Adapting the multi-Bernoulli filter to phased array observations using MUSIC as pseudo-likelihood

Praveen B. Choppala, Paul D. Teal, Marcus R. Frean
School of Engineering and Computer Science
Victoria University of Wellington
New Zealand
Email: {praveen, pault, marcus}@ecs.vuw.ac.nz

***Abstract*—In this paper, we consider Bayesian multi-target tracking using phased array of sensors. Although joint Bayesian filtering is theoretically the optimal approach to multi-target tracking, the method suffers from high computational complexity for large numbers of targets. The PHD and multi-Bernoulli filters avoid this complexity by operating in the dimensionality of a single target space. However, these filters do not possess a mathematical framework to operate directly on signals from the phased sensor array. Therefore, it is necessary for the sensor signals to be first converted to a beamformer image (in the multi-Bernoulli filter) which is then thresholded (in the PHD filter). Moreover, resolving close targets in a beamformer image is difficult. This paper proposes the use of MUSIC as a pseudo-likelihood in the multi-Bernoulli filter. The merits of this proposal are that the multi-Bernoulli filter is able to operate more directly on sensor array signals, and that close targets are effectively resolved. We show the efficacy of our approach in reverberant and noisy environments.**

I. Introduction

Much scientific research and many practical applications depend on the estimation of an unknown entity (called a *target*). A target is characterised by its state, which captures all the information of interest concerning the target, for example, its dynamics. The aim is then to estimate the state of the target from its effects on a data stream, usually consisting of sensor measurements. This paper addresses multiple target tracking (MTT) using a phased array of sensors that receive noisy signals generated by or reflected from the targets [1] at discrete time samples, such as occurs in the tracking of aircraft from RADAR or fish from SONAR. The availability of successive sensor observations allows *sequential Bayes' filtering* [2] to construct a posterior probability distribution function (pdf) of the target state using previous target state information as the prior. However for MTT, the Bayes' filter becomes intractable due to the high-dimensional nature of the posterior [3]. This problem can be overcome by the particle filter (PF) [4], which approximates the posterior by a set of weighted samples ("particles") [5, 6].

If the number of targets is known, then the *Bayesian joint filter* approach is strictly the correct way of casting the multi-target tracking problem. Here, the single target state space is augmented as a single probabilistic entity to form a joint multi-target probability density (JMPD) [7]. In the joint PF (the particle implementation of the Bayesian joint filter), every particle is a concatenation of individual single target hypotheses, i.e., it is a *complete hypothesis* about the locations of *all* the actual targets present. One general problem of the joint filter approach is that it is uncertain as to which observation should be associated with which target — the "data association problem," for which a number of solutions have been proposed [8–11]. Frequently, the number of targets is unknown and must itself be modelled as random variable (see [12–14] for approaches addressing this limitation). Futhermore, the joint filter maximises the likelihood of *only* the joint target state that hypothesises the number of targets and all their states correctly. Consequently, in the joint PF, the information contained in partially correct particles — those that hypothesise only a few targets correctly, or those that hypothesise all the targets correctly but have additional wrong hypotheses — is completely lost. In typical applications, most particles are only partially correct. The major limitation of joint filtering, however, is the *exponential increase* in its complexity [15] for increasing numbers of targets. This resource complexity continues to be a challenge to the research community [16].

A class of Bayesian sub-optimal MTT filters are the random finite set (RFS) filters [17]: the probability hypothesis density (PHD) [18] filter and the multi-target multi-Bernoulli (MeMBer) filter [2, 19]. These filters can also be implemented using particles. The PHD filter operates by approximating the JMPD by its first moment while the MeMBer filter [20] approximates the posterior using its multi-Bernoulli RFS parameters. The fundamental advantage of these filters is that they *operate in the dimensionality of a single target space* and thereby avoid the computational complexity and data association problems of the joint filter. However, since the filters are sub-optimal, they could be flawed in many ways, for example, in high noise scenarios. The major challenge, nonetheless, is that the measurement feed to the filters is a finite set containing noisy target state estimates, thereby *impeding the direct use of phased array data.* This drawback is only partially overcome when the MeMBer filter has been extended to image observations [21, 22]. To date, the RFS filters do not provide a suitable mathematical formalism to evaluate the likelihood function for phased array data, because these filters operate in the dimensionality of a single target, while the phased array data requires a joint state hypothesis to evaluate the likelihood. In the PHD filter, this problem is avoided by first converting the array data into an image [23, 24] by means of a single target beamformer or FFT [25], and then obtaining the finite set measurements using pre-processing, such as segmentation and thresholding [26]. The MeMBer filter for image data does not require image pre-processing, but still necessitates the conversion of the array data into a image. Closely related work in [27, 28] deals with super-positional sensors that generate a "single measurement" from multiple targets. However, this single measurement was considered to be a point estimate of the targets rather than a signal; implying that the array data has to be converted to a beamformer image. These intermediary conversion processes cause substantial loss of information contained in the phased array sensor signals, and could be detrimental for real time applications [29] at high noise conditions. In addition, it becomes difficult to resolve close targets in beamformer images.

The motivating rationale of this paper is to find a function similar to the PHD, which will act as a likelihood function while having the dimensionality of the state of a single target. The MUltiple SIgnal Classification (MUSIC) pseudo-spectrum [30] is such a function. MUSIC is one of what are called "super-resolution techniques" and provides a suitable framework to feed the array signals more directly to the MeMBer filter. Here we propose to use MUSIC as the pseudo-likelihood in the MeMBer filter for image observations. The advantage of this proposition is that filtering can now be conducted more directly on the sensor signals. Moreover, close targets are better resolved by virtue of the super-resolution property of MUSIC. Through a simulated study, we validate our approach in accurately tracking multiple appearing/disappearing targets in an acoustic sensing environment. We also present a detailed study of the effects of MUSIC on the posterior pdf.

The rest of the paper is organised as follows. In section II we set the notation and describe the Bayes' joint filter. In section III, we outline the MeMBer filter for image observations. After stating the problem in section IV, we present our proposal in section V. We then present our simulation results in section VI and conclude in section VII.

II. Bayesian inference

In this section, we introduce the notation and the mathematical formulations for Bayesian target tracking. The state of a target x_k at time k may be characterised, for example, by its position and velocity in a two-dimensional space as

$$\mathrm{x}_k = [x, v_x]^T \tag{1}$$

where x and v_x are the position and velocity. $(.)^T$ denotes matrix transpose. The process and the observation models are described as

$$\mathrm{x}_k = f_k(\mathrm{x}_{k-1}, q_{k-1}) \tag{2}$$

$$z_k = h_k(\mathrm{x}_k, r_k) \tag{3}$$

where the nonlinear functions f_k and h_k are the state evolution and measurement functions respectively. q_{k-1} and r_k are the process and observation noise sequences respectively.

Target inference is achieved by estimating the state of a target x_k at time k using the sensor data $z_{1:k}$ received until the kth time step. The sequential Bayes' filter facilitates this estimation by using the previous posterior pdf $p(\mathrm{x}_{k-1}|z_{1:k-1})$ to construct $p(\mathrm{x}_k|z_{1:k})$. In the joint approach of Bayesian MTT, x_k is replaced by the joint target state X_k which is a concatenation of the individual target states

$$\mathrm{X}_k = [\mathrm{x}_k^1, \mathrm{x}_k^2, ..., \mathrm{x}_k^N]^T \tag{4}$$

where N is the number of targets present at time k. The single target posterior pdf $p(\mathrm{x}_k|z_{1:k})$ is replaced by the joint multi-target posterior $p(\mathrm{X}_k|z_{1:k})$. The Bayes' joint filter recurses in two stages: a) prediction, and b) update. The prediction of the state from $k-1$ to k is given by

$$p(\mathrm{X}_k|z_{1:k-1}) = \int p(\mathrm{X}_k|\mathrm{X}_{k-1})p(\mathrm{X}_{k-1}|z_{1:k-1})d\mathrm{X}_{k-1} \tag{5}$$

Once the new sensor data z_k is obtained, Bayes' rule is used to update the prediction as

$$p(\mathrm{X}_k|z_{1:k}) = \frac{p(z_k|\mathrm{X}_k)p(\mathrm{X}_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \tag{6}$$

The recurrence of the prediction and correction processes is the property that forms the basis for the Bayes' recursive filter. An estimate of X_k can then be extracted from $p(\mathrm{X}_k|z_{1:k})$ using any familiar Bayesian state estimators.

III. Multi-Bernoulli Bayesian approximation

In this section, we present the multi-Bernoulli approximation to the Bayes' filter [19, 20, 31]. A Bernoulli set X has a probability $1-\epsilon$ of being a null set, and has a probability ϵ of containing a single element x that is distributed according to a pdf p. A multi-Bernoulli RFS X can be considered as union of a fixed number of independent Bernoulli sets, such that

$$\mathrm{X} = \bigcup_{m=1}^{M} \mathrm{X}^m \tag{7}$$

where the mth Bernoulli set is described by its two parameters; the existence probability ϵ^m and the pdf $p(\mathrm{x}^m)$. The MeMBer filter recursively propogates these Bernoulli parameters and hence is considered as a parameterised approximation to the sequential Bayes' filter. The JMPD at time k can then be approximated as

$$p(\mathrm{X}_k|z_{1:k}) \approx \{\epsilon_k^m, p(\mathrm{x}_k^m)\}_{m=1}^{M_k} \tag{8}$$

and the corresponding PHD is

$$D(\{\mathrm{x}_k^1, ...\mathrm{x}_k^{M_k}\}|z_{1:k}) = \sum_{m=1}^{M_k} \epsilon_k^m p(\mathrm{x}_k^m) \tag{9}$$

In the following subsection, we will describe the operation of the MeMBer filter for image observations.

A. MeMBer filter for image observations – Vo et al. [21]

If the multi-Bernoulli parameters that describe the multi-target posterior at time $k-1$ are known and given by

$$p(\mathrm{X}_{k-1}|z_{1:k-1}) \approx \{\epsilon_{k-1}^m, p(\mathrm{x}_{k-1}^m)\}_{m=1}^{M_{k-1}} \tag{10}$$

then (5) can be approximated [21] as

$$p(\mathrm{X}_k|z_{1:k-1}) \approx \{\epsilon_{P,k}^m, p(\mathrm{x}_{P,k}^m)\}_{m=1}^{M_{k-1}} \cup \{\epsilon_{B,k}^m, p(\mathrm{x}_{B,k}^m)\}_{m=1}^{M_{B,k}} \tag{11}$$

where

$$\epsilon_{P,k}^m = \epsilon_{k-1}^m \int p_{\mathrm{S}}(\mathrm{x}_{k-1}^m)p(\mathrm{x}_{k-1}^m)d\mathrm{x}_{k-1}^m \tag{12}$$

$$p(\mathrm{x}_{P,k}^m) = \frac{\int p_{\mathrm{S}}(\mathrm{x}_{k-1}^m)p(\mathrm{x}_k^m|\mathrm{x}_{k-1}^m)p(\mathrm{x}_{k-1}^m)d\mathrm{x}_{k-1}^m}{\int p_{\mathrm{S}}(\mathrm{x}_{k-1}^m)p(\mathrm{x}_{k-1}^m)d\mathrm{x}_{k-1}^m} \tag{13}$$

where $p_{\mathrm{S}}(\mathrm{x}_{k-1}^m)$ is the survival probability of the mth target state x^m in the multi-Bernoulli set at time $k-1$. The set $\{\epsilon_{B,k}^m, p(\mathrm{x}_{B,k}^m)\}_{m=1}^{M_{B,k}}$ contains the multi-Bernoulli parameters for the $M_{B,k}$ newborn targets at time k. $M_{B,k}$ is usually modelled as Poisson distributed with mean λ_B.

Once the new image observation z_k and the predicted multi-Bernoulli parameters $\{\hat{\epsilon}_k^m, p(\hat{\mathrm{x}}_k^m)\}_{m=1}^{M_k}$ are obtained, the JMPD in (6) can be approximated [21] as

$$p(\mathrm{X}_k|z_{1:k}) \approx \{\epsilon_k^m, p(\mathrm{x}_k^m)\}_{m=1}^{M_k} \tag{14}$$

where

$$\epsilon_k^m = \frac{\hat{\epsilon}_k^m \int l(z_k|\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)d\hat{\mathrm{x}}_k^m}{1-\hat{\epsilon}_k^m + \hat{\epsilon}_k^m \int l(z_k|\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)d\hat{\mathrm{x}}_k^m} \tag{15}$$

$$p(\mathrm{x}_k^m) = \frac{l(z_k|\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)}{\int l(z_k|\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)d\hat{\mathrm{x}}_k^m} \tag{16}$$

Using the observation model described in Ch. 11 of [3], the likelihood ratio of the mth hypothesis (see (3) of [21]) is

$$l(z_k|\hat{\mathrm{x}}_k) \triangleq \prod_{t\in T} \frac{\mathcal{N}(g_k^t(\hat{\mathrm{x}}_k), \sigma^2)}{\mathcal{N}(0,\sigma^2)} \tag{17}$$

where $\mathcal{N}(a,b)$ is a normal distribution with mean a and variance b. T includes the high illumination regions of target activity obtained by thresholding the image z_k using an illumination threshold I_l. t indexes the set of pixels that fall within T. From (11.19) in Ch. 11 of [3], (17) can be written (up to a normalising constant) as

$$l(z_k|\hat{\mathrm{x}}_k) = \prod_{t\in T} \exp\Bigg(-\frac{g_k^t(\hat{\mathrm{x}}_k)(g_k^t(\hat{\mathrm{x}}_k) - 2z_k^t)}{2\sigma^2}\Bigg) \tag{18}$$

where σ^2 is the noise variance at each pixel. $g_k^t(\hat{\mathrm{x}}_k)$ is the intensity contribution of a point target hypothesis $\hat{\mathrm{x}}_k$ at the tth pixel. The hypothesis measure $g_k(\hat{\mathrm{x}}_k)$ (which is a image) is obtained, as from (3), according to

$$g_k(\hat{\mathrm{x}}_k) = h_k(\hat{\mathrm{x}}_k) \tag{19}$$

If the existence probability ϵ_k^m of a hypothesised track is less than a threshold $\mathcal{T}_{\text{prune}}$, the track is eliminated. If the distance between the estimates of two targets is closer than a threshold $\mathcal{T}_{\text{merge}}$, the tracks are merged by combining their Bernoulli parameters [2].

In the particle implementation of the MeMBer filter, for the multi-target posterior approximation in (14), each probability density $\{p(\mathrm{x}_k^m)\}_{m=1}^{M_k}$ is approximated using a set of weighted particles $\{\mathrm{x}_k^{m,i}, w_k^{m,i}\}_{i=1}^{I}$, where i is the particle index and I is the total number of particles. The formulation for the particle prediction, the weight update and the existence probability update can be found in [21, 22]. Resampling strategies for the PF can be found in [32, 33].

IV. Problem Statement

The goal of this paper is to implement a particle based MeMBer filter more directly on the signals generated by the phased sensor array [1]. Assume an array of M sensors. Let there be N targets and the nth target at location ζ_k^n generates a signal s_k^n. In this paper, we assume the signals are narrowband, and of frequency ω. Assuming near field conditions, the sensor model can be summarised as

$$z_k = A_k s_k + r_k \tag{20}$$

where z_k is the vector of sensor signals, s_k is the signal vector and the noise vector $r_k \sim \mathcal{CN}(0,R)$. The $M \times N$ steering matrix

$$A_k = [a_k^1, ..., a_k^N] \tag{21}$$

whose n-th column is given by

$$a_k^n = \left[\frac{e^{-j\ w\ \frac{d_{\zeta_k^n,1}}{e}}}{d_{\zeta_k^n,1}},\ \frac{e^{-j\ w\ \frac{d_{\zeta_k^n,2}}{e}}}{d_{\zeta_k^n,2}},\ ...,\ \frac{e^{-j\ w\ \frac{d_{\zeta_k^n,M}}{e}}}{d_{\zeta_k^n,M}}\right]^T \tag{22}$$

describes the delay and attenuation s_k^n undergoes in reaching the mth sensor. e is the propagation velocity and $d_{\zeta_k^n,m}$ is the distance between the nth target and the mth sensor. For multiple targets, signals are added at the sensors. The likelihood equation for this model, assuming normal sensor noise of covariance R, is

$$p(z_k|A_k,s_k) = \frac{1}{\pi^M\ \det(R)}\exp(-(z_k - A_k s_k)^H\ R^{-1}\ (z_k - A_k s_k)) \tag{23}$$

The PHD and MeMBer filters do not possess a framework to evaluate (23). This non-provision can be bypassed using image models [24] that induce considerable information loss. Moreover, the MeMBer filter suffers from high computation load in its particle implementation since the evaluation of (18) requires construction of the beamformer image g_k for each particle and each target hypothesis in the multi-Bernoulli RFS.

Our proposal is motivated from the fact that MUSIC, that allows the particles to operate in the dimension of a single target space, provides a framework to feed the sensor signals more directly to the MeMBer filter. Moreover, the super resolution property of the MUSIC aids in separating very close targets. The objective now is to design a particle based MeMBer filter using MUSIC to localise and track the targets at $\zeta_k^1, ..., \zeta_k^N$ using the z_k from (20).

V. Multiple Signal Classification

A. Classical MUSIC

Here, we will first describe how the non-Bayesian MUSIC [30] can be used for target localisation, and then present our idea of using it in the MeMBer filter. Consider we have received z_k using (20). Assuming that the signals and noise are uncorrelated, the spatial correlation matrix can be given by

$$\tilde{R}_k = E\{A_k\ s_k\ (A_k\ s_k)^H\} + E\{v_k\ v_k^H\} \tag{24}$$

where $(.)^H$ denotes Hermitian transpose. After performing eigendecomposition of $\tilde{R}_k$, the eigenvectors corresponding to the smallest $M - \hat{N}$ eigenvalues are represented as

$$U_{\text{n}} = [u_1, ..., u_{M-\hat{N}}] \tag{25}$$

U_{n} contains the noise eigenvectors and its columns span the noise sub-space. Ideally, $\hat{N}$ should be equal to the actual number of targets N. If any location $\hat{\zeta}_k$ in the field of view is hypothesised to be a target, its steering vector $a_k^{\hat{\zeta}}$ can be formed using (22) and the MUSIC pseudo-spectrum value of $\hat{\zeta}_k$ is given by

$$S_{\text{MU}}(\hat{\zeta}_k) = \frac{1}{(a_k^{\hat{\zeta}})^H\ U_{\text{n}}\ U_{\text{n}}^H\ a_k^{\hat{\zeta}}} \tag{26}$$

It is known that the covariance matrix and the eigenvalues are sufficient statistics for estimating the target location from the phased array data [34, 35].

B. MUSIC as a pseudo-likelihood in the MeMBer filter

In this paper, *we propose to use the MUSIC value as a pseudo-likelihood of a particle.* Hence, (6), for a single target posterior, is now written as

$$p(\mathrm{x}_k|z_{1:k}) = \frac{S_{\mathrm{MU}}(\mathrm{x}_k)p(\mathrm{x}_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \tag{27}$$

In the MeMBer filter implementation, we propose to evaluate (15) and (16) according to

$$\epsilon_k^m = \frac{\hat{\epsilon}_k^i \int S_{\mathrm{MU}}(\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)d\hat{\mathrm{x}}_k^m}{1-\hat{\epsilon}_k^m + \hat{\epsilon}_k^m \int S_{\mathrm{MU}}(\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)d\hat{\mathrm{x}}_k^m} \tag{28}$$

$$p(\mathrm{x}_k^m) = \frac{S_{\mathrm{MU}}(\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)}{\int S_{\mathrm{MU}}(\hat{\mathrm{x}}_k^m)p(\hat{\mathrm{x}}_k^m)d\hat{\mathrm{x}}_k^m} \tag{29}$$

$p(\mathrm{X}_k|z_{1:k})$ which is then obtained from (14) is largest for those particles that are close to the *target(s)* and smallest for those particles that are far from the *target(s)*. Since we ideally require $\hat{N} = N$ while evaluating (25), we propose to use the MMSE-based-MDL algorithm [36] that provides an almost accurate estimate $\hat{N}$ of N. However, it is known that the MUSIC algorithm does not catastrophically fail if N is incorrectly estimated [37].

The fundamental contribution of this proposition is that, while the MeMBer filter in its originally proposed form requires the sensor observation z_k in (3) to be an image for evaluating (18), our proposed technique allows (3) to be expressed directly as the phased array observation model in (20). This is because the likelihood function $p(z_k|\hat{\mathrm{x}}_k^m)$, which is now chosen to be $S_{\mathrm{MU}}(\hat{\mathrm{x}}_k^m)$ is evaluated according to (26) which requires the observation to be phased array sensor signals. Consequently, this proposition will overcome the information loss problem from which the MeMBer and PHD filters suffer. However, inasmuch as $S_{\mathrm{MU}}(\hat{\mathrm{x}}_k^m)$ is not truly a likelihood, $p(\mathrm{X}_k|z_{1:k})$ is not truly a multi-target posterior state distribution, but possesses the attributes necessary for tracking the mode of the posterior which is often an adequate estimate of the target state [38]. A few other pseudo-likelihood functions that have been successfully used in PF based target tracking can be found in [39, 40]. Our proposal also aids in significantly reducing the computational cost — while the MeMBer filter requires the formation of complete beamformer image to evaluate (18) for each particle and each hypothesis, i.e., $M_k I$ images for every time sample, our technique requires the evaluation of only $M_k I$ MUSIC values, each consisting of a single number.

VI. Evaluation

All simulations in this paper consider tracking acoustic sources in a two-dimensional environment which includes 3 echo paths with 1 second reverberation time for 60 dB decay. The targets are constrained to move with position 0m to $x_{\max}$m in a line located 10m from another parallel line containing 5 sensors at positions 13m, 35m, 63m, 81m and 91m. The state of the target x_k is defined as in (1). The targets are assumed to maneuver according to a constant velocity model.

In Fig. 1, we demonstrate the validity of the proposed technique for tracking single/multiple, appearing/disappearing and crossing targets. It can be seen that the filter quickly locks on and tracks the targets with high accuracy. If a target disappears, the existence probability of the Bernoulli RFS goes low, thereby eliminating the track. The use of the Markov transition prior as the proposal distribution is known to be sub-optimal and the performance of the proposed scheme could be enhanced using an improved importance distribution. Nevertheless, it can be seen that the track initialisation and deletion are conducted very effectively.

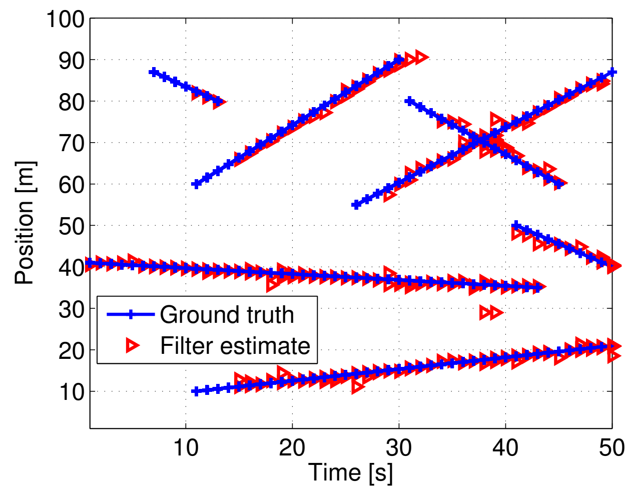


Fig. 1: Filter estimates compared to the ground truth. The observation space is a one dimensional grid of $x_{\max} = 100$m width. The number of particles that approximate the pdf of each element in the multi-Bernoulli RFS is $I = 1000$. $f = 4$KHz and $\nu = 1500$m/s. The SNR is 30dB. The Poisson mean of the number of newborn targets is $\lambda_{\mathrm{B}} = 10$. The probability of target survival (chosen to be constant) is $p_{\mathrm{S}} = 0.8$. The pruning threshold is $\mathcal{T}_{\text{prune}} = 0.6$. The merging threshold is $\mathcal{T}_{\text{merge}} = 1$m. The filter runs for 50 time steps.

For the multi-target scenario described in Fig. 1, Fig. 2 shows the MMSE-based-MDL estimate $\hat{N}$ of the number of targets used to evaluate (25). It can be observed that the

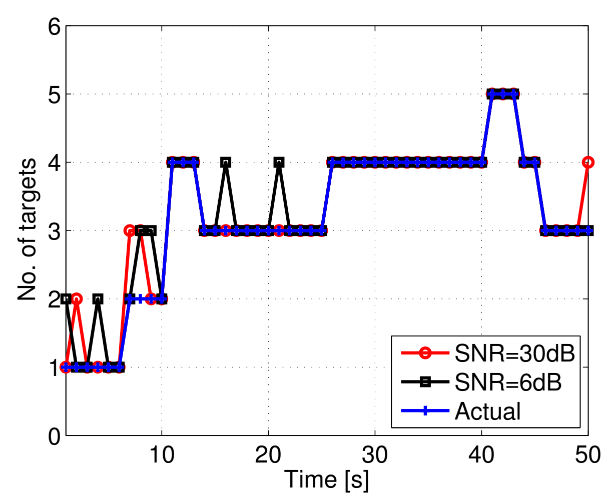


Fig. 2: The MMSE-based-MDL estimate of the number of targets, versus time.

technique provides an accurate estimate of N (even at low SNR levels).

Secondly, we demonstrate the ability of the proposed technique to effectively resolve close targets. For this, we use the scenario shown in Fig. 3, in which, two targets cross each other twice. This motion is realistic for real-world scenarios where targets often move in straight lines with constant velocity and occasionally make an abrupt maneuver [41].

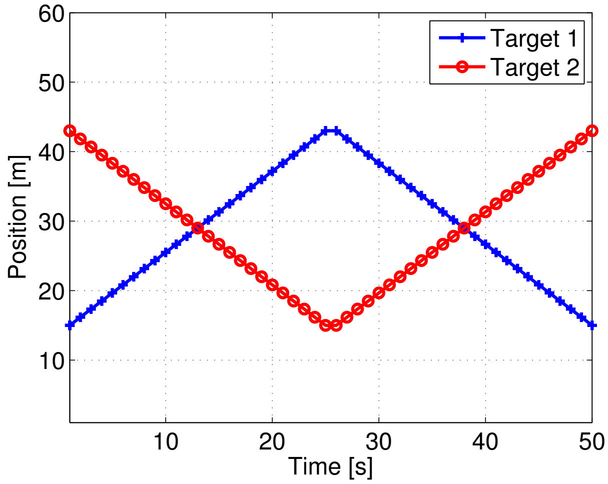
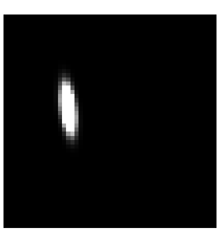
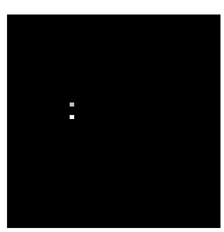


Fig. 3: The ground truth of the two targets used to obtain the evaluation results in Fig. 5, Fig. 6, Fig. 7, Fig. 9, Fig. 10 and Fig. 13. The target positions coincide exactly at $k = 13$ and at $k = 38$.

In view of the fact that the PHD and the MeMBer filters cannot operate directly on array data, for comparison with the proposed method we first convert the sensor acoustic signals into the minimum variance distortionless response (MVDR) beamformer image which is the observation for the MeMBer filter. For the PHD filter, the image is further thresholded to generate the finite set measurement feed. Our proposed MeMBer filter, however, operates more directly on the acoustic signals. The super resolution property of MUSIC aids our filter in resolving close targets and keeping the tracks separate for a longer duration. Fig. 4 illustrates the resolution property of MUSIC when compared with a MVDR beamformer image.



(a) Beamformer image.



(b) MUSIC image.

Fig. 4: Beamformer and MUSIC outputs with two closely spaced targets in a two dimensional environment.

Fig. 5, Fig. 6 and Fig. 7 show the particle evolution (for the ground truth in Fig. 3) for the proposed MeMBer, the MeMBer for image and the PHD filters respectively. It can be seen that the proposed filter resolves close targets better than the others. Before the first crossing at $k = 13$, it can be seen that while the PHD and the MeMBer filters keep the targets separate only until the 6th and 9th time samples respectively, the proposed MeMBer filter keeps the targets separate until the 11th time step. After the second crossing at $k = 38$, the PHD and MeMBer filters resolve the targets only after $k = 46$ and $k = 44$ respectively, and the proposed MeMBer filter separates the targets from the 40th time sample.

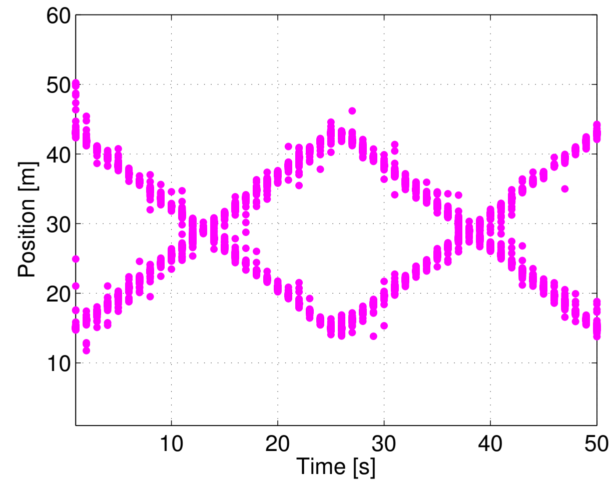


Fig. 5: Evolution of the particles with time in the proposed MeMBer filter. The specifications mentioned in the caption of Fig. 1 apply to this figure also, except that the observation space is now a one dimensional grid of $x_{\max} = 60$m width.

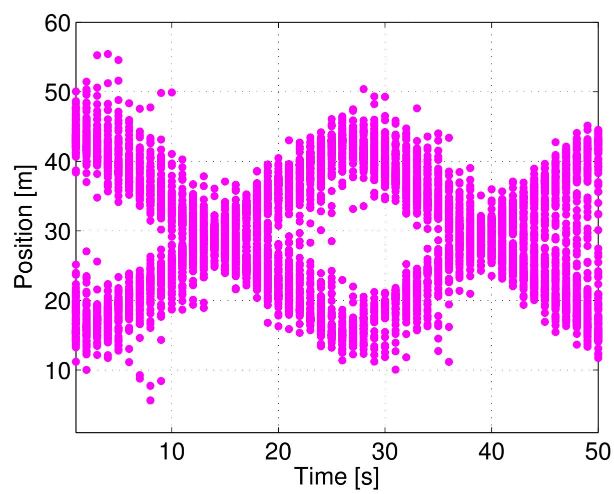


Fig. 6: Evolution of the particles with time in the particle based MeMBer filter (for image data) [21]. The system specifications described in the caption of Fig. 5 apply to this figure also, except that the knowledge about the number of targets is assumed to be known. The illumination threshold is $I_t = 0$ so that the entire image is fed to the filter.

To further illustrate the filters' resolving ability, we evaluate the normalised PHD function for the proposed MUSIC MeMBer and MeMBer filters (from (9)) and the PHD filter. Fig. 8 shows this PHD corresponding to the ground truth in Fig. 3 at the 11th time step. The proposed method successfully resolves the two targets while the other filters merge them inextricably (as is also seen in Fig. 5, Fig. 6 and Fig. 7).

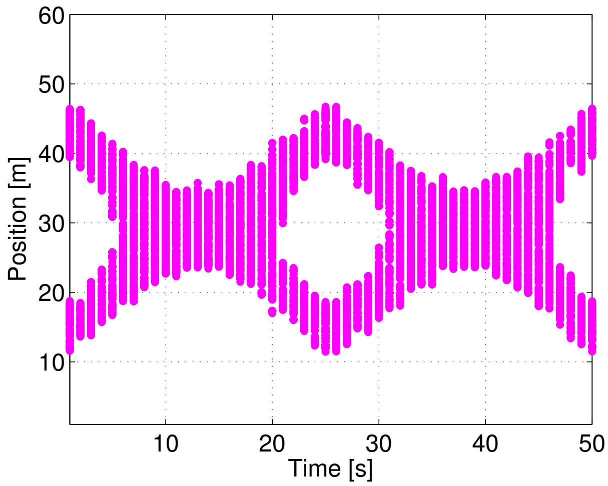


Fig. 7: Evolution of the particles with time in the particle based PHD filter [42]. The system specifications described in the caption of Fig. 5 apply to this figure also, except that the knowledge about the number of targets is assumed to be known. The detection probability is $p_{\mathrm{D}} = 0.95$. Target birth is not considered. Thresholding is done at 70% of the maximum peak in the beamformer image. 1000 particles are used. To separate the targets, we use clustering as proposed in [43].

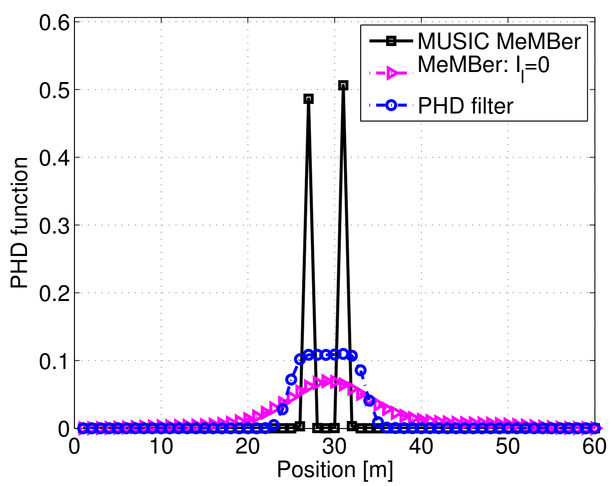


Fig. 8: Normalised PHD function of the MUSIC MeMBer, MeMBer and PHD filters at the 11th time step.

Thirdly, we test the accuracy of the filters (using the ground truth shown in Fig. 3) for varying SNR levels and Fig. 9 indicates that the proposed technique is very accurate. Its performance advantage as against the PHD and the MeMBer filters can be attributed to the more direct use of the acoustic signals. At low SNR, spurious peaks in the beamformer image degrade the MeMBer filter performance (error increases with increasing I_l). These unwanted peaks, in turn, increase the false alarms in the PHD filter observation set and adversely effects its performance. The joint filter operates using the true likelihood function in (23) and performs better than the others.

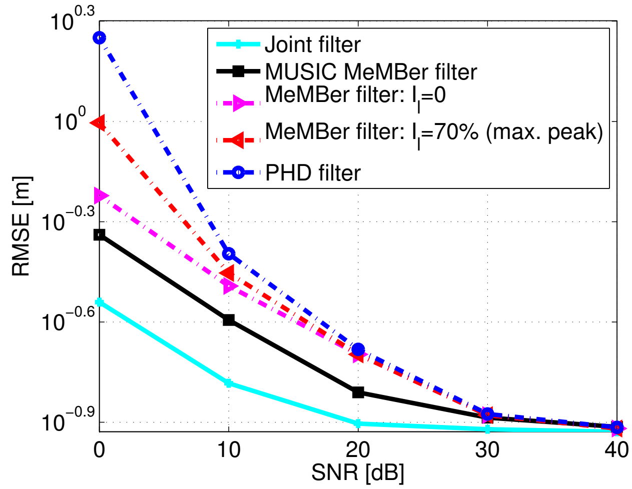


Fig. 9: rmse versus SNR. The specifications described in the caption of Fig. 5, Fig. 6 and Fig. 7 apply to this figure also. For MeMBer and PHD filters, the thresholding is done at 70% of the highest peak in the beamformer image. For all the filters, we assume knowledge about the number of targets. The results are averaged over 20 trials, each having 50 time steps.

Fourthly, we show in Fig. 10, the complexity of the filters (again using the ground truth in Fig. 3) by measuring the computational time to evaluate one complete time step for varying numbers of particles. It can be observed that our technique requires fewer computational resources than the MeMBer filter operating on the image data. This is because the MeMBer filter needs to construct $2I$ images, each having 60 pixels for the evaluation of (18) while the proposed technique requires only $2I$ MUSIC values. In the particle based PHD filter, the majority of the time is consumed while evaluating the beamformer image and hence the filter will be slower than the others when $I \lessapprox x_{\max}$; $x_{\max} = 60$m is width of the surveillance region. The joint filter uses I particles of double the dimension, and its complexity becomes unfavourable for larger numbers of targets.

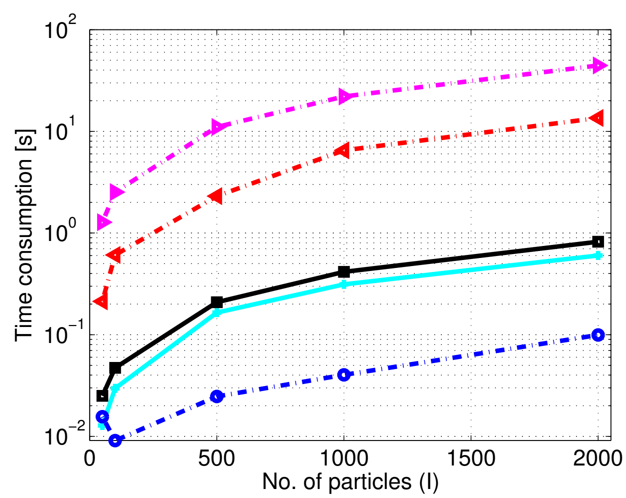


Fig. 10: Time consumption for one time step versus I. The specifications described in the caption of Fig. 9 apply to this figure also. The legend of Fig. 9 applies to this figure also.

Finally, given that we use MUSIC as the pseudo-likelihood of a particle, it is crucial to investigate the faithfulness of MUSIC in accurately representing the *true* likelihood. Fig. 11 shows the true likelihood and the MUSIC pseudo-likelihood evaluated over the entire 60m grid in the presence of a single target. It can be observed that the MUSIC pseudo-spectrum has a much narrower peak at the target location than that of the true likelihood function, and hence will provide a good estimate about the target location. The corresponding cumulative distribution functions (CDFs), shown in Fig. 12, indicate that the narrow peak in MUSIC pseudo-spectrum causes its CDF to be truncated near the tails of the CDF of the maximum-likelihood. This is shown using double arrows. Hence, although MUSIC pseudo-likelihood provides a good estimate of the mode of the target, it does not provide accurate information about the uncertainty of the estimate.

To numerically assess how faithfully the proposed filter approximates the actual posterior, we use the well known Kolomogorov-Smirnov (KS) statistic [44] that provides a reliable measure of the accuracy of the estimate of the posterior. The KS test evaluates the misfit between the two CDFs [44] by using the largest absolute difference. KS testing for multi-dimensional distributions is conducted using the two-sample test [45]. Although used in [46], KS testing has not yet received widespread use in the PF literature. In this paper, we conduct the two-sample KS test on the proposed filter using the ground truth shown in Fig. 3, except that now, only one target is considered. Fig. 13 shows the KS statistic misfit of the proposed filter with the theoretically optimal PF operated with the true likelihood equation in (23). We also show the KS values for the MeMBer filter in which the likelihood is: a) the MVDR value, and b) combined value of MVDR and MUSIC. As expected, we notice considerable disagreement in all the three filter approximations. Although the disagreement decreases with increasing I, it can be observed that the statistic begins to exhibit asymptotic behaviour from $I \approxeq 500$.

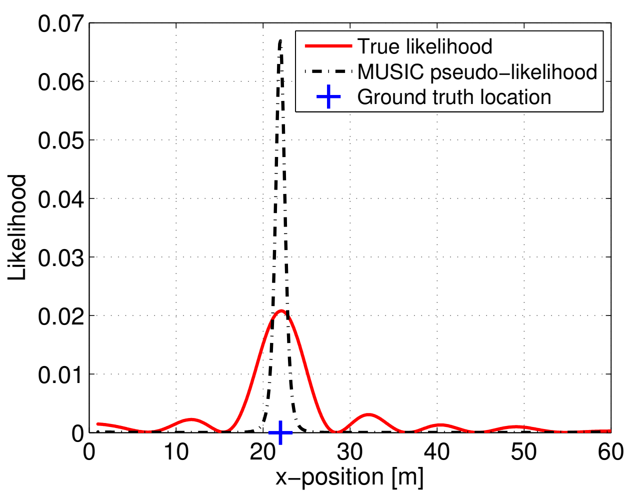


Fig. 11: The true likelihood and MUSIC pseudo-likelihood spectrums. The target is exactly at 22m.

VII. Conclusion

In this paper, we address the problem of Bayesian multi-target tracking from phased array observations. The theoretically optimal joint filter approach suffers from high computational complexity for large numbers of targets. Although the PHD and MeMBer filters overcome this complexity by operating in the dimensionality of a single target state space, they do not possess the framework to operate directly on phased array data and instead operate on the images derived from the data. Converting the phased array data causes substantial information and resolution loss, especially at high noise conditions. The key innovation given here is to use for the likelihood in the MeMBer filter, a function which has as its domain the state-space of a single target. The MUSIC pseudo-spectrum is such a function and has not previously been used in this context. The merits of the proposal (as demonstrated using simulations) are that the MeMBer filter can now operate more directly on array data, resulting in improved accuracy at low SNR. Moreover, the high computational effort, that is otherwise required by the MeMBer filter, is drastically reduced. Finally, close targets can now be easily separated by virtue of the super-resolution property of the MUSIC.

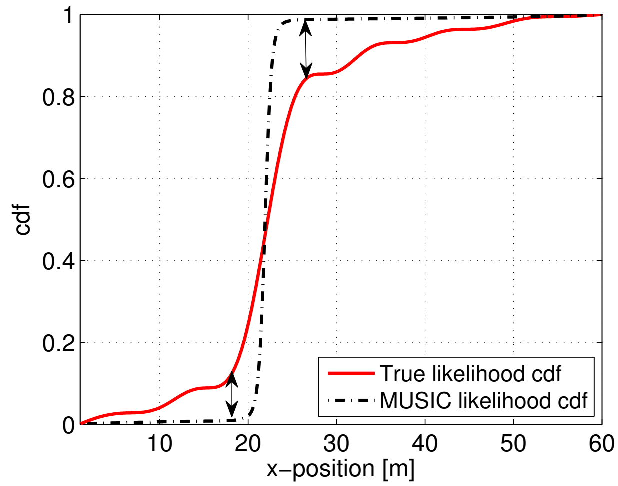


Fig. 12: CDFs for the normalised results in Fig. 11. The locations of maximum deviation are shown with the arrows.

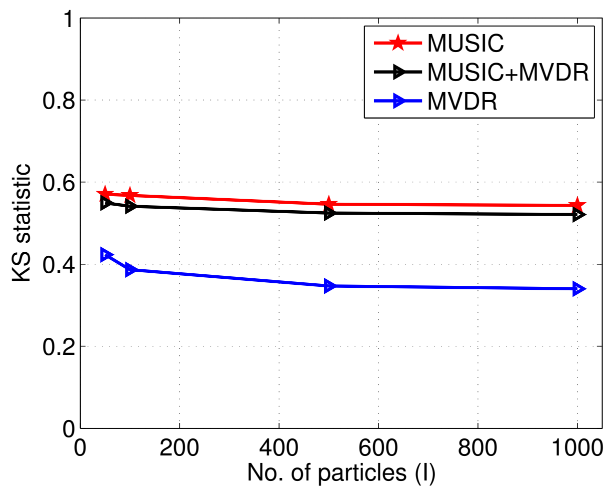


Fig. 13: KS statistic disagreement versus the number of particles. The specifications described in the caption of Fig. 5 apply to this figure also. The results are averaged over 20 trials, each comprising 50 time steps.

References

[1] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. on Signal Process.*, vol. 39, no. 5, pp. 1110–1121, 1991.

[2] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House Publishers, 2007, vol. 685.

[3] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House Pub., 2004.

[4] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to non-linear/non-Gaussian Bayesian state estimation," in *Proc. Inst. Elect. Eng.*, vol. 140, no. 2, 1993, pp. 107–113.

[5] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, 2002.

[6] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. M. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, 2003.

[7] C. Kreucher, K. Kastella, and A. O. Hero III, "Multitarget tracking using the joint multitarget probability density," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1396–1414, 2005.

[8] D. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[9] J. Vermaak, S. J.Godsill, and P. Pérez, "Monte Carlo filtering for multi target tracking and data association," *IEEE Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 309–332, 2005.

[10] M. Ekman, "Particle filters and data association for multi-target tracking," in *Proc. IEEE Information Fusion*, 2008, pp. 1–8.

[11] T. Yang, G. Huang, and P. G. Mehta, "Joint probabilistic data association-feedback particle filter for multiple target tracking applications," in *IEEE, Am. Control Conf.*, 2012, pp. 820–826.

[12] M. Orton and W. Fitzgerald, "A Bayesian approach to tracking multiple targets using sensor arrays and particle filters," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 216–223, 2002.

[13] C. Kreucher and B. Shapo, "Multitarget detection and tracking using multisensor passive acoustic data," *IEEE J. Ocean. Eng.*, vol. 36, no. 2, pp. 205–218, 2011.

[14] M. Fallon and S. J. Godsill, "Acoustic Source Localization and Tracking of a Time-Varying Number of Speakers," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 4, pp. 1409–1415, 2012.

[15] F. Daum and J. Huang, "Curse of dimensionality and particle filters," in *IEEE Proc. Aerosp. Conf.*, 2003, pp. 1979–1993.

[16] W. Yi, M. Morelande, L. Kong, and J. Yang, "A computationally efficient particle filter for multi-target tracking using an independence approximation," *IEEE Trans. on Signal Process.*, vol. 61, no. 4, pp. 843–856, 2013.

[17] B. T. Vo, B. N. Vo, and A. Cantoni, "Bayesian filtering with random finite set observations," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1313–1326, 2008.

[18] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 39, no. 4, pp. 1152–1178, 2003.

[19] B. T. Vo, B. N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 409–423, 2009.

[20] 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, 2013.

[21] B. N. Vo, B. T. Vo, N. T. Pham, and D. Suter, "Bayesian multi-object estimation from image observations," in *IEEE Proc. Inf. Fusion*, 2009, pp. 890–898.

[22] ——, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5129–5141, 2010.

[23] D. E. Clark and J. Bell, "Bayesian multiple target tracking in forward scan sonar images using the PHD filter," in *Proc. IEE Radar Sonar Navig.*, vol. 152, no. 5, 2005, pp. 327–334.

[24] D. E. Clark, I. T. Ruiz, Y. Petillot, and J. Bell, "Particle PHD filter multiple target tracking in sonar image," *IEEE Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 409–416, 2007.

[25] B. Balakwmar, A. Sinha, T. Kirubarajan, and J. P. Reilly, "PHD filtering for tracking an unknown number of sources using an array of sensors," in *IEEE Workshop Statistical Signal Process.*, 2005, pp. 43–48.

[26] E. Özkan, M. B. Guldogan, U. Orguner, and F. Gustafsson, "Ground multiple target tracking with a network of acoustic sensor arrays using PHD and CPHD filters," in *IEEE Proc. Inf. Fusion*, 2011, pp. 1–8.

[27] F. Thouin, S. Nannuru, and M. Coates, "Multi-target tracking for measurement models with additive contributions," in *IEEE Proc. Inf. Fusion*, 2011, pp. 1–8.

[28] S. Nannuru and M. Coates, "Multi-Bernoulli filter for superpositional sensors," in *IEEE Proc. Inf. Fusion*, 2013, pp. 1632–1637.

[29] M. Tobias and A. D. Lanterman, "Techniques for birth-particle placement in the probability hypothesis density particle filter applied to passive radar," *Proc. IET Radar, Sonar & Navig.*, vol. 2, no. 5, pp. 351–365, 2008.

[30] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[31] B. T. Vo, B. N. Vo, R. Hoseinnezhad, and R. P. S. Mahler, "Multi-Bernoulli filtering with unknown clutter intensity and sensor field-of-view," in *IEEE Proc. Inf. Sciences and Syst.* IEEE, 2011, pp. 1–6.

[32] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *J. of Computational and Graphical Statistics*, pp. 1–25, 1996.

[33] P. B. Choppala, P. D. Teal, and M. R. Frean, "Soft resampling for improved information retention in particle filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 13, 2013, pp. 4036–4040.

[34] Y. Abramovich and N. K. Spencer, "Locally optimal maximum-likelihood completion of a partially specified toeplitz covariance matrix," in *Proc. IEEE Workshop Statistical Signal Process.*, 2001, pp. 229–232.

[35] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930–3941, 2009.

[36] L. Huang, T. Long, E. Mao, and H. C. So, "MMSE-based MDL method for accurate source number estimation," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 798–801, 2009.

[37] B. M. Radich and K. M. Buckley, "The effect of source number underestimation on MUSIC location estimates," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 233–236, 1994.

[38] N. Vaswani, "Particle filtering for large-dimensional state spaces with multimodal observation likelihoods," *IEEE Trans. on Signal Process.*, vol. 56, no. 10, pp. 4583–4597, 2008.

[39] D. B. Ward and R. C. Williamson, "Particle filter beamforming for acoustic source localization in a reverberant environment," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 2, 2002, pp. 1777–1780.

[40] M. Kawamoto, F. Asano, H. Asoh, and K. Yamamoto, "Particle Filtering Algorithms for Tracking Multiple Sound sources using Microphone Arrays," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, 2007, pp. I–129–I–132.

[41] W. Ng, J. Li, S. Godsill, and S. pang, "Multitarget initiation, tracking and termination using Bayesian Monte Carlo methods," *The Computer J., British Computer Soc.*, vol. 50, no. 6, pp. 674–693, 2007.

[42] T. Zajic and R. P. S. Mahler, "Particle-systems implementation of the PHD multitarget-tracking filter," in *Proc. SPIE*, vol. 5096, 2003, pp. 291–299.

[43] D. Clark and J. Bell, "Data association for the PHD filter," in *Proc. IEEE Intell. Sensors, Sensor Networks and Inf. Process.*, 2005, pp. 217–222.

[44] J. A. Peacock, "Two-dimensional goodness-of-fit testing in astronomy," *Royal Astronomical Society*, vol. 202, pp. 615–627, 1983.

[45] G. Fasano and A. Franceschini, "A multidimensional version of the Kolmogorov-Smirnov test," *Royal Astronomical Society*, vol. 225, pp. 155–170, 1987.

[46] P. M. Djurić and J. Míguez, "Assessment of Nonlinear Dynamic Models by Kolmogorov–Smirnov Statistics," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5069–5079, 2010.