

# The Shifted Rayleigh Mixture Filter for Bearings-Only Tracking of Maneuvering Targets

J. M. C. Clark, S. A. Robbiati, and R. B. Vinter, *Fellow, IEEE*

**Abstract**—This paper introduces the shifted Rayleigh mixture filter (SRMF), which is based on jump Markov linear systems. The formulation permits the presence of clutter. For bearings-only tracking problems involving maneuvering targets, the conditional density of the target state given the available measurements evolves as a growing mixture of probability density functions associated with a history of manoeuvre “modes.” Similar to other “mixture” algorithms, the SRMF approximates this conditional density by a Gaussian mixture of fixed order. Unlike the extended or unscented Kalman filters, the shifted Rayleigh filter incorporates an *exact* calculation of the posterior density, when the prior is assumed to be Gaussian, given the latest bearings measurement. Computer simulations are provided to demonstrate the performance of the algorithm.

**Index Terms**—Bearings-only tracking, Gaussian mixture reduction, jump Markov linear models, mixture algorithms, particle filter (PF), shifted Rayleigh filter, unscented Kalman filter.

## I. INTRODUCTION

**I**N THIS paper, we propose the shifted Rayleigh mixture filter (SRMF) for tracking a maneuvering target, given noisy bearing measurements of target position relative to a sensor platform. The design of the filter is based on a description of the motion of the target and the sensor platform using a discrete time linear system driven by Gaussian inputs with a discrete set of random coefficients describing the current maneuvering mode. Such models, referred to as jump Markov linear models, are widely used [1], [2] because of their versatility in describing a wide range of maneuvering target motions.

For the tracking problem considered, the conditional density of the target state given the available measurements, is a mixture of probability densities. Each mixture component can be regarded as the density of the target state conditioned on the available measurements and on a mode history. The number of components in the mixture grows in time because the number of possible mode histories increases geometrically.

Our approach is to obtain estimates of target state using Gaussian mixture approximations of the conditional density. Constructing these approximations involves a moment matching operation to generate a Gaussian approximation to

each component of the mixture, and also involves removing the insignificant elements in the Gaussian mixture to reduce the computational burden.

A number of tracking algorithms have been proposed, based on propagating a Gaussian mixture approximation of the conditional density. These mixture filters differ mainly by the techniques employed to restrict the number of elements in the mixture. Algorithms such as the Generalized Pseudo-Bayes algorithm [2] or the Interacting Multiple Model (IMM) algorithm [2] merge together and retain  $N$  mixture components. This number is strictly linked to the number of possible mode histories. Other techniques eliminate the least significant terms and allow full flexibility in the choice of  $N$ . The elimination process can be deterministic (such as in the detection-estimation algorithm [3]), random (such as in the random sampling algorithm [4] and in mixture Kalman filters [5], [6]) or a mixture of both (as proposed by Fearnhead and Clifford [7]). In the case study of Section IV, we use the last method for the mixture reduction step in our implementation of the SRMF, because it is computationally efficient and satisfies certain optimality criteria explained in [7].

The above filters are based on linear/Gaussian models describing the evolution of the state variable and the measurement process, for a given manoeuvre mode history. Hence, the parameters of each mixture component can be obtained from the Kalman filter equations.

For bearings-only tracking the measurement process equation is nonlinear. The novelty of the SRMF lies in the way that it deals with this nonlinearity. To generate Gaussian approximations to the conditional densities of the state, associated with each manoeuvre mode history, it employs the shifted Rayleigh filter [8] based on an *exact* calculation of the first and second moments of the conditional density of the state.

The proposed algorithm takes account of clutter by associating with each manoeuvre mode an additional clutter mode (cf. [9]). The algorithm can be adapted also to take account of multiple sensors. The multiple sensor tracking problem can be treated as a single sensor tracking problem, in which the measurements are artificially ordered and regarded as separated in time by an interval of 0 s duration.

The performance of the new filter is assessed in simulations. The scenario considered is a variant of the one used by Marrs [10], which describes a target maneuvering through a cluster of sonobuoys, to illustrate the effectiveness of a particle filter (PF) in circumstances when extended Kalman filters fail. Comparisons are made with the generic PF using stratified sampling [11] in the resampling step and also with a modified version of our filter based on the unscented Kalman filter [12]. In this scenario, the performance of the SRMF is comparable to that of the PF, while reducing the computational burden by an order of

Manuscript received September 15, 2005; revised October 18, 2006. This work was supported in part by the Ministry of Defence. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jaume Riba.

J. M. C. Clark and R. B. Vinter are with the Department of Electrical and Electronic Engineering, Imperial College, London SW7 2BT, U.K. (e-mail: j.m.c.clark@imperial.ac.uk; r.vinter@imperial.ac.uk).

S. A. Robbiati is with Aspect Capital Limited, London W1U 1QS, U.K. (e-mail: stefano.robbiati@aspectcapital.com).

Digital Object Identifier 10.1109/TSP.2007.894378

magnitude. The SRMF also significantly outperforms the unscented Kalman filter version of the algorithm, with respect to both accuracy and the computational demands.

*Notation:* In Sections II–IV,  $\mathbf{z}_{\sigma:\tau}$  denotes the subsequence  $\{\mathbf{z}_\sigma, \dots, \mathbf{z}_\tau\}$  of a process  $\{\mathbf{z}_t\}$  between times  $\sigma$  and  $\tau$  with  $\sigma \leq \tau$ . We shall often use a generic “ $p$ ” for probability densities, probability mass functions and their joint extensions. It is clear from the context which function is being referred to. For example,  $p(\theta_t = \psi_t | r_t = j)$  refers to the density of the variable  $\theta_t$ , conditional on the discrete variable  $r_t$  being  $j$ , evaluated at  $\psi_t$ . The multivariate normal density with mean  $\mathbf{x}$  and covariance  $\mathbf{P}$  is denoted  $\mathcal{N}(\mathbf{x}, \mathbf{P})$ .

## II. SHIFTED RAYLEIGH MIXTURE FILTER (SRMF)

1) *Formulation of the Tracking Problem:* Let the state process  $\{\mathbf{x}_t\}_{t \geq 0}$  be a sequence of  $M$ -dimensional random variables, which describe the combined motion of the target in the plane and the sensor platforms, from which the target motion is observed. Let the measurement process  $\{\psi_t\}_{t \geq 1}$  be a sequence of random variables, which are noisy bearing measurements in the plane of the target position relative to the platform in the presence of clutter. The tracking problem is to estimate the state  $\mathbf{x}_t$  and the mode  $r_t$  of the target, given current and past bearings measurements  $\psi_{1:t}$ .

It is assumed that the state process is generated by a linear system driven by an  $M$ -dimensional white-noise process,  $\{\mathbf{v}_t\}_{t \geq 1}$ . To capture the maneuvering nature of target motion at time  $t$ , we allow the system matrices to be functions of the mode process  $\{r_t\}_{t \geq 1}$ , a Markov chain taking values in the discrete space  $S = \{1, \dots, d\}$ .

Our model of the measurement process involves additional processes. In the following equations we introduce: 1) the 2-D displacements  $\mathbf{d}_t$ , which are target positions relative to a sensor platform; 2) the augmented measurements  $\mathbf{y}_t$ , 2-D random variables which are the displacements corrupted by a process of variables  $\mathbf{w}_t$ ; 3) the clutter process, a sequence of scalar random variables  $\{U_t\}_{t \geq 1}$ ; and 4) a sequence of random variables  $\{\kappa_t\}_{t \geq 1}$ , which take values 0 or 1 indicating the absence or presence of clutter

$$\mathbf{x}_t = \mathbf{F}_t(r_t)\mathbf{x}_{t-1} + \mathbf{u}_t^s(r_t) + (\mathbf{Q}_t^s)^{1/2}(r_t)\mathbf{v}_t \quad (1)$$

$$\mathbf{d}_t = \mathbf{H}_t(r_t)\mathbf{x}_t + \mathbf{u}_t^m(r_t) \quad (2)$$

$$\mathbf{y}_t = \mathbf{d}_t + (\mathbf{Q}_t^m)^{1/2}(r_t)\mathbf{w}_t \quad (3)$$

$$\theta_t = \tan^{-1} \left( \frac{y_t^1}{y_t^2} \right) \quad (4)$$

$$\psi_t = (1 - \kappa_t)\theta_t + \kappa_t U_t \quad (5)$$

where  $\mathbf{y}_t = (y_t^1, y_t^2)^T$ , and for each  $r \in S$  and time  $t$ , and

$\mathbf{F}_t(r)$  is the  $M \times M$  systems matrix;

$\mathbf{Q}_t^s(r)$  is the  $M \times M$  state noise covariance matrix;

$\mathbf{H}_t(r)$  is the  $2 \times M$  output matrix;

$\mathbf{Q}_t^m(r)$  is the  $2 \times 2$  measurement noise covariance matrix.

The vector  $\mathbf{u}_t^s(r)$  is the  $M$ -dimensional exogenous state input and  $\mathbf{u}_t^m(r)$  is the 2-D exogenous measurement input. Although not explicitly shown, these matrices and vectors may depend on the past measurement history  $\psi_{1:t-1}$ . The expression  $\tan^{-1}(y_t^1/y_t^2)$  in (4) is interpreted in all four quadrants, as

the angle of clockwise rotation (in radians) of the vector  $\mathbf{y}_t$  from the positive  $y$ -axis. The noise processes  $\{\mathbf{v}_t\}$  and  $\{\mathbf{w}_t\}$  are sequences of independent and identically distributed (i.i.d.) random variables with the normal distributions

$$\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{I}_{M \times M}) \text{ and } \mathbf{w}_t \sim \mathcal{N}(0, \mathbf{I}_{2 \times 2}) \text{ for all } t.$$

The process  $\{r_t\}$  is assumed to be a discrete, stationary Markov process with transition probabilities

$$p_{ij} = P(r_t = j | r_{t-1} = i), \quad i, j \in S. \quad (6)$$

It is further assumed that the random variables  $U_t$  are i.i.d., with a density  $p(U_t)$  supported on the interval  $[-\pi, +\pi]$ , and that the random variables  $\kappa_t$  are i.i.d. with mass function

$$P(\kappa = 1) = p_c \text{ and } P(\kappa = 0) = 1 - p_c$$

for some clutter probability  $p_c$  ( $0 \leq p_c \leq 1$ ). For the special case of isotropic clutter,  $p(U_t)$  is just the uniform density, with value  $(2\pi)^{-1}$ . Finally, it is assumed that  $\{r_t\}$ ,  $\{\mathbf{v}_t\}$ ,  $\{\mathbf{w}_t\}$ ,  $\{U_t\}$  and  $\{\kappa_t\}$  are independent processes.

2) *The Measurement Model:* In the state formulation, a novel approach is taken to model the noisy bearing measurement, where a noise term with normal density  $\mathcal{N}(0, \mathbf{Q}^m(r))$  is added to the displacement  $\mathbf{d}_t$  before the evaluation of the bearing. Such a structure is effective for modelling noise occurring at the sensor, as well as random fluctuations of the line-of-sight over an extended target. In contrast, the conventional approach is to model the noisy bearing  $\theta'_t$  by adding an angular wrapped normal noise term after the evaluation of the arctan function at the displacement,  $\mathbf{d}_t = (d_t^1, d_t^2)^T$ . That is, we replace (2)–(4) by the equation

$$\theta'_t = \tan^{-1} \left( \frac{d_t^1}{d_t^2} \right) + \sigma(r_t)w'_t \text{ (modulo } 2\pi). \quad (7)$$

where  $\sigma(r)$  is a known function of  $r$ ,  $w'_t$  is sampled from  $\mathcal{N}(0, 1)$  and is independent of  $\{\mathbf{v}_t\}$ , etc.

To apply the tracking algorithm in this paper to situations where the noisy bearing measurement is described in the traditional manner (7), for a given  $\sigma(r)$ , we replace the traditional noisy bearings measurement (7) by the corresponding (2)–(4) in our formulation, with

$$\mathbf{Q}_t^m(r) = E[\|\mathbf{d}_t\|^2 | r_t = r, \psi_{1:t-1}] \sigma^2(r) \mathbf{I}_{2 \times 2}. \quad (8)$$

Intuitively, this is the right thing to do: if the process noise is inserted before evaluation of the arctan function, its covariance must be scaled up by an appropriate amount, to compensate for the downscaling by the arctan function. Detailed analysis in [8] justifies the choice of the covariance in (8) and demonstrates that interchanging the two models of the measurement process has little effect on the resulting measurement processes.

3) *Evolution of Conditional Distributions:* The SRMF propagates Gaussian mixture approximations to the conditional distribution of  $\mathbf{x}_t$  given  $\psi_{1:t}$ . To motivate the algorithm, we first examine the joint probability distribution of  $(\mathbf{x}_t, r_t)$  given  $\psi_{1:t}$ , under the hypothesis that the prior is of the form

$$\begin{aligned} p(\mathbf{x}_{t-1}, r_{t-1} = i | \psi_{1:t-1}) \\ = \sum_{n \in \{m: r_{t-1}^m = i\}} w_{t-1}^n \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)(\mathbf{x}_{t-1}) \end{aligned} \quad (9)$$

where, for a subset  $A$  of  $\{1, \dots, N\}$ ,  $\sum_{n \in A}$  denotes summation over the elements of  $A$ . For each  $n$ ,  $\hat{\mathbf{x}}_{t-1}^n$  is a known  $M \times 1$ -vector,  $\mathbf{P}_{t-1}^n$  is an  $M \times M$  covariance matrix, and  $r_{t-1}^n \in S$  is a mode value. The positive weights  $w_{t-1}^1, \dots, w_{t-1}^N$  are required to satisfy

$$\sum_{n=1}^N w_{t-1}^n = 1.$$

Marginalizing either the  $\mathbf{x}_{t-1}$  or the  $r_{t-1}$  variable from the above distribution, we find

$$p(\mathbf{x}_{t-1} | \psi_{1:t-1}) = \sum_{n=1}^N w_{t-1}^n \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)(\mathbf{x}_{t-1}) \quad (10)$$

and

$$p(r_{t-1} = i | \psi_{1:t-1}) = \sum_{n \in \{m: r_{t-1}^m = i\}} w_{t-1}^n. \quad (11)$$

We see from (10) that, as a consequence of our hypothesis on the prior, the conditional density of  $\mathbf{x}_t$  is an  $N$ -fold Gaussian mixture.

The analysis summarized in the Appendix yields the formulae presented in Table I. Notice the simplified notation used in Table I and in what follows for partially marginalized densities; for example,  $p(\mathbf{x}_t | \psi_t, r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n))$  stands for  $\int_{\mathbb{R}^M} p(\mathbf{x}_t | \psi_t, r_t, \mathbf{x}_{t-1}) \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)(\mathbf{x}_{t-1}) d\mathbf{x}_{t-1}$ . Equation (T-6) shows that, if the conditional density at time  $t-1$  is an  $N$ -fold Gaussian mixture, then the conditional density at time  $t$  is an  $(N \times d)$ -fold mixture of non-Gaussian densities.

4) *Derivation of the Tracking Posterior Distributions:* We now assemble some formulae that permit the construction of computationally tractable approximations to the probability distribution (12).

The mean  $\hat{\mathbf{x}}_t$  and covariance  $\mathbf{P}_t$  of the density

$$p(\mathbf{x}_t | \theta_t, r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}, \mathbf{P}_{t-1}))$$

are derived as follows. Define  $\hat{\mathbf{x}}_{t|t-1}$ ,  $\hat{\mathbf{y}}_{t|t-1}$ ,  $\mathbf{P}_{t|t-1}$  and  $\mathbf{S}_t$

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t(r_t) \hat{\mathbf{x}}_{t-1} + \mathbf{u}_t^s(r_t) \quad (12)$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t(r_t) \mathbf{P}_{t-1} \mathbf{F}_t^T(r_t) + \mathbf{Q}_t^s(r_t) \quad (13)$$

$$\hat{\mathbf{y}}_{t|t-1} = \mathbf{H}_t(r_t) \hat{\mathbf{x}}_{t|t-1} + \mathbf{u}_t^m(r_t) \quad (14)$$

$$\mathbf{S}_t = \mathbf{H}_t(r_t) \mathbf{P}_{t|t-1} \mathbf{H}_t^T(r_t) + \mathbf{Q}_t^m(r_t). \quad (15)$$

Define the function of the scalar variable

$$\rho(z) = \frac{ze^{-(z^2/2)} + \sqrt{2\pi}(z^2 + 1) F_{\mathcal{N}}(z)}{e^{-(z^2/2)} + \sqrt{2\pi} z F_{\mathcal{N}}(z)}$$

in which  $F_{\mathcal{N}}(z)$  is the cumulative distribution function of a unit variance, zero mean normal distribution. Let  $\mathbf{b}_t = (\sin \theta_t, \cos \theta_t)^T$  and define

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^T(r_t) \mathbf{S}_t^{-1} \quad (16)$$

$$z_t = (\mathbf{b}_t^T \mathbf{S}_t^{-1} \mathbf{b}_t)^{-(1/2)} \mathbf{b}_t^T \mathbf{S}_t^{-1} \hat{\mathbf{y}}_{t|t-1} \quad (17)$$

$$\gamma_t = (\mathbf{b}_t^T \mathbf{S}_t^{-1} \mathbf{b}_t)^{-(1/2)} \rho(z_t) \quad (18)$$

TABLE I  
FORMULA FOR CONDITIONAL DISTRIBUTIONS

For  $t \geq 1$ , the conditional distribution of  $(\mathbf{x}_t, r_t)$  given  $\psi_{1:t}$  can be expressed as follows:

$$p(\mathbf{x}_t, r_t = j | \psi_{1:t}) = \sum_{n=1}^N \mu_t^{nj} p(\mathbf{x}_t | \psi_t, r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) \quad (T-1)$$

where

$$\mu_t^{nj} = \frac{1}{c} w_{t-1}^n \pi_{r_{t-1}^n, j} p(\psi_t | r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) \quad (T-2)$$

and  $c$  is a constant chosen to ensure that

$$\sum_{n=1}^N \sum_{j=1}^d \mu_t^{nj} = 1. \quad (T-3)$$

Furthermore, the density (T-2) admits to the following expansions:

$$p(\psi_t | r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) = (1 - p_c) \tilde{p}(\theta_t = \psi_t) + p_c p(U_t = \psi_t). \quad (T-4)$$

Hence

$$p(\mathbf{x}_t | \psi_t, r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) = \frac{(1 - p_c) \tilde{p}(\theta_t = \psi_t) p(\mathbf{x}_t | \theta_t = \psi_t, r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n))}{(1 - p_c) \tilde{p}(\theta_t = \psi_t) + p_c p(U_t = \psi_t)} + \frac{p_c p(U_t = \psi_t) p(\mathbf{x}_t | r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n))}{(1 - p_c) \tilde{p}(\theta_t = \psi_t) + p_c p(U_t = \psi_t)} \quad (T-5)$$

where  $\tilde{p}(\theta_t = \psi_t)$  stands for  $p(\theta_t = \psi_t | r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n))$ . Note that from (T1) we have

$$p(\mathbf{x}_t | \psi_{1:t}) = \sum_{n=1}^N \sum_{j=1}^d \mu_t^{nj} p(\mathbf{x}_t | \psi_t, r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) \quad (T-6)$$

$$\delta_t = (\mathbf{b}_t^T \mathbf{S}_t^{-1} \mathbf{b}_t)^{-1} (2 + z_t \rho(z_t) - \rho^2(z_t)). \quad (19)$$

Then

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\gamma_t \mathbf{b}_t - \hat{\mathbf{y}}_{t|t-1}), \quad (20)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t(r_t)) \mathbf{P}_{t|t-1} + \delta_t \mathbf{K}_t \mathbf{b}_t \mathbf{b}_t^T \mathbf{K}_t^T. \quad (21)$$

Equations (12)–(21) define functions  $\hat{\mathbf{x}}_t = \hat{\mathbf{g}}(\hat{\mathbf{x}}_{t-1}, \mathbf{P}_{t-1}, \theta_t, r_t, t)$  and  $\mathbf{P}_t = \mathbf{G}(\hat{\mathbf{x}}_{t-1}, \mathbf{P}_{t-1}, \theta_t, r_t, t)$  that describe how the mean and covariances of the conditional distributions are updated in the absence of clutter for a given  $\theta_t$  and  $r_t$ , under the assumption that the prior distribution is Gaussian. Next, we provide a computable formula for the density  $p(\theta_t | r_t, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}, \mathbf{P}_{t-1}))$  in (T-5). In the following, to make its arguments explicit, this density is denoted by  $f(\hat{\mathbf{x}}_{t-1}, \mathbf{P}_{t-1}, \theta_t, r_t, t)$ .

Given  $\hat{\mathbf{x}}_{t-1}$ ,  $\mathbf{P}_{t-1}$  and  $r_t$ , we obtain  $\mathbf{S}_t$  and  $\hat{\mathbf{x}}_{t|t-1}$  from (12)–(15). Define the  $2 \times 2$  matrix

$$[a_{ij}] = \mathbf{S}_t^{-1/2}$$

and the functions

$$h_t(\theta) = \tan^{-1} \left( \frac{a_{11} \sin \theta + a_{12} \cos \theta}{a_{21} \sin \theta + a_{22} \cos \theta} \right)$$

$$g_t(\theta) = [(\sin \theta, \cos \theta) \mathbf{S}_t^{-1} (\sin \theta, \cos \theta)^T]^{-1/2}.$$

For given  $\sigma > 0$ , also define the function

$$\alpha_\sigma(\theta) = \frac{e^{-(1/2\sigma^2)}}{2\pi} \times \left( 1 + \sqrt{2\pi} \frac{\cos \theta}{\sigma} F \left( \frac{\cos \theta}{\sigma} \right) e^{(1/2)(\cos \theta/\sigma)^2} \right).$$

Now, let

$$\hat{\theta}_t = \tan^{-1} \left( \frac{\hat{y}_{t|t-1}^1}{\hat{y}_{t|t-1}^2} \right)$$

where

$$\left( \hat{y}_{t|t-1}^1, \hat{y}_{t|t-1}^2 \right)^T = \mathbf{H}_t(r_t) \hat{\mathbf{x}}_{t|t-1} + \mathbf{u}_t^m(r_t).$$

Then

$$f(\hat{\mathbf{x}}_{t-1}, \mathbf{P}_{t-1}, \theta_t, r_t, t) = \frac{g_t^2(\theta_t)}{(\det \mathbf{S}_t)^{1/2}} \alpha_{g_t(\hat{\theta}_t)}(h_t(\theta_t) - h_t(\hat{\theta}_t)). \quad (22)$$

This function gives the evaluation of the density of the measured bearing  $\theta_t$ , given the assumed value for the current mode and the parameters of the prior density, in the absence of clutter.

5) *The Shifted Rayleigh Mixture Filter*: The SRMF summarized in Table II generates  $N$ -fold Gaussian mixture approximations, with parameters

$$\{\hat{\mathbf{x}}_t^n, \mathbf{P}_t^n, r_t^n, w_t^n\}_{n=1}^N$$

to the conditional distribution of  $(\mathbf{x}_t, r_t)$ , for all  $t$ .

Suppose that an approximation to the conditional distribution of the state at time  $t - 1$  is available in the form of an  $N$ -fold Gaussian mixture (9). Then, the resulting conditional distribution of the state at time  $t$  is an  $(N \times d)$ -fold mixture of non-Gaussian densities, as shown by (T-6). This formula is not practical for computing an approximation for the conditional distribution at time  $t$ , because the number of terms in the mixture approximation grows, by a factor of  $d$ , at every iteration. A key idea behind the shifted Raleigh mixture filter is to introduce analytical approximations to the update formula (T-6). First, we replace the constituent (non-Gaussian) densities by Gaussian densities with matched first and second moments. Here, the moment calculations are based on (12)–(21), which define the mean and the covariance, and on (22). Second, we approximate the resulting  $(N \times d)$ -fold Gaussian mixture by an  $N$ -fold Gaussian mixture, using one of a variety of available mixture reduction techniques. Because  $\psi_t$  may be clutter, each density in (T-6) is a weighted sum of two densities

$$p(\mathbf{x}_t | r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n))$$

and

$$p(\mathbf{x}_t | \theta_t = \psi_t, r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)),$$

corresponding to the cases where  $\psi_t$  is clutter and  $\psi_t$  is not clutter [see (16)]. In Table II, the parameters of these two Gaussian densities are denoted as  $(\hat{\mathbf{x}}_t^{nj}(0), \mathbf{P}_t^{nj}(0))$  and  $(\hat{\mathbf{x}}_t^{nj}(1), \mathbf{P}_t^{nj}(1))$ , respectively. The pair  $(\bar{\mathbf{x}}_t^{nj}, \bar{\mathbf{P}}_t^{nj})$  is the mean and covariance of the two-fold mixture. Other variables appearing in the algorithm description are associated with weight calculations.

TABLE II  
ALGORITHM

- Initialization. Assume knowledge of  $\hat{\mathbf{x}}_0^n, \mathbf{P}_0^n, r_0^n, w_0^n$ , for  $n = 1, \dots, N$ , defining a  $N$ -fold Gaussian mixture associated with the conditional distribution of the state at time  $t = 0$ .
- For  $t \geq 1$ , given  $\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n, r_{t-1}^n$ , and  $w_{t-1}^n$ , proceed through the following steps:

### 1) Distribution Update

For  $n = 1, \dots, N$  and  $j = 1, \dots, d$ , calculate  $(\bar{\mathbf{x}}_t^{nj}, \bar{\mathbf{P}}_t^{nj}, r_t^{nj}, \bar{w}_t^{nj})$  from the relations:

$$\begin{aligned} \hat{\mathbf{x}}_t^{nj}(0) &= \hat{\mathbf{g}}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n, \psi_t, j, t), \\ \mathbf{P}_t^{nj}(0) &= \mathbf{G}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n, \psi_t, j, t), \\ \hat{\mathbf{x}}_t^{nj}(1) &= \mathbf{F}_t(j) \hat{\mathbf{x}}_{t-1}^n + \mathbf{u}_t^s(j), \\ \mathbf{P}_t^{nj}(1) &= \mathbf{F}_t(j) \mathbf{P}_{t-1}^n \mathbf{F}_t(j)^T + \mathbf{Q}_t^s(j), \\ f_t^{nj} &= f(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n, \psi_t, j, t), \\ L_t^{nj} &= (1 - p_c) f_t^{nj} + p_c \mathcal{P}(U_t = \psi_t), \\ q_t^{nj}(0) &= (L_t^{nj})^{-1} (1 - p_c) f_t^{nj}, \\ q_t^{nj}(1) &= (L_t^{nj})^{-1} p_c \mathcal{P}(U_t = \psi_t) \\ \bar{\mathbf{x}}_t^{nj} &= q_t^{nj}(0) \hat{\mathbf{x}}_t^{nj}(0) + q_t^{nj}(1) \hat{\mathbf{x}}_t^{nj}(1), \\ \bar{\mathbf{P}}_t^{nj} &= q_t^{nj}(0) \left[ \mathbf{P}_t^{nj}(0) + (\hat{\mathbf{x}}_t^{nj}(0) - \bar{\mathbf{x}}_t^{nj}) (\hat{\mathbf{x}}_t^{nj}(0) - \bar{\mathbf{x}}_t^{nj})^T \right] \\ &\quad + q_t^{nj}(1) \left[ \mathbf{P}_t^{nj}(1) + (\hat{\mathbf{x}}_t^{nj}(1) - \bar{\mathbf{x}}_t^{nj}) (\hat{\mathbf{x}}_t^{nj}(1) - \bar{\mathbf{x}}_t^{nj})^T \right], \\ \bar{w}_t^{nj} &= c_t^{-1} w_{t-1}^n \pi_{r_{t-1}^n, j} L_t^{nj}, \\ \bar{r}_t^{nj} &= j, \end{aligned}$$

where the normalizing constant  $c_t$  is  $\sum_{n=1}^N \sum_{j=1}^d w_{t-1}^n \pi_{r_{t-1}^n, j} L_t^{nj}$ .

### 2) State Estimate and Error Covariance Update

$$\begin{aligned} \hat{\mathbf{x}}_t &= \sum_{n=1}^N \sum_{j=1}^d \bar{w}_t^{nj} \bar{\mathbf{x}}_t^{nj}, \\ \mathbf{P}_t &= \sum_{n=1}^N \sum_{j=1}^d \bar{w}_t^{nj} [\bar{\mathbf{P}}_t^{nj} + (\bar{\mathbf{x}}_t^{nj} - \hat{\mathbf{x}}_t) (\bar{\mathbf{x}}_t^{nj} - \hat{\mathbf{x}}_t)^T]. \end{aligned}$$

### 3) Mixture Reduction

Approximate, by using the method of [7] (or any other mixture reduction method), the  $(N \times d)$ -fold Gaussian mixture with parameters  $\{\bar{\mathbf{x}}_t^{nj}, \bar{\mathbf{P}}_t^{nj}, \bar{r}_t^{nj}, \bar{w}_t^{nj}\}_{n=1, j=1}^{N, d}$  by an  $N$ -fold Gaussian mixture with parameters  $\{\hat{\mathbf{x}}_t^n, \mathbf{P}_t^n, r_t^n, w_t^n\}_{n=1}^N$ , and return to (1) for the next step.

The precise form that the SRMF takes in a particular application depends on the manner in which  $(N \times d)$ -fold Gaussian mixtures are approximated to  $N$ -fold Gaussian mixtures in Step 3. A variety of approaches have been proposed for tackling “mixture reduction” problems of this type. Some involve both the elimination of some elements from the mixture and also the combination of other, in some sense similar, components (see [13] or [14]). This procedure, which involves a careful analysis and modification of the spatial distribution of the elements in the mixture, has a heavy computational overhead. For this reason, the most commonly used mixture reduction methods leave the constituent densities unaltered and retain or reject the mixture components, depending on the size of their weights. In this latter category, available methods include purely deterministic selection procedures (as in the detection-estimation algorithm [3]), purely randomized selection procedures (as in the random-sampling algorithm [4]), and “partial randomization” procedures

where some elements are chosen deterministically and some randomly. In the simulations reported in Section IV, we employ a partial randomization algorithm, proposed by Fearnhead and Clifford [7], [15]. A distinctive feature of [7] and [15] is that the threshold on weight magnitude, determining whether an index value and its accompanying weight is automatically retained, is adapted to the distribution of weights, in a manner which is optimal as defined in [7] and [15].

### III. MULTIPLE SENSORS

Consider now a variant of the tracking problem of Section II, in which a total of  $K$  sensors provide independent, simultaneous bearing measurements of target position. In this problem, the state (1) and mode transition matrix (6) remain the same, but the measurement (2)–(5) are replaced by  $4 \times K$  equations for the displacements  $\{\mathbf{d}_t^k\}_{k=1}^K$ , augmented measurements  $\{\mathbf{y}_t^k\}_{k=1}^K$ , noisy measurements  $\{\theta_t^k\}_{k=1}^K$  and clutter corrupted measurements  $\{\psi_t^k\}_{k=1}^K$

$$\begin{aligned} \mathbf{d}_t^k &= \mathbf{H}_t^k(r_t)\mathbf{x}_t + \mathbf{u}_t^{m,k}(r_t) \\ \mathbf{y}_t^k &= \mathbf{d}_t^k + (\mathbf{Q}_t^{m,k})^{1/2}(r_t)\mathbf{w}_t^k \end{aligned} \quad (23)$$

$$\theta_t^k = \tan^{-1} \left( \frac{y_t^{k,1}}{y_t^{k,2}} \right) \quad (24)$$

$$\psi_t^k = (1 - \kappa_t^k)\theta_t^k + \kappa_t^k U_t^k \quad (25)$$

for  $k = 1, \dots, K$ . (Here, a superscript  $k$  indicates a variable or coefficient corresponding to the  $k$ th sensor). These equations involve sets of output parameters  $\left( \left\{ \mathbf{H}_t^k(r_t) \right\}_{k=1}^K, \text{ etc } \dots \right)$  and independent measurement noise and clutter processes  $(\{\mathbf{w}_t^k\})$  and  $\{\kappa_t^k, U_t^k\}$ .

The SRMF can be adapted to apply to this multi-sensor tracking problem as follows [9]. Impose an arbitrary ordering of the sensor measurements at each update time (the position of a particular sensor measurement in the ordering is indicated by the value of the index  $k$ ), and regard the occurrence of sensor measurements as separated in time by a time period of 0 s.

Suppose that, at time  $t$ , there is an available Gaussian mixture approximation to  $p(\mathbf{x}_{t-1}|\psi_{1:t-1}^1, \dots, \psi_{1:t-1}^K)$ . Then, a Gaussian mixture approximation to  $p(\mathbf{x}_t|\psi_{1:t}^1)$  (the updated density based on the first measurement) is obtained by applying one step of the SRMF to the system and measurement (1) and (23)–(25) (with  $k = 1$ ).

Next, a Gaussian mixture approximation to  $p(\mathbf{x}_t|\psi_{1:t}^1, \psi_{1:t}^2)$  is obtained by applying one step of the SRMF to the state equation

$$\begin{aligned} \mathbf{x}_t &= \mathbf{x}_{t-1} + 0 + 0 \\ \pi_{ij} &= \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

and measurement (23)–(25) (with  $k = 2$ ). The prior state distribution here is taken to be the previously calculated Gaussian mixture approximation to  $p(\mathbf{x}_t|\psi_{1:t}^1)$ . Estimators based on all remaining measurements are calculated in the same manner.

Notice that, in the above procedure, the order assigned to measurements is a design decision, because in some cases one ordering might produce better estimates than another. One sensible procedure for assigning an order would be to choose measurements according to their predicted error variances,

with the most accurate first. Alternatively, the interested reader can find more refined approximation procedures for multiple sensor problems in [9] (Section III), for the nonmaneuvering target case.

### IV. MULTISENSOR TRACKING OF A SEA-BORNE TARGET

In this section, we report on simulation experiments comparing the SRMF with a PF and also with a variant of our algorithm in which the update calculations are computed using the unscented Kalman filter ([12], [16], [17]). We refer to the last filter as the unscented mixture filter (UMF).

The bearings-only tracking problem considered here is related to those earlier investigated in [9], [10], and [18]. The motion of a submerged maneuvering target is tracked by three drifting sonobuoys providing noisy bearings measurements of target position, subject to isotropic clutter. A mobile airborne vehicle provides noisy, clutter-free bearings measurements of the sonobuoy positions.

1) *Sensor-Target Model*: Each of the three sonobuoys is assumed to move as the sum of an independent low-intensity Brownian motion and an integrated Brownian motion, common to all three, that represents the effect of a bulk drift [10] correlating their behaviour. The model permits three possible modes of manoeuvre for the target ( $d = 3$ ); two are known in the tracking literature [19] as left and right coordinated turns and the third as a nearly constant velocity model.

The continuous component of the state (describing the positions and velocities of the target and the observation platforms for each mode) is 12-D

$$\mathbf{x}_t = [x_t^0, \dot{x}_t^0, y_t^0, \dot{y}_t^0, x_t^1, y_t^1, x_t^2, y_t^2, x_t^3, y_t^3, u_{x,t}, u_{y,t}]^T. \quad (26)$$

Here  $x_t$  and  $y_t$  denote coordinates of position in the plane, and  $\dot{x}_t$  and  $\dot{y}_t$  the corresponding velocities (at time  $t$ ). The target and the three sonobuoys are labelled by the superscripts 0, 1, 2, and 3, respectively. The vector  $(u_{x,t}, u_{y,t})$  describes the bulk drift term common to the equations governing the positions of the three sonobuoys.

The model governing the state is a version of (1), in which the exogenous state term is zero

$$\mathbf{x}_t = \mathbf{F}(r_t)\mathbf{x}_{t-1} + (\mathbf{Q}_t^s(r_t))^{1/2}\mathbf{v}_t.$$

Here, the normalized input noise term  $\mathbf{v}_t$  has density  $\mathcal{N}(0, \mathbf{I}_{12 \times 12})$  and

$$\mathbf{F}(r) = \text{diag}(\mathbf{A}(r), \mathbf{I}_{4 \times 4}, \mathbf{I}_{4 \times 4}) + h[\mathbf{O}_{12 \times 10}, \mathbf{D}^T].$$

( $h$  is the sample period.)  $\mathbf{O}_{m \times n}$  is an  $m \times n$  matrix of zeros and

$$\mathbf{D} = [\mathbf{O}_{2 \times 4}, \mathbf{I}_{2 \times 2}, \mathbf{I}_{2 \times 2}, \mathbf{I}_{2 \times 2}, \mathbf{O}_{2 \times 2}].$$

For a given real number  $\omega$ , define

$$\mathbf{A}_\omega = \begin{bmatrix} 1 & \frac{\sin(\omega h)}{\omega} & 0 & \frac{\cos(\omega h) - 1}{\omega} \\ 0 & \cos(\omega h) & 0 & -\sin(\omega h) \\ 0 & \frac{1 - \cos(\omega h)}{\omega} & 1 & \frac{\sin(\omega h)}{\omega} \\ 0 & \sin(\omega h) & 0 & \cos(\omega h) \end{bmatrix}$$

and

$$\mathbf{A}_0 = \begin{bmatrix} 1 & h & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & h \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

For  $\omega \neq 0$ , define the matrices

$$\mathbf{R}_\omega^D = \begin{bmatrix} \frac{2(\omega h - \sin(\omega h))}{\omega^3} & \frac{(1 - \cos(\omega h))}{\omega^2} \\ \frac{(1 - \cos(\omega h))}{\omega^2} & h \end{bmatrix}$$

$$\mathbf{R}_\omega^C = \begin{bmatrix} 0 & \frac{(\omega h - \sin(\omega h))}{\omega^2} \\ \frac{-(\omega h - \sin(\omega h))}{\omega^2} & 0 \end{bmatrix}$$

and

$$\mathbf{R}_\omega = \begin{bmatrix} \mathbf{R}_\omega^D & \mathbf{R}_\omega^C \\ -\mathbf{R}_\omega^C & \mathbf{R}_\omega^D \end{bmatrix}.$$

Let

$$\mathbf{R}_0 = \begin{bmatrix} \frac{h^3}{3} & \frac{h^2}{2} & 0 & 0 \\ \frac{h^2}{2} & h & 0 & 0 \\ 0 & 0 & \frac{h^3}{3} & \frac{h^2}{2} \\ 0 & 0 & \frac{h^2}{2} & h \end{bmatrix}.$$

For  $\omega_0 > 0$ , we set

$$\mathbf{A}(r=1) = \mathbf{A}_{-\omega_0}$$

$$\mathbf{A}(r=2) = \mathbf{A}_0$$

$$\mathbf{A}(r=3) = \mathbf{A}_{+\omega_0}.$$

The mode-dependent system noise covariance matrices are

$$\mathbf{Q}_t^s(r) = \text{diag}(q_1 \mathbf{R}(r), q_2 \mathbf{I}_{2 \times 2}, q_2 \mathbf{I}_{2 \times 2}, q_2 \mathbf{I}_{2 \times 2}, q_3 \mathbf{I}_{2 \times 2}),$$

where

$$\mathbf{R}(r=1) = \mathbf{R}_{-\omega_0}$$

$$\mathbf{R}(r=2) = \mathbf{R}_0$$

$$\mathbf{R}(r=3) = \mathbf{R}_{+\omega_0}.$$

The equations above govern the motion of the variables  $(x_t^0, y_t^0)$ , and are obtained by sampling (with a period of  $h$  seconds) the exact response of the stochastic differential equation

$$\ddot{x}^0 = -\omega \dot{y}^0 + q_1 n_x$$

$$\ddot{y}^0 = \omega \dot{x}^0 + q_1 n_y$$

where  $n_x$  and  $n_y$  are independent Gaussian white noise processes with unit intensity. When  $q_1 = 0$ ,  $(x^0(t), y^0(t))$  executes a circular motion with a constant angular speed of  $|\omega|$  rad  $s^{-1}$ , anticlockwise if  $\omega > 0$  and clockwise if  $\omega < 0$ . The motion is linear if  $\omega = 0$ . Dynamic models of this general type are elsewhere referred to as coordinated turn models [19].

2) *Measurement Model*: Six simultaneous measurements are made at each time step. The three monitor-sensor measurements are identified by index values ( $k=1, 2, 3$ ), and the three sensor-target measurements by index values ( $k=4, 5, 6$ ).

The displacements  $\mathbf{d}_t^k = (d_t^{k,1}, d_t^{k,2})^T$ , ( $k=1, \dots, 6$ ) associated with the six measurements are

$$\mathbf{d}_t^k = \begin{cases} (x_t^k - x_t^{\text{ms}}, y_t^k - y_t^{\text{ms}})^T, & \text{for } (k=1, 2, 3) \\ (x_t^0 - x_t^{k-3}, y_t^0 - y_t^{k-3})^T, & \text{for } (k=4, 5, 6). \end{cases}$$

Here  $(x_t^{\text{ms}}, y_t^{\text{ms}})$  is the position of the monitoring sensor platform, which we regard in this exercise as a deterministic signal (an exogenous input).

TABLE III  
SIMULATION PARAMETERS

$h$	= 5 s	sample period
$T$	= 500 s	time horizon
$\omega_0$	= 0.02 rad $s^{-1}$	turn rate
$q_1$	= 0.005	target perturbation intensity
$q_2$	= 30	sensor perturbation intensity
$q_3$	= 0.005	bulk drift perturbation intensity
$\sigma_1$	= 0.283 rad ( $\sim 16^\circ$ )	standard deviation of sensor-target measurement noise
$\sigma_2$	= 0.028 rad ( $\sim 1.6^\circ$ )	standard deviation of monitor-sensor measurement noise
$p_c$	= 0.5	probability that the sensor-target bearing is clutter

The measurements  $\psi_t^k$  are generated by the following equations

$$\theta_t^k = \tan^{-1} \left( \frac{d_t^{k,1}}{d_t^{k,2}} \right) + \alpha_k w_t^k$$

$$\psi_t^k = (1 - \kappa_t^k) \theta_t^k + \kappa_t^k U_t \quad (27)$$

for ( $k=1, \dots, 6$ ). Here  $w_t^k$  is a scalar  $\mathcal{N}(0, 1)$  distributed noise variable,  $\kappa_t^k$  is a discrete variable taking values 0 or 1 with

$$P(\kappa_t^k = 1) = p_c$$

and  $U_t$  is a variable, uniformly distributed on  $[-\pi, \pi]$ .

It is assumed that the variables  $\{\psi_t^k\}$ ,  $\{w_t^k\}$ ,  $\{U_t\}$  are independent. We suppose that

$$\alpha_k = \begin{cases} \sigma_1, & \text{for } k=1, 2, 3 \\ \sigma_2, & \text{for } k=4, 5, 6 \end{cases}$$

and

$$p_c^k = \begin{cases} 0, & \text{for } k=1, 2, 3 \\ p_c, & \text{for } k=4, 5, 6. \end{cases}$$

(There is no clutter on sensor-monitor measurements).

Notice that the measurement model adopted in the above formulation of the tracking problem is of the traditional kind (bearing angle of the displacement plus scalar noise). For purposes of implementing the SRMF algorithm, we approximate it by replacing (27) by the following equations:

$$\mathbf{y}_t^k = \mathbf{d}_t^k + (\mathbf{Q}_t^{m,k}(r))^{1/2} \mathbf{w}_t^k \quad (28)$$

$$\theta_t^k = \tan^{-1} \left( \frac{y_t^{k,1}}{y_t^{k,2}} \right). \quad (29)$$

in which  $\{\mathbf{w}_t^k\}$  is an  $\mathcal{N}(0, \mathbf{I}_{2 \times 2})$  distributed noise term and

$$\mathbf{Q}_t^{m,k}(r) = E \left[ \left\| \mathbf{d}_t^k \right\|^2 \middle| \psi_{1:t-1}^k \right] \alpha_k^2 \mathbf{I}_{2 \times 2}.$$

In all our simulations, the traditional measurement model (27) is used to generate the data and is also used in the construction of the PF. These choices put the SRMF at a disadvantage, because there is a mismatch between the measurement model implicit in the filter and the model generating the measurements, as compared with the PF, where there is no such mismatch.

The parameters used throughout the simulations are given in Table III. The monitoring sensor platform is initially at the location (500, -2500) and moves uniformly in a straight line to the location (500, 500) (all distances are in meters). The changing

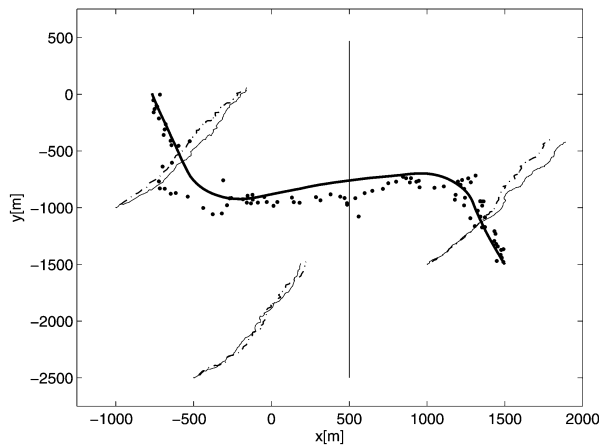


Fig. 1. Sample Ssimulation. Target's trajectory (thick black line) and a sample realization of the SRMF (dots). The position of the three buoys (thin lines) and their estimated position (dash-dot lines) is also displayed. The thin straight line shows the trajectory of the airborne monitor.

coordinates  $x_t^{\text{ms}}$ ,  $y_t^{\text{ms}}$  of the monitoring sensor platform appear as terms in  $\mathbf{u}_t^m$ . The prior state mean is

$$\hat{\mathbf{x}}_{0|0} = (1500, -2.5, -1500, 5.5, 1000, -1500, -500, -2500, -1000, -1000, 2, 2)^T$$

and the prior state covariance is  $P_{0|0} = \text{diag}(10000, 1, 10000, 1, 0, \dots, 0, 0.02, 0.02)$ .

The target manoeuvres as follows:  $r_{1:20} = 2$ ,  $r_{21:35} = 3$ ,  $r_{36:65} = 2$ ,  $r_{66:80} = 1$ , and  $r_{81:100} = 2$ . As the actual mode switching probabilities are unknown to the observer, we develop a SRMF using the following Markov transition matrix<sup>1</sup>

$$\pi = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.05 & 0.9 & 0.05 \\ 0.05 & 0.05 & 0.9 \end{bmatrix}.$$

For all the tracking algorithms considered, the procedure described in Section III is used to accommodate the multiple measurements. That is, the simultaneous measurements were regarded as sequential, with a zero time interval separating them. No special attention was given to the order in which the measurements were considered.

Fig. 1 shows a sample realization of the SRMF for the scenario just described with  $p_c = 0.5$ . The number of mixture components was  $N = 50$ . It can be seen that the filter successfully tracked the maneuvering target.

Fig. 2 provides a comparison of the average RMS errors for the SRMF, UMF and a version of the Multiple Model Bootstrap Filter [20] using "stratified sampling" (Algorithm #2 in [11]) in the final resampling phase. The errors are plotted against computational time and for varying number of particles. As the graph shows, the SRMF achieves asymptotic convergence with a computational time of approximately 10 s, and is as such the most efficient algorithm. The PF—while eventually achieving similar results—requires a very large number of samples (10,000 for the same accuracy) to deal with all the nonlinearities present in this scenario. Finally, the UMF appears to be less accurate than its competitors and to make greater computational demands. It should be pointed out, however, that the

<sup>1</sup>Simulations show that, within reason, the choice of parameters has little effect

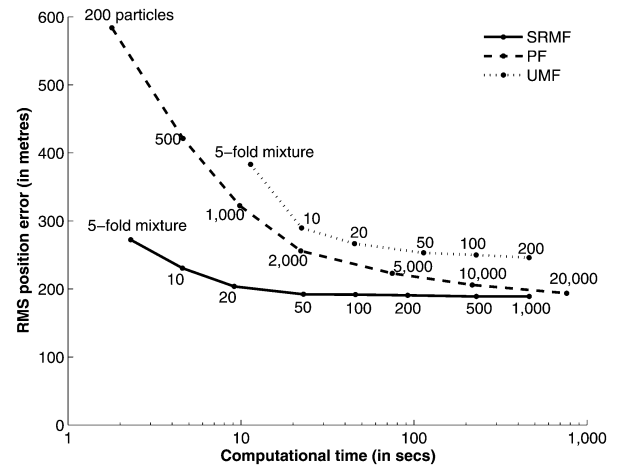


Fig. 2. Time-averaged RMS position error versus computational time. (100 Monte Carlo simulations.) The numbers beside the graph points indicate either the corresponding number of particles used in the calculation for the PF, or the number of mixture components for the SRMF and the UMF. The indicated computational time is the average running time for the entire trajectory.

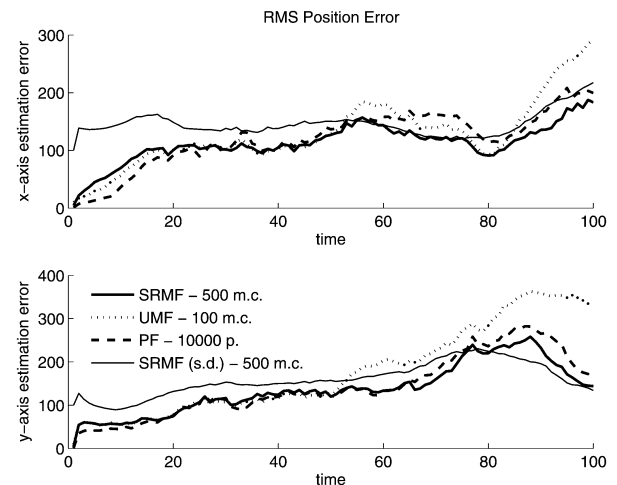


Fig. 3. Position RMS estimation error. (100 Monte Carlo simulations.) This figure shows the average  $x$ - $y$  axis RMS position error throughout the trajectory by using 500 mixture components for the SRMF, 100 for the UMF, and 10000 particles of the PF. These numbers are chosen such that the computation times of the three filters are comparable. The additional graphs are the RMS estimated standard deviations generated by the SRMF.

UMF used in these simulations is a straightforward application of the general purpose unscented Kalman filter described in [12] and [17]. A more specific design of the UMF, which takes into account the particular structure of the bearings only model, is likely to be computationally more efficient.

In Figs. 3 and 4, we also show the average RMS error (position and velocity, respectively) for the three algorithms using component and particle numbers requiring similar computational times.

Fig. 5 illustrates the effect of clutter on the estimates provided by the SRMF and the PF, matched to impose the same computational load. We can see that both algorithms can tolerate a high level of clutter up to approximately  $p_c \sim 0.667$  (i.e., two thirds of measurements are clutter). Above that, tracking performance becomes erratic and the increase in RMS error reflects occasional track losses.

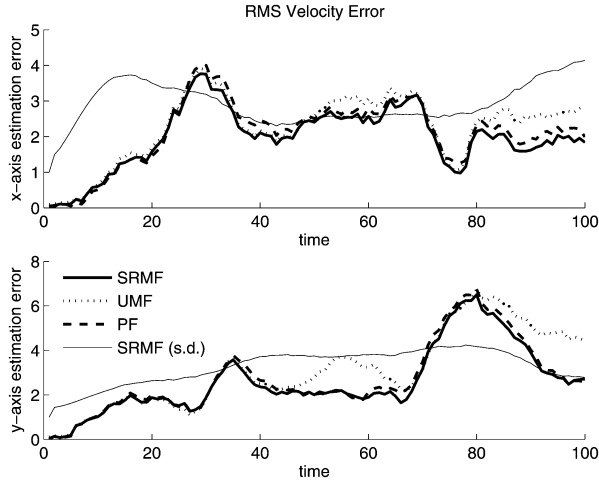


Fig. 4. Velocity RMS estimation error. (100 Monte Carlo simulations.) The parameter settings are the same as for Fig. 3.

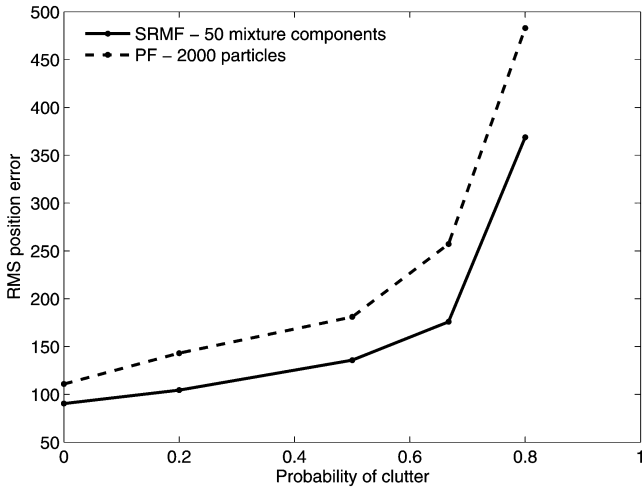


Fig. 5. Time-averaged RMS position error versus probability of clutter. (100 Monte Carlo simulations.) This figure shows the effect of the clutter probability for the SRMF and the PF with comparable computation times.

## V. CONCLUSION

In this paper, the SRMF—based on Gaussian mixture approximation of conditional densities—has been proposed for bearings-only maneuvering target tracking problems. The SRMF's effectiveness stems from the use of advanced mixture reduction techniques and the use of a refined probability analysis, implicit in the recently proposed shifted Rayleigh filter, to achieve accurate approximations.

Simulations demonstrate the effectiveness of the filter in a challenging scenario. It achieves the accuracy of a PF (while reducing the computational burden by an order of magnitude). Furthermore, it improves significantly on a related mixture filter, based on the unscented Kalman filter rather than the shifted Rayleigh filter, both as regards accuracy and computational demands.

## APPENDIX

In this Appendix, we briefly sketch the derivation of relationships (T-1)–(T-5). In our analysis, all distributions are condi-

tioned on  $\psi_{1:t-1}$ . It is convenient to introduce a new discrete random variable  $\nu$ , taking values in  $\{1, \dots, N\}$  such that the variables  $(\mathbf{x}_{t-1}, r_{t-1}, \nu)$  have a joint density-probability function

$$p(\mathbf{x}_{t-1}, r_{t-1} = i, \nu = n) = \begin{cases} w_{t-1}^n \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)(\mathbf{x}_{t-1}), & \text{if } i = r_{t-1}^n \\ 0, & \text{otherwise.} \end{cases}$$

Notice that the marginal of this distribution with respect to the  $(\mathbf{x}_{t-1}, r_{t-1})$  variables coincides with (9). Furthermore

$$p(r_{t-1} = r_{t-1}^n | \nu = n) = 1, \quad p(\nu = n) = w_{t-1}^n$$

and

$$p(\mathbf{x}_{t-1} | \nu = n) = \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n). \quad (30)$$

It follows that

$$\begin{aligned} p(\psi_t, r_t = j, \nu = n) &= p(\psi_t | r_t = j, \nu = n) p(r_t = j | \nu = n) p(\nu = n) \\ &= w_{t-1}^n \pi_{r_{t-1}^n, j} p(\psi_t | r_t = j, \nu = n). \end{aligned} \quad (31)$$

We calculate

$$\begin{aligned} p(\mathbf{x}_t, r_t = j | \psi_t) &= \sum_n p(\mathbf{x}_t, r_t = j, \nu = n | \psi_t) \\ &= \sum_n p(\mathbf{x}_t | \psi_t, r_t = j, \nu = n) \mu_t^{nj} \end{aligned}$$

where from (31)

$$\begin{aligned} \mu_t^{nj} &= p(r_t = j, \nu = n | \psi_t) \\ &= \left(\frac{1}{c}\right) \pi_{r_{t-1}^n, j} w_{t-1}^n p(\psi_t | r_t = j, \nu = n). \end{aligned}$$

Here,  $c = p(\psi_t)$ . We have  $\sum_n \sum_j \mu_t^{nj} = 1$ . In view of (30)

$$\begin{aligned} p(\psi_t | r_t = j, \nu = n) &= \int p(\psi_t | r_t = j, \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \nu = n) d\mathbf{x}_{t-1} \\ &= \int p(\psi_t | r_t = j, \mathbf{x}_{t-1}) \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)(\mathbf{x}_{t-1}) d\mathbf{x}_{t-1} \\ &= p(\psi_t | r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)). \end{aligned}$$

Next, notice that

$$\begin{aligned} p(\psi_t | r_t = j, \mathbf{x}_{t-1}) &= p(\psi_t | \kappa_t = 0, r_t = j, \mathbf{x}_{t-1}) p(\kappa_t = 0) \\ &\quad + p(\psi_t | \kappa_t = 1, r_t = j, \mathbf{x}_{t-1}) p(\kappa_t = 1) \\ &= p(\theta_t = \psi_t | r_t = j, \mathbf{x}_{t-1}) (1 - p_c) \\ &\quad + p(U_t = \psi_t | r_t = j, \mathbf{x}_{t-1}) p_c. \end{aligned}$$

Equation (T-4) follows from these last two relations. Writing  $c_{nj} = p(\psi_t | r_t = j, \nu = n)$ , we have

$$\begin{aligned} p(\mathbf{x}_t | \psi_t, r_t = j, \nu = n) &= (c_{nj})^{-1} p(\mathbf{x}_t, \psi_t | r_t = j, \nu = n) \\ &= (c_{nj})^{-1} [(1 - p_c) p(\mathbf{x}_t, \psi_t | r_t = j, \nu = n, \kappa_t = 0) \\ &\quad + p_c p(\mathbf{x}_t, \psi_t | r_t = j, \nu = n, \kappa_t = 1)] \\ &= (c_{nj})^{-1} [(1 - p_c) p(\mathbf{x}_t | \theta_t = \psi_t, r_t = j, \nu = n) \end{aligned}$$



$$\begin{aligned}
& \times p(\theta_t = \psi_t | r_t = j, \nu = n) \\
& + p_c p(\mathbf{x}_t | r_t = j, \nu = n) p(U_t = \psi_t) \\
= & (c_{nj})^{-1} \left[ (1 - p_c) \times p(\mathbf{x}_t | \theta_t = \psi_t, r_t = j, \mathbf{x}_{t-1} \right. \\
& \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n) \times p(\theta_t = \psi_t | r_t = j, \mathbf{x}_{t-1} \\
& \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) + p_c p(U_t = \psi_t) \\
& \left. \times p(\mathbf{x}_{t-1} | r_t = j, \mathbf{x}_{t-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{t-1}^n, \mathbf{P}_{t-1}^n)) \right].
\end{aligned}$$

This verifies (T-1)–(T-3), and (T-5).

#### ACKNOWLEDGMENT

The assistance of the reviewers, whose comments led to numerous improvements on an early version of the paper, is gratefully acknowledged. J. M. C. Clark and R. B. Vinter are members of the Data and Information Defence Technology Centre. The authors also wish to thank S. Maskell of QinetiQ Ltd. for his assistance in constructing the particle filter of Section IV.

#### REFERENCES

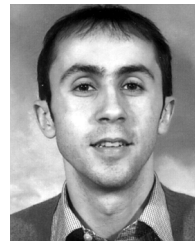
- [1] W. P. Malcolm, A. Doucet, and S. Zollo, "Sequential Monte Carlo tracking schemes for maneuvering targets with passive ranging," in *Proc. 5th Int. Conf. Inform. Fusion*, 2002, vol. 1, pp. 482–488.
- [2] Y. Bar-Shalom and X.-R. Li, *Estimation and Tracking: Principles, Techniques, and Software*. Norwood, MA: Artech House, 1993.
- [3] J. K. Tugnait, "Detection and estimation for abruptly changing systems," *Automatica*, vol. 18, pp. 607–615, 1982.
- [4] H. Akashi and H. Kumamoto, "Random sampling approach to state estimation in switching environments," *Automatica*, vol. 13, pp. 429–434, 1977.
- [5] R. Chen and J. S. Liu, "Mixture Kalman filters," *J. Roy. Statist. Soc. B*, vol. 62, no. 3, pp. 493–508, 2000.
- [6] A. Doucet, N. J. Gordon, and V. Krishnamurthy, "Particle filter for state estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, Mar. 2001.
- [7] P. Fearnhead and P. Clifford, "On-line inference for hidden Markov models via particle filters," *J. Roy. Statist. Soc. B*, vol. 65, no. 4, pp. 887–899, 2003.
- [8] J. M. C. Clark, R. B. Vinter, and M. M. Yaqoob, "The shifted Rayleigh filter: A new algorithm for bearings-only tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, 2007, to be published.
- [9] J. M. C. Clark, R. B. Vinter, and M. M. Yaqoob, "The shifted Rayleigh filter for bearings only tracking," presented at the 8th. Int. Conf. Inform. Fusion, Philadelphia, PA, Jul. 2005, CD-ROM A3-3.
- [10] A. D. Marrs, "Asynchronous multi-sensor tracking in clutter with uncertain sensor locations using Bayesian sequential Monte Carlo methods," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, 2001, vol. 5, pp. 2171–2178.
- [11] J. Carpenter, P. Clifford, and P. Fearnhead, "Improved particle filter for nonlinear problems," *Proc. IEE Radar, Sonar Navigation*, vol. 146, no. 1, pp. 2–7, 1999.
- [12] S. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proc. SPIE*, Orlando, FL, 1997, pp. 182–193.
- [13] D. J. Salmond, "Mixture reduction algorithms for target tracking in clutter," *SPIE Signal Data Process. Small Targets*, vol. 1305, pp. 434–444, 1990.
- [14] J. L. Williams and P. S. Maybeck, "Cost-function-based Gaussian mixture reduction for target tracking," in *Proc. 6th Int. Conf. Inform. Fusion*, 2003, vol. 2, pp. 1047–1054.
- [15] P. Fearnhead, "Sequential Monte Carlo Methods in Filter Theory," D.Phil. dissertation, Univ. Oxford, Oxford, U.K., 1998.
- [16] S. Julier and J. K. Uhlmann, A general method for approximating nonlinear transformations of probability distributions Robotics Res. Group, Dept. Eng. Sci., Univ. Oxford, Oxford, U.K., 1996 [Online]. Available: <http://www.robots.ox.ac.uk/siju>
- [17] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proc. IEEE Symp. Adaptive Syst. Signal Process., Commun., Contr.*, 2000, pp. 153–158.
- [18] A. D. Karan and S. E. Wang, "Bearings-only tracking with sea trial sonar data from multiple asynchronous sonobuoys," presented at the Proc. Underwater Defence Technol. Conf., Sydney, Australia, 1998.
- [19] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation: Theory Algorithms and Software*. New York: Wiley, 2001.
- [20] S. McGinnity and G. Irwin, "Multiple model bootstrap filter for maneuvering target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 3, pp. 1006–1012, May 2000.



**J. M. C. Clark** received the B.A. degree in mathematics and engineering from Cambridge University, Cambridge, U.K., in 1961 and the Ph.D. degree from Imperial College, London, U.K., in 1966.

He was a Postdoctoral Fellow at Stanford University, Stanford, CA, from 1966 to 1967 and was then employed in the Mathematics division of Bell Laboratories, Murray Hill, NJ from 1967 to 1968. Since 1969, he has worked at Imperial College, where he is now Emeritus Reader in Stochastic Systems. He is currently engaged in a project concerned with the

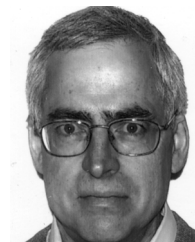
tracking of difficult targets for the MOD-funded Data and Information Defence Technology Centre. He has held visiting positions at numerous institutions, including the Universities of California at Berkeley and Davis, Harvard, MIT, and the Australian National University. His main fields of interest are stochastic modelling and estimation, and he has made contributions in nonlinear filtering, the robustness analysis of nonlinear filters, the computational aspects of stochastic processes and other areas. Recently, he has worked also in the field of target tracking, with emphasis on new, computationally efficient, filters for bearings only tracking problems.



**S. A. Robbiati** received the M.Eng. degree in electrical engineering and the Ph.D. degree from Imperial College, London, U.K., in 2002 and 2006, respectively.

During the period of his research, he worked in the Control and Power group in the Department of Electrical and Electronic Engineering, and was sponsored by BAe Systems. He is now employed at Aspect Capital Ltd., London, U.K. His research interests are in tracking and estimation with special emphasis on computationally efficient algorithms for tracking

maneuvering targets.



**R. B. Vinter** (F'98) received the B.Sc. degree in engineering science at Oxford University, Oxford, U.K., the Ph.D. degree in control engineering and the D.Sc. degree in mathematics, both from Cambridge University, Cambridge, U.K.

As a recipient of a Harkness Fellowship, he was a Postdoctoral Researcher at the Electronic Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA, for two years. Since 1974, he has been on the academic staff of Imperial College, London, U.K., where he is now Professor of control

theory. He is Head of the Control and Power Group in the Department of Electrical and Electronic Engineering. His research has centered on nonlinear control, optimization, differential games, tracking, and estimation. He has published over 140 scientific papers and two text books.