

# AN ACOUSTIC MULTIPLE TARGET TRACKER

*Volkan Cevher and James H. McClellan*

Georgia Institute of Technology  
Atlanta, GA 30332-0250

## ABSTRACT

We propose a particle filter acoustic tracker to track multiple maneuvering targets using a state space formulation. The state update is formulated through a locally linear motion model. The observations are a batch of direction-of-arrival (DOA) estimates at various frequencies. The data likelihood incorporates the possibility of missing data as well as spurious DOA observations. By imposing smoothness constraints on the target motion, the particle filter is able to avoid the data association problems. To make the filter computationally efficient, a proposal strategy based on approximating the full posterior is employed. Computer simulations are presented to show the performance of the algorithm.

## 1. INTRODUCTION

Tracking the bearing angles of multiple maneuvering targets using acoustic arrays is a classical problem with many challenging aspects. In the literature, the problem is usually formulated in terms of state-space models where target direction-of-arrivals (DOA) are related to the acoustic microphone outputs through some observation equation and the updates on the state (which may include variables other than DOA alone) are explained by locally linear motion models [1–3]. The performance of these algorithms relies heavily on how accurate these models represent the observed natural phenomena.

One important observation model is the far-field narrow-band observation model where the array response is related to a source with a constant narrow-band frequency at DOA,  $\theta$ , through some steering vectors under isotropic and non-dispersive medium assumptions [4]. This model possesses a well-understood mathematical structure and has a significant impact on the DOA estimation problem since it has led to many practical DOA estimation techniques (e.g., MUSIC, MVDR, eigenvalue beamformers, etc.). However, the estimation performance of the algorithms using this model deteriorate when the target signals exhibit wide-band structure, or when rapid target motion spreads the array spatial spectrum. The complexity of this observation model increases significantly as some of the assumptions in the model are relaxed to more accurately represent reality, rendering the tracking algorithms analytically intractable.

The presence of multiple targets also increases the tracking complexity because of data association issues, where a mechanism is needed, in effect, to sort the received data for each target. The association problem is handled in many ways: (i) probabilistic data

association methods estimate the states by summing over all the association hypothesis weighted by the probabilities obtained by the likelihood [5], (ii) smoothness assumptions on the target (motion) states allows a natural ordering of the data [6], (iii) computationally costly ML/EM methods use the likelihood functions to search for a global maximum, and (iv) nearest neighbor methods provide easy heuristics to perform measurement updates. Most of these methods use the mean and the covariance approximation on the sufficient statistics for the state, which may be estimated with a Kalman filter; however, for nonlinear state-spaces with general noise assumptions, Monté-Carlo methods should be used to adequately capture the dynamic, possibly multi-modal, statistics.

A particle filter naturally accommodates solutions for state-space problems where the observations arrive in sequence. The state probability density function is represented by discrete state samples (particles) distributed according to the underlying distribution (as explained by the state-space) either directly or by proper weighting [7]. Hence, the filter can approximate any statistics of the distribution arbitrarily accurately by increasing the number of particles with proven convergence results. In the particle filtering framework, the data association problem is undertaken implicitly by the state-space model interaction; however, to increase the efficiency of the algorithm, various methods are proposed, such as the partitioning approach [2], or other Bayesian approaches [8].

In this paper, a particle filter tracker is presented for multiple target DOA tracking using a constant velocity motion model on the targets. Instead of directly using the signals as explained by the narrow-band model, the algorithm uses a sufficient statistics for the state vector: a batch of DOA estimates  $\mathbf{y}_{t,f} = \{y_{t+m\tau,f}(j_m)\}_{m=0}^{M-1}$  out of a beamformer appropriate for the local frequency characteristics of the target signals. The filter state consists of the DOA  $\theta_k(t)$ , heading direction  $\phi_k(t)$ , and (logarithm of) velocity over range  $Q_k(t) = \log v_k/r_k(t)$  for each target  $k$  where the total number of targets  $K$  is assumed known. The state is reported at regular time intervals of  $T = M\tau$ . The batch DOAs may also have depend on frequency  $f$  (e.g., DOAs calculated at different narrow-band frequencies), which is suitable in specific tracking scenarios. The number of DOA estimates  $j_m$  at time  $t + m\tau$  is also modelled as time-dependent.

The observations are assumed normally distributed around the true DOA tracks with constant missing data probability and clutter rate. To increase the efficiency of the algorithm, the filter uses an approximation of the full posterior to generate particles where a robust Newton search method is used to approximate the data-likelihood. Data association is handled automatically by the particle filter by imposing the smoothness constraints on the target motion. Lastly, all DOA observations within the batch are assumed locally stationary to avoid angle spread due to target motion since the batch sampling period  $\tau$  is assumed small.

---

Prepared through collaborative participation in the Advanced Sensors Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-02-0008.

Organization of the paper is as follows. Section 2 is intended to give a brief overview of the motion model and provide the needed data-likelihood expressions. The particle filter details are covered in Sect. 3 where a constrained Newton method is introduced to approximate the filter proposal function. Finally, a representative set of simulation results is given in Sect. 4.

## 2. STATE-SPACE FORMULATION

The particle filter state vector  $\mathbf{x}_k(t)$  consists of the concatenation of the individual motion vectors  $x_k(t) \triangleq [\theta_k(t), Q_k(t), \phi_k(t)]^T$  as  $\mathbf{x}_t = [x_1^T(t), x_2^T(t), \dots, x_K^T(t)]^T$ . Parameters  $\theta_k(t)$  and  $\phi_k(t)$  are measured c.w. w.r.t the  $x$ -axis. The state update equation can be derived from the geometry imposed by the assumed constant velocity motion and is nonlinear [1, 3]:

$$x_k(t+T) = h_T(x_k(t)) + u_k(t) \quad (1)$$

where  $u_k(t) \sim \mathcal{N}(0, \Sigma_u)$  with  $\Sigma_u = \text{diag}\{\sigma_{\theta,k}^2, \sigma_{q,k}^2, \sigma_{\phi,k}^2\}$  and  $h_T(x_k(t)) =$

$$\left[ \begin{array}{c} \tan^{-1} \left\{ \frac{\sin \theta_k(t) + T \exp Q_k(t) \sin \phi_k(t)}{\cos \theta_k(t) + T \exp Q_k(t) \cos \phi_k(t)} \right\} \\ Q_k(t) - \frac{1}{2} \log \left\{ 1 + 2T \exp Q_k(t) \cos(\theta_k(t) - \phi_k(t)) + T^2 \exp 2Q_k(t) \right\} \\ \phi_k(t) \end{array} \right] \quad (2)$$

The observations  $\mathbf{y}_{t,f} = \{y_{t+m\tau,f}(j_m)\}_{m=0}^{M-1}$  consist of all the DOA estimates out of a beamformer at each batch index  $m$ . Hence, the acoustic data of length  $T$  is segmented into  $M$  segments of length  $\tau$ . These segments are processed by a proper beamformer based on the temporal frequency structure of the signals to calculate possible DOA estimates. This procedure may be repeated  $F$  times for narrow-band signals for each frequency indexed by  $f$ . Note that only the peak locations are kept in the beamformer power pattern. Moreover, the peak values, indexed by  $j_m$ , need not be ordered or associated with the previous index  $m-1$  and the number of peaks to retain can even be time-dependent.

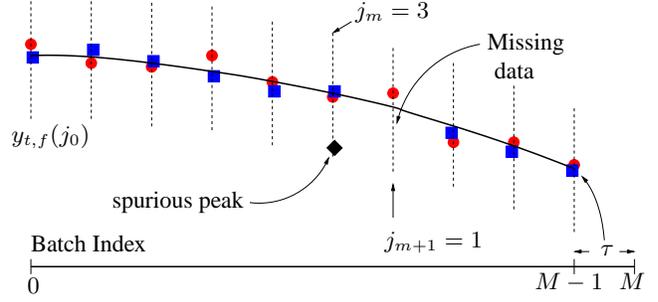
Figure 1 visualizes the observation model. It is assumed that the batch of DOAs,  $\mathbf{y}_{t,f}$ , form a normally distributed cloud around the true target DOA tracks with a constant missing data probability  $\kappa^1$ ; and may have spurious peaks Poisson distributed with rate  $\lambda$ . The variance of DOAs,  $\sigma^2$ , is assumed constant and may be estimated using the DOAs from the previous estimation period in conjunction with the ML estimation techniques on how the specific beamformer works. For example, for an ML estimator using the narrow-band model, there is a formula that calculates the additive array noise variance given in [1]. The array noise then can be related to an expected DOA noise through the formulas in [10].

This specific observation model is very similar to the one used in active contour image tracking problems [9]. It can be shown that the data likelihood given the state under the assumptions described above can be written as  $p(\mathbf{y}_t|\mathbf{x}_t) \propto$

$$\prod_f \prod_k \prod_m \left\{ 1 + \frac{1}{\sqrt{2\pi\kappa\lambda}} \sum_{j_m} \exp -\frac{(h_{m\tau}^\theta(\mathbf{x}_t) - y_{t+m\tau,f}(j_m))^2}{2\sigma^2} \right\} \quad (3)$$

where  $\mathbf{y}_t = \{\mathbf{y}_{t,f}\}_{f=1}^F$  denotes the cumulative DOA data calculated in time interval  $[t, t+T)$  and  $h^\theta$  refers to the DOA compo-

<sup>1</sup>This model assumes that only one DOA at each  $f$  belongs to the target or the target is missed. By assuming that the probability of the true measurement being in the observed data is equal for each  $j_m$ , a constant data miss probability  $\kappa$  may be assumed [9].



**Fig. 1.** The circles and squares are the DOA estimates at different frequencies calculated using the acoustic data received during a period of length  $\tau$ . Given the observations  $\mathbf{y}_{t,f}$ , the objective of the tracker is to determine the state  $x_k(t)$  which completely parameterizes the solid curve.

nents. Equation 3 also assumes that each observation through the dimension  $f$  are independent. If there is a known internal correlation structure, a joint density can be formulated to replace (3).

## 3. PARTICLE FILTER DETAILS

Given the state-space description of a problem, the particle filtering solutions are well defined. The efficiency of the algorithm depends on the proposal functions that determine the random support of the particles to be *properly* weighted for estimation. In this paper, a proposal function, denoted as  $g(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1})$ , is derived to approximate the target posterior density directly:

$$g(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1}) \approx p(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1}) \propto p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1}) \quad (4)$$

where  $p(\mathbf{x}_t|\mathbf{x}_{t-1}) \sim \mathcal{N}(f_T(\mathbf{x}_{t-1}), \Sigma_u)$  and  $p(\mathbf{y}_t|\mathbf{x}_t)$  is given by (3). Moreover, the proportionality in (4) is independent from the current state  $\mathbf{x}_t$ . This approximation, in effect, moves the particle stream towards high probability regions of the posterior so that more particles survive the final resampling step, producing better future states as the system evolves [7].

The posterior density should be approximated such that the resulting proposal function is as close to the posterior as possible and, at the same time, is easy to sample from. Hence, various Gaussian approximations to the full posteriors are commonly used in the literature [11]. In our case, we first approximate the data-likelihood by a Gaussian so that the proposal function will also be Gaussian, therefore achieving the objective. Hence, a mean  $\mu_y$  and a covariance  $\Sigma_y$  should be first determined from the observed data  $\mathbf{y}_t$ . Then, by using the Gaussian parameters of the state update, an analytical relation for the proposal function is given.

The mode of  $p(\mathbf{y}_t|\mathbf{x}_t)$ , denoted as  $\mathbf{x}_M$ , is a good candidate for the parameter  $\mu_y$ . Then, the Hessian  $H$  of (3) at the mode can also be used as the covariance estimate:  $\Sigma_y = H^{-1}$ . To calculate these parameters, a Newton search algorithm can be used on the negative log-likelihood of the data [1]. However, the resulting analytical relations have numerical sensitivity issues. Hence, as an alternative, we propose to use the following cost function to

determine the mode  $\mathbf{x}_M$ :

$$J = - \sum_f \sum_k \sum_m \sum_{j_m} \exp \left\{ - \frac{(h_{m\tau}^\theta(\mathbf{x}_t) - y_{t+m\tau, f}(j_m))^2}{2\sigma^2} \right\} + \frac{1}{2} (\mathbf{x}_t - \mathbf{x}_0)^T \Sigma^{-1} (\mathbf{x}_t - \mathbf{x}_0) \quad (5)$$

This cost function consists of two terms: the first term has the same minima as the negative log-likelihood function of the data distribution; and the second term is a regularization term forcing the solution  $\mathbf{x}_M$  to lie close to some vector  $\mathbf{x}_0$  w.r.t. some weighted distance measure  $\Sigma$ .

Nominally, the mode should be within the particle cloud coming from the previous iteration after being propagated through the state update. Hence, an easy way to determine the mode would be to choose the particle best explaining the current data set (denoted as  $\mathbf{x}_0$  in (5)). Unfortunately, when the targets maneuver,  $\mathbf{x}_0$  may fall outside actual data observations. This necessitates a correction accomplished by a Newton search algorithm to determine the actual mode  $\mathbf{x}_M$ .

Note that the cost function without the regularization term only depends on the angle distances. The gradients in that case may lead to physically infeasible (motion) changes in the parameters  $Q(t)$  and  $\phi(t)$  to account for fractional angle errors while determining  $\mathbf{x}_M$ . Hence, the regularization is introduced to prevent this issue by constraining the solution space to lie in the  $\Sigma$  neighborhood of  $\mathbf{x}_0$ , and, at the same time, imposing the smoothness of the target motion. The parameter  $\Sigma$  also bounds the covariance of the data-likelihood approximation.

If we define  $G = \partial J / \partial \mathbf{x}$  and  $\hat{H} = \partial^2 J / \partial \mathbf{x} \partial \mathbf{x}^T$ , the Newton recursion is given by the familiar expression:  $\mathbf{x}_M^{l+1} = \mathbf{x}_M^l - \mu_l \hat{H}_l^{-1} G_l$  and  $\mathbf{x}_M = \mathbf{x}_0$ . The step size  $\mu_l$  should be decreased adaptively making sure that the cost function is always decreasing. Although time-consuming, it is straightforward to derive analytical expressions for  $G$  and  $\hat{H}$  (similar calculations can be found in [1].) They are omitted here due to lack of space.

Even with the available analytical relations, the calculation of the Hessian still poses problems. If the Hessian of (5) is directly calculated from the exact formulas, it is possible to show that the resulting expression for  $\hat{H}$  is not guaranteed to be positive definite and modifications are necessary to make the Newton correction  $\hat{H}^{-1} G$  effective at each iteration<sup>2</sup>. Hence, while calculating the final expression of the Hessian, the terms including second order derivatives are neglected from the analytical formula. In this case, the Hessian is a function of the outer product of the gradient, and it is possible to prove that it is positive definite.

After the Gaussian approximation to the data-likelihood described above (note that  $\Sigma_y = H^{-1} \approx \hat{H}^{-1}$ ), the final expression for the proposal function to be used in the particle filter is given by

$$g(\mathbf{x}_t | \mathbf{y}_t, \mathbf{x}_{t-1}) \sim \mathcal{N}(\mu_g, \Sigma_g), \quad \text{where} \quad (6)$$

$$\Sigma_g = (\Sigma_y^{-1} + \Sigma_u^{-1})^{-1}, \quad \mu = \Sigma_g (\Sigma_y^{-1} \mathbf{x}_M + \Sigma_u^{-1} h_T(\mathbf{x}_{t-1})). \quad (7)$$

Then, the particle filter incremental weights are given by

$$u^{(i)} = \frac{p(\mathbf{y}_t | \mathbf{x}_t^{(i)}) p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)})}{g(\mathbf{x}_t^{(i)} | \mathbf{y}_t, \mathbf{x}_{t-1}^{(i)})} \quad (8)$$

The implementation of the filter used in the simulations also has a resampling (with replacement) stage after estimation.

<sup>2</sup>The same issue applies to the the negative log-likelihood of (3).

**Table 1. Simulation Parameters**

Fig.	N	T	M	F	$\sigma$	$\sigma_{u,\theta}$	$\sigma_{u,Q}$	$\sigma_{u,\phi}$
2, 4	100	1	10, 20	1	2°	1°	0.05	10°
3	100	1	10	2	3°	1°	0.05	10°

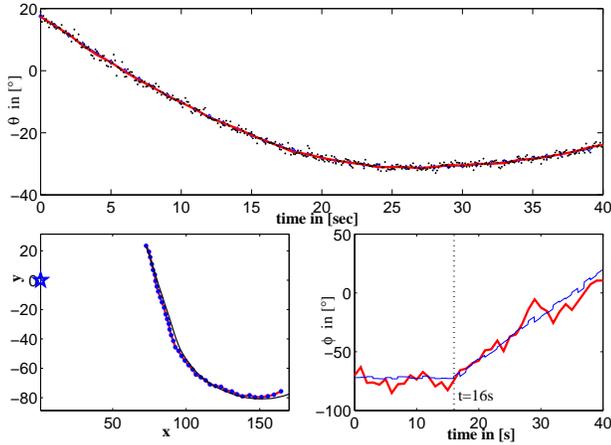
## 4. SIMULATIONS

The objective of the simulations is to demonstrate the performance of the algorithm, using data generated according to the assumed models. For comparison, in the last example, we also generated acoustic data based on the narrow-band array model and calculated the observations using the minimum variance distortionless response (MVDR) beamformer. To initialize the trackers, we used a particle cloud with the correct mean, and a covariance of  $\Sigma_0 = \text{diag}\{(2^\circ)^2, 0.1^2, (4^\circ)^2\}$ . For the Newton algorithm that approximates the data-likelihood for the particle proposal, an initial step size of 0.1 is used, which is decreased adaptively until 1000 iterations are reached. In the simulations, we used  $\Sigma = \sqrt{2}\Sigma_u$ .

Simulation parameters are summarized in Table 1. Figure 2 demonstrates a single target tracking scenario. The observed DOAs are Gaussian distributed around the true DOA track with variance  $(2^\circ)^2$ . The filter does a good job of catching the target as it maneuvers through the large heading process noise. Figure 3 shows a much more difficult scenario for the tracker where two independent layers of DOA estimates are given each with the correct mean and a variance of  $(3^\circ)^2$ . There is a small bias in the filter DOA estimates between  $t = 4s$  and  $t = 6s$  where the targets are crossing. This bias is, in part, also due to the target maneuvers, which start at  $t = 4s$ . The filter maintains the track coherence in this difficult case using the independent frequency observations (when only one of them is present, we observed that the filter can confuse the targets). Although its estimates deteriorate in the region where targets are crossing as well as maneuvering, it locks back on the targets after the transient region.

For the last example (Fig. 4), acoustic data sampled at  $F_s = 1000\text{Hz}$  is generated for two targets each with narrow-band frequencies  $f_1 = 40\text{Hz}$  and  $f_2 = 80\text{Hz}$  using the narrow-band observation model [4]. The filter state is also augmented to include a frequency variable:  $x_k(t) \triangleq [\theta_k(t), Q_k(t), \phi_k(t), f_k(t)]^T$  [3]. Then, Gaussian noise is added to the array data where the noise standard deviation is equal to one-tenth of the sinusoid amplitudes. The microphone array used for the simulation has 15 microphones situated uniformly on a circle such that the minimum inter-microphone distance is .45 times the wavelength of the second signal. The acoustic data, then, is run through the MVDR beamformer where the highest 3 peaks are picked with no particular order at each batch index. 50 data samples are used to calculate each DOA, hence  $M = 20$ . DOAs are used to output per state vector at a period of  $T = 1s$ .

Note that in this case, ignoring the model dependent angle bias of  $(\approx 0.2^\circ)$ , the calculated DOA's are distributed around the true DOA track with much less variance  $(\approx (0.4^\circ)^2)$  than the assumed variance  $\sigma^2 = (2^\circ)^2$ . Hence, the actual data-likelihood is narrower than what is assumed by the tracker. This has the impact of decreasing the number of effective particles that contribute to the estimation accuracy. This is to be expected since the filter is not matched to the data; however, this also demonstrates the robustness of the algorithm under the unmatched prior case.



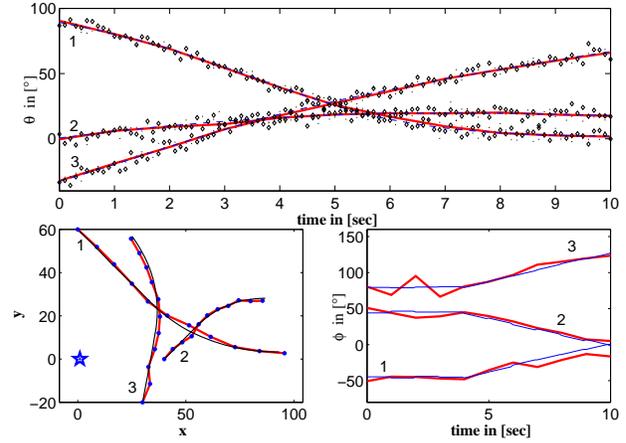
**Fig. 2.** *Top:* Black dots are the DOA observations generated by adding Gaussian noise to the true target DOA track. The filter estimate is indistinguishable from the true DOA track. *Left:* Ground truth vs. filter estimate. The sensor array is shown with the star. The filter track estimate is constructed by using the filter motion outputs with the correct initial target position. *Right:* The filter heading estimates vs. the true target heading.

## 5. CONCLUSIONS

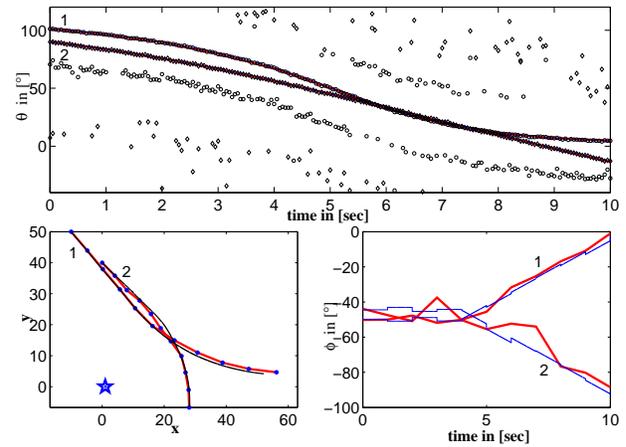
In this paper, a robust acoustic tracker is formulated in a flexible framework that has minimal assumptions on the observations. Even though the filter requires a batch of data to process, it can be implemented for online applications because the data-likelihood approximation can be done as the data is being accumulated. Since the filter uses the Bayesian framework, it can avoid the data association problems commonly encountered in target tracking. The multiple target tracking performance can be further improved by altering the data likelihood so that a confusion probability is assigned to each data while the targets are crossing.

## 6. REFERENCES

- [1] Y. Zhou, P.C. Yip, and H. Leung, "Tracking the direction-of-arrival of multiple moving targets by passive arrays: Algorithm," *IEEE Trans. on Signal Processing*, vol. 47, no. 10, pp. 2655–2666, October 1999.
- [2] M. Orton and W. Fitzgerald, "A Bayesian approach to tracking multiple targets using sensor arrays and particle filters," *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 216–223, February 2002.
- [3] V. Cevher and J. H. McClellan, "General direction-of-arrival tracking with acoustic nodes," *IEEE Trans. on Signal Processing*, vol. 53, no. 1, pp. 1–12, January 2002.
- [4] D.H. Johnson and D.E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall, 1993.
- [5] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*, Academic-Press, 1988.
- [6] R. Karlsson and F. Gustafsson, "Monte Carlo data association for multiple target tracking," in *IEE Target Tracking: Algorithms and Applications*, Netherlands, 16-17 Oct. 2001.



**Fig. 3.** DOAs, represented by diamonds and dots, are generated by independent noise and are input to the filter unsorted.



**Fig. 4.** The filter automatically separates DOAs with respect to the frequency planes defined by  $f_k$  in the state. This way the filter survives the period between times  $t = 5s$  and  $t = 8s$  where the motion vectors for both targets are very similar. Without the help of the frequency association, the filter confuses the targets. Some of the spurious peaks (circles) form a track similar to the true target track due to array spatial aliasing.

- [7] J.S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American Statistical Association*, vol. 93, pp. 1032–1044, September 1998.
- [8] M. Isard and J. MacCormick, "BraMBLE: A Bayesian multiple-blob tracker," in *8<sup>th</sup> International Conference on Computer Vision*, 2001.
- [9] M. Isard and A. Blake, *Active Contours*, Springer, 2000.
- [10] P. Stoica and A. Nehorai, "Music, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. on ASSP*, vol. 37, no. 5, pp. 720–741, May 1989.
- [11] A. Doucet, N. Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, 2001.