tracking performance is evaluated in two scenarios, both with sources located North of the sensor array's broadside. In the crossing contacts case, Source 1, modelled after a loud modern cargo ship [10], is 40 km away, while Source 2 is quieter and 10 km away; both travel on intersecting paths at 10 m/s. In the close contacts scenario, both sources are approximately 60 km away and remain in close angular proximity for extended periods, challenging the filter's ability to discriminate overlapping measurements. Performance is assessed over 25 SNR values from –10 dB to 2 dB in 0.5 dB steps, with 100 Monte Carlo trials per SNR. SNR is controlled by varying ambient noise between 68 dB and 80 dB, with a fixed detection threshold of 0 dB. Signal parameters are summarised in Table 1, and example trials are shown in Figure 1.

Parameter	Units	Source 1	Source 2
Source Level	dB re 1 µPa at 1 m	[168, 155, 158]	[110, 118]
Frequency	Hz	[97, 72, 580]	[650, 370]
Phase	Degrees	[–112, –151, 86]	[15, 40]

Table 1: Signal tonal parameters for the two sources used in the example tracking scenarios.



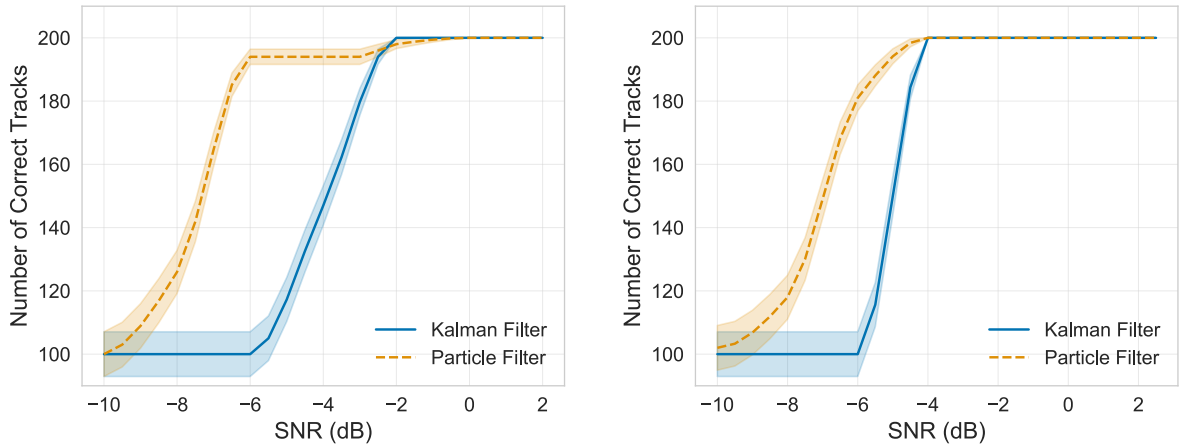
(a) Crossing contacts. *(b) Close contacts.*

Figure 1: Two scenarios showing beamformer output (left), Kalman filter tracks (upper right), and particle filter tracks (lower right). Source/Truth 2 has a SNR of –6 dB.

As shown in Figure 2, the Kalman filter's performance declines sharply below –2 dB in the crossing contacts scenario and below –4 dB in the close contacts case, failing to track the weaker

source at –6 dB and below. The particle filter remains effective down to –10 dB, showing greater robustness at low SNR, while the Kalman filter performs optimally at higher SNRs.

A closer analysis reveals that under low-SNR conditions, the Kalman filter often produces two tracks that follow the same strong source, neglecting the weaker contact entirely. This occurs because the Kalman filter, being a unimodal estimator, tends to converge on the most consistent measurement stream. Degradation of the particle filter's performance is typically due to ambiguity in measurement-to-track association, leading to tracks that intermittently swap between sources. This reflects the particle filter's sensitivity to multi-modal uncertainty.



(a) Crossing contacts. *(b) Close contacts.*

Figure 2: Tracking accuracy across SNR values, showing the number of correctly tracked sources. A track is correct if it remains within 5 degrees of the ground truth bearing. Shaded regions show ± 1 *standard deviation over 100 trials.*

In practice, a SNR threshold may serve as a decision boundary for selecting between Kalman and particle filters in real-time systems. This concept naturally extends to a scalable, hierarchical tracking framework, where multiple filters are available and selected adaptively based on the operating environment. By leveraging the strengths of different algorithms, such a system can maintain robust tracking while optimising computational efficiency across varying conditions.

5. CONCLUSIONS

A comparative analysis of Bayesian tracking filters for passive sonar in bearings-only scenarios has been presented, focusing on how signal-to-noise ratio (SNR) affects filter performance. This work promotes a scalable, hierarchical approach to multi-target tracking. Results from simulations demonstrate that while Kalman filters offer efficient tracking under high-SNR conditions, particle filters are more robust in low-SNR environments where non-linearities and ambiguities are more pronounced. These findings suggest that no single filter is universally optimal across all operating conditions. Instead, identifying a practical decision boundary, such as an SNR threshold, for selecting between filters enables hierarchical multi-target filtering strategies that balance tracking accuracy and computational efficiency by matching algorithm complexity to the difficulty of the tracking task, supporting practical real-time sonar applications.

6. ACKNOWLEDGEMENTS

This work was supported by the EPSRC Centre for Doctoral Training in Distributed Algorithms through a Research Studentship under Grant EP/S023445/1.

REFERENCES

[1] M. A. Ainslie. *Principles of Sonar Performance Modelling*. Springer, 2010.

[2] M Sanjeev Arulampalam et al. "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking". In: *IEEE Transactions on signal processing* 50.2 (2002), pp. 174–188.

[3] Paul T Arveson and David J Vendittis. "Radiated noise characteristics of a modern cargo ship". In: *The Journal of the Acoustical Society of America* 107 (2000), pp. 118–129.

[4] Yaakov Bar-Shalom, Fred Daum, and Jim Huang. "The probabilistic data association filter". In: *IEEE Control Systems Magazine* 29.6 (2009), pp. 82–100.

[5] Richard M Heitmeyer, Stephen C Wales, and Lisa A Pflug. "Shipping noise predictions: capabilities and limitations". In: *Marine Technology Society Journal* 37.4 (2003), pp. 54–65.

[6] Genshiro Kitagawa. "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models". In: *Journal of Computational and Graphical Statistics* 5 (1996), pp. 1–25.

[7] A Vincent van Leijen, Willem L van Norden, and Fok Bolderheij. "Unification of radar and sonar coverage modeling". In: *2009 12th International Conference on Information Fusion*. IEEE. 2009, pp. 1673–1678.

[8] Junhai Luo, Ying Han, and Liying Fan. "Underwater acoustic target tracking: A review". In: *Sensors* 18.1 (2018), p. 112.

[9] Alan Marrs, Simon Maskell, and Yaakov Bar-Shalom. "Expected likelihood for tracking in clutter with particle filters". In: *Signal and Data Processing of Small Targets 2002*. Vol. 4728. SPIE. 2002, pp. 230–239.

[10] Megan F McKenna et al. "Underwater radiated noise from modern commercial ships". In: *The Journal of the Acoustical Society of America* 131 (2012), pp. 92–103.

[11] Walter H Munk. "Sound channel in an exponentially stratified ocean, with application to SOFAR". In: *The Journal of the Acoustical Society of America* 55 (1974), pp. 220–226.

[12] Michael B Porter. "The bellhop manual and user's guide: Preliminary draft". In: *Heat, Light, and Sound Research, Inc., CA, USA, Tech. Rep* 260 (2011).

[13] Barry D Van Veen and Kevin M Buckley. "Beamforming: A versatile approach to spatial filtering". In: *IEEE assp magazine* 5.2 (1988), pp. 4–24.

[14] Greg Welch, Gary Bishop, et al. "An introduction to the Kalman filter". In: (1995).

[15] Gordon M Wenz. "Acoustic ambient noise in the ocean: spectra and sources". In: *The journal of the acoustical society of America* 34.12 (1962), pp. 1936–1956.

[16] Wei Yi et al. "Particle filtering based track-before-detect method for passive array sonar systems". In: *Signal Processing* 165 (2019), pp. 303–314.