SIViP (2007) 1:149–161
DOI 10.1007/s11760-007-0012-9

ORIGINAL PAPER

Target tracking by fusion of random measures

Mahesh Vemula · Mónica F. Bugallo · Petar M. Djurić

Received: 16 October 2006 / Revised: 23 March 2007 / Accepted: 30 March 2007 / Published online: 25 April 2007
© Springer-Verlag London Limited 2007

Abstract In this paper we propose fusion methods for tracking a single target in a sensor network. The sensors use sequential Monte Carlo (SMC) techniques to process the received measurements and obtain random measures of the unknown states. We apply standard particle filtering (SPF) and cost-reference particle filtering (CRPF) methods. For both types of filtering, the random measures contain particles drawn from the state space. Associated to the particles, the SPF has weights representing probability masses, while the CRPF has user-defined costs measuring the quality of the particles. Summaries of the random measures are sent to the fusion center which combines them into a global summary. Similarly, the fusion center may send a global summary to the individual sensors that use it for improved tracking. Through extensive simulations and comparisons with other methods, we study the performance of the proposed algorithms.

Keywords Multisensor fusion · Target tracking · Particle filtering · Cost-reference particle filtering

This work has been supported by the National Science Foundation under Award CCF-0515246 and the Office of Naval Research under Award N00014-06-1-0012.

M. Vemula (✉)· M. F. Bugallo · P. M. Djurić
Department of Electrical and Computer Engineering,
Stony Brook University, Stony Brook,
NY 11794-2350, USA
e-mail: vema@ece.sunysb.edu

M. F. Bugallo
e-mail: monica@ece.sunysb.edu

P. M. Djurić
e-mail: djuric@ece.sunysb.edu

1 Introduction

Multisensor data fusion refers to the processing and synergistic combination of data from different sensors to provide improved accuracy and reduced uncertainty about events of interest [1,2]. Fusion of data from multiple sensors improves the robustness and reliability of the system and has a wide range of application including military, geosciences, robotics, statistical sciences, manufacturing and medicine [2]. In this paper we study the problem of target tracking by fusion of information in a sensor network framework as shown in Fig. 1. There each sensor applies a sequential Monte Carlo (SMC) method to obtain a random measure of the target state. The obtained information is transmitted to the fusion center (FC), which combines the received data and provides an estimate of the target state. Clearly, a better performance could be obtained by transmitting to the FC all the measurements received by the sensors without any processing and running the SMC method at the FC. However, the transmission of all these measurements is often not practical, and therefore we consider local processing at the sensors. Moreover, we assume that the fusion processing occurs *periodically* or *by request* of the FC. A challenge associated to the proposed scheme comes from the fact that the local SMC methods produce random measures represented by large sets of samples and weights/costs. The transmission of the complete measures is therefore prohibitive. We propose solutions that summarize the random measures and allow for reduced overall communication load.

The problem of target tracking is usually represented by using a discrete-time state space (DSS) model. In scenarios where the noise processes in the DSS are linear and Gaussian, the Kalman filter can be used by the sensors for obtaining local estimates of the target state [3,4]. Linear combination schemes have been proposed in this context to combine local

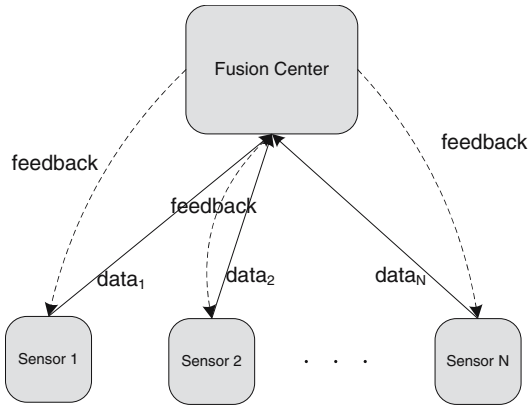

Fig. 1 Pictorial representation of the considered sensor network framework

sensor estimates [5,6]. When assumptions of linearity and Gaussianity do not hold, linearization methods and extensions of the Kalman filters (e.g. EKF) have to be considered. However, these approximations lead to poor estimates, divergence, and large uncertainty.

During the past decade, a class of SMC algorithms known as particle filtering and referred to in this paper as standard particle filtering (SPF), have been exhaustively studied [7]. SPF has shown very good performance for non-linear and non-Gaussian DSS models. This methodology represents the unknown states by evolving random measures which are composed of particles that represent the possible values of the states and weights associated to the particles that express probability masses. In order to achieve an acceptable performance, SPF requires knowledge about the distributions of the noise processes.

Recently, another type of SMC methods, known as cost-reference particle filtering (CRPF), has been proposed [8]. CRPF deals with situations where the noise distributions are not known. These methods are also based on evolving random measures composed of particles and associated costs.

Fusion of SPF-processed sensor data in the context of target tracking has been previously addressed, whereas fusion of CRPF-processed data has not. In [9], the authors propose two methods for the fusion of SPF-processed data. There the fusion rule for obtaining the joint random measure is obtained as a product of the individual clique's random measures which is not an optimal fusion rule. In our paper, we utilize optimal rules for fusion of random measures. In [10], two distributed particle filters for fusion of random measures are presented. The first one, closer in spirit to this paper, is based on factorization of likelihood terms and relies on assumptions such as sensor nodes maintaining same particles and random number seeds. We do not utilize any such assumptions in our methods. A decentralized sensor fusion framework with an information theoretic approach for sensors to collect measurements is adopted in [11]. The sensors' belief measure is approximated by a small subset of randomly chosen particles for transmission to neighboring nodes. Also, the formulated fusion expressions require the transmission of "belief" particles of past states. These requirements can be quite prohibitive even when the chosen subset of particles is small. In contrast, our methods have lower communication and power requirements. Finally, compression of the random measures using support vector machine methods is proposed and studied in [12].

The main contributions of this paper are:

1. Distributed algorithms with low communication and power requirements for fusion of SPF and CRPF-processed data are proposed. These methods are generic and can be easily adapted to various sensor network architectures. Previously proposed fusion methods either have large communication requirements or are applied to specific architectures.
2. We show the feasibility of our methods for target tracking with and without feedback from the FC. To the best of our knowledge previously proposed methods for fusion of random measures do not consider any feedback.
3. We provide simulation studies of target tracking in flat and hierarchical sensor networks with fusion of random measures.

The organization of the paper is as follows. In Sect. 2, we state the problem of data fusion for target tracking in a sensor network. In Sect. 3, we briefly describe the SPF and CRPF methods used to locally process the sensor measurements. Sect. 4 provides the theory for fusion of information when the complete posterior probability distributions of the sensors is known at the FC. In Sect. 5, we describe our proposed method for presenting the summaries of the random measures of the sensors to the FC. In Sect. 6 we show the validity of our approach through computer simulations. We consider two examples: tracking using bearings only measurements and tracking in a hierarchical sensor network. We conclude the paper with Sect. 7.

2 Problem statement

The DSS model describing the considered system is given by[1]

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}) + \gamma(\mathbf{u}_t) \tag{1}$$

$$\mathbf{y}_t^n = h^n(\mathbf{x}_t) + \mathbf{v}_t^n, \tag{2}$$

where $\mathbf{x}_t$ represents the time-varying unobservable target state, $\mathbf{y}_t^n$ the information sensed by the nth sensor, $\mathbf{u}_t$ is a

[1] Our methods can also handle models of the form $\mathbf{x}_t = f(\mathbf{x}_{t-1}, \boldsymbol{u}_t)$, $\mathbf{y}_t = h(\mathbf{x}_t, \mathbf{v}_t)$.

driving state noise process, $\mathbf{v}_t^n$ is a measurement noise process, $f(\cdot)$ is a function describing the evolution of the state, $\gamma(\cdot)$ is a deterministic function of $\mathbf{u}_t$, and $h^n(\cdot)$ is a function transforming the state to a "signal" in the measurement space [3].

The target state is obtained using a standard model of constant velocity [4] according to

$$\mathbf{x}_t = \mathbf{F}_x \mathbf{x}_{t-1} + \mathbf{\Gamma}_u \boldsymbol{u}_t \tag{3}$$

where $\mathbf{x}_t = [x_{1,t}, x_{2,t}, \dot{x}_{1,t}, \dot{x}_{2,t}]^\top \in \mathbb{R}^4$ comprises the target position and velocity in the two-dimensional space, and $\boldsymbol{u}_t$ is the random noise vector. Denoting by $\mathbf{I}_2$ and $\mathbf{0}_2$ the 2×2 identity and zero matrices, the transition matrices are given by

$$\mathbf{F}_\mathbf{x} = \begin{pmatrix} \mathbf{I}_2 & T_s\mathbf{I}_2 \\ \mathbf{0}_2 & \mathbf{I}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{\Gamma}_u = \begin{pmatrix} \frac{T_s^2}{2}\mathbf{I}_2 \\ T_s\mathbf{I}_2 \end{pmatrix}$$

with T_s being the sampling period. The sensor measurement signals $\mathbf{y}_t^n$ are typically non-linear functions of the target dynamics. We denote by $\mathbf{x}_{0:t}$ the target dynamics from time instant 0 to time instant t, and by $\mathbf{y}_{1:t}^n$ the data observed by the nth sensor up to time instant t.

When the distribution of the noise processes are known, each sensor using a SPF method approximates $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}^n)$ by a random measure $\chi_t^n = \{\mathbf{x}_{0:t}^{(m),n}, w_t^{(m),n}\}_{m=1}^M$, where $\mathbf{x}_{0:t}^{(m),n}$ are particles of the random measure, $w_t^{(m),n}$ are the weights associated to the particles, and M denotes the number of particles. When the distribution of the noise processes are unknown, each sensor using CRPF methods obtains a random measure of the state $\mathbf{x}_{0:t}$, which is represented by $\zeta_t^n = \{\mathbf{x}_{0:t}^{(m),n}, c_t^{(m),n}\}_{m=1}^M$, where $c_t^{(m),n}$ denotes the costs assigned to the particles. The objective of the proposed fusion methods is to obtain the joint random measures χ_t or ζ_t from the individual sensor random measures χ_t^n or ζ_t^n, $n = 1, \ldots, N$, and from them obtain estimates of the unknown state $\mathbf{x}_t$.

3 SMC methods

In this section we describe the SMC methods that take place at each of the sensors. For simplicity in the notation we drop the superscript n.

3.1 Standard particle filtering (SPF)

SPF methods allow for recursive approximation of the posterior density of the unknown state by $p(\mathbf{x}_{0:t},|\mathbf{y}_{1:t}) \approx \sum_{m=1}^M \delta\left(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(m)}\right) w_t^{(m)}$, where $\delta(\cdot)$ denotes the Dirac delta function [7]. At time instant t, the algorithm updates the random measure $\chi_{t-1} = \{\mathbf{x}_{0:t-1}^{(m)}, w_{t-1}^{(m)}\}_{m=1}^M$ to $\chi_t = \{\mathbf{x}_{0:t}^{(m)}, w_t^{(m)}\}_{m=1}^M$ by three main steps:

1. *Particle generation:* New particles are drawn from a proposal distribution function $\pi(\cdot)$ i.e., $\mathbf{x}_t^{(m)} \sim \pi(\mathbf{x}_t|\mathbf{x}_{t-1}^{(m)}, \mathbf{y}_{1:t})$.
2. *Weight update:* Upon reception of the measurement $\mathbf{y}_t$, the weights are updated as

$$\tilde{w}_t^{(m)} = w_{t-1}^{(m)} \frac{p(\mathbf{y}_t|\mathbf{x}_{0:t}^{(m)}, \mathbf{y}_{1:t-1}) p(\mathbf{x}_t^{(m)}|\mathbf{x}_{t-1}^{(m)})}{\pi(\mathbf{x}_t^{(m)}|\mathbf{x}_{t-1}^{(m)}, \mathbf{y}_{1:t})}$$

   and normalized such that $\sum_{m=1}^M w_t^{(m)} = 1$.
3. *Resampling:* With progress in time all but a few particles have negligible weights which may cause the filter to degenerate and diverge. Therefore, to counteract such situations particles with larger weights are proportionately replicated. This process is called resampling and it is a necessary step in the implementation of SPF methods [7].

3.2 Cost-reference particle filtering (CRPF)

In SPF schemes, the generation of particles and calculation of their weights requires the knowledge of the noise distributions in the DSS model. By contrast, the CRPF does not require assumptions about the noise distributions in the model; however, the first moments of the distributions are assumed known [8]. The random measures are constructed by applying user-defined costs. If the random measure at time instant $t-1$ is $\zeta_{t-1} = \left\{\mathbf{x}_{0:t-1}^{(m)}, c_{t-1}^{(m)}\right\}_{m=1}^M$, where $c_{t-1}^{(m)}$ are the costs assigned to the particles $\mathbf{x}_{t-1}^{(m)}$, then upon the reception of the measurement $\mathbf{y}_t$, ζ_{t-1} is updated to $\zeta_t = \left\{\mathbf{x}_{0:t}^{(m)}, c_t^{(m)}\right\}_{m=1}^M$ following the SPF structure and the principle of *survival of the fittest*. The main steps of this scheme are:

1. *Selection of the most promising paths:* This step resembles the resampling procedure in SPF schemes. The most promising paths are selected using risk functions defined as $r_t^{(m)} = \lambda c_{t-1}^{(m)} + \Delta r(\mathbf{x}_{t-1}^{(m)}|\mathbf{y}_t)$ where

$$\Delta r(\mathbf{x}_{t-1}^{(m)}|\mathbf{y}_t) = \left|\left|\mathbf{y}_t - h\left(f(\mathbf{x}_{t-1}^{(m)})\right)\right|\right|^q, \tag{4}$$

   with λ being a forgetting factor to avoid attributing excessive importance to the past, $q > 0$ and $||\cdot||$ denoting norm of a vector. These risk functions measure the adequacy of the particles at time instant $t-1$ given the new measurement $\mathbf{y}_t$ [8]. For resampling, a probability mass function (pmf), $\hat{\pi}_{r,t}^{(m)}$, is created to allow for assignment of weights to each particle, $\hat{\pi}_{r,t}^{(m)} \propto \mu_r(r_t^{(m)})$, where $\mu_r : \mathbb{R} \to [0, +\infty)$ is a monotonically decreasing

function.[2] A simple formulation of this pmf is given by

$$\mu_r\left(r_t^{(m)}\right) \propto \frac{1}{r_t^{(m)}} \tag{5}$$

which is modified into a proper pmf through normalization. Following resampling, a new stream $\hat{\zeta}_{t-1} = \left\{\hat{\mathbf{x}}_{0:t-1}^{(m)}, \hat{c}_{t-1}^{(m)}\right\}_{m=1}^{M}$ is obtained.

2. *Particle generation:* New particles are proposed using a proposal density, $p_t(\mathbf{x}_t|\hat{\mathbf{x}}_{t-1}^{(m)})$. In [8], the authors approximate $p_t(\mathbf{x}_t|\hat{\mathbf{x}}_{t-1}^{(m)})$ as a Gaussian kernel with statistics

$$\mathbb{E}(\mathbf{x}_t) = f(\hat{\mathbf{x}}_{t-1}^{(m)}), \quad \text{and} \quad \text{Cov}(\mathbf{x}_t) = \sigma_t^{2,(m)}\mathbf{I}_d,$$

where d is the dimension of the state and the variance $\sigma_t^{2,(m)}$ is recursively updated as

$$t \le \tau_0 \quad \sigma_t^{2,(m)} = \sigma_{t-1}^{2,(m)},$$

$$t > \tau_0 \quad \sigma_t^{2,(m)} = \frac{t-1}{t}\sigma_{t-1}^{2,(m)} + \frac{\left\|\mathbf{x}_t^{(m)} - f(\hat{\mathbf{x}}_{t-1}^{(m)})\right\|^2}{td},$$

with τ_0 being the time instant until which the filter obtains adequate measurements for learning the statistics of the state process.

3. *Costs update:* Costs measure the quality of the estimated state and are recursively updated as

$$c_t^{(m)} = \lambda c_{t-1}^{(m)} + \triangle c(\mathbf{x}_t^{(m)}|\mathbf{y}_t). \tag{6}$$

A possible choice of the incremental cost function is

$$\triangle c(\mathbf{x}_t|\mathbf{y}_t) = ||\mathbf{y}_t - h(\mathbf{x}_t)|\,|^q. \tag{7}$$

Depending upon system requirements such as robustness, other cost functions such as the Huber loss function or the fair function can be incorporated [13].

4. *State estimation:* A simple estimation scheme consists of choosing the particle with the minimum cost as the state estimate. Alternatively, one can construct another artificial pmf $\tilde{\pi}_{c,t}^{(m)} \propto \mu_c(c_t^{(m)})$, $m = 1, \ldots, M$, and obtain estimates such as the weighted mean of the particles.

[2] High risks indicate poor predictions of the state and lower risks indicate good predictions of the state.

4 Fusion with and without feedback

In this section we describe the theoretical expressions for fusion of probability distributions in two scenarios.[3] In the first, the global probability distribution (GPD) obtained by fusion of the individual probability distributions (IPDs) from the sensors is not reported back to the sensors, and in the second, the FC broadcasts the GPD back to the sensors, which is then used for improved tracking. These scenarios are depicted in Fig. 1. The strategy with feedback is particularly advantageous in situations when some of the filters start diverging.

4.1 Fusion without feedback

For the sake of simplicity let us consider fusion of densities from two sensors. The obtained result can readily be generalized to an arbitrary number of sensors. The GPD, $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}^1, \mathbf{y}_{1:t}^2)$, can be written as

$$\begin{aligned} p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}^1, \mathbf{y}_{1:t}^2) &= p(\mathbf{x}_{0:t}|\mathbf{y}_t^1, \mathbf{y}_t^2, \mathbf{y}_{1:t-1}^1, \mathbf{y}_{1:t-1}^2) \\ &\propto p(\mathbf{y}_t^1, \mathbf{y}_t^2|\mathbf{x}_t, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^1, \mathbf{y}_{1:t-1}^2) \\ &\quad \times p(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^1, y_{1:t-1}^2) \\ &\quad \times\ p(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}^1, \mathbf{y}_{1:t-1}^2). \end{aligned} \tag{8}$$

Assuming independence among the sensor measurements $\{\mathbf{y}_t^1, \mathbf{y}_t^2\}$ conditioned on $\mathbf{x}_t$, we have

$$\begin{aligned} p(\mathbf{y}_t^1, \mathbf{y}_t^2|\mathbf{x}_t, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^1, \mathbf{y}_{1:t-1}^2) &= p(\mathbf{y}_t^1|\mathbf{x}_t, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^1) \\ &\quad \times p(\mathbf{y}_t^2|\mathbf{x}_t, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^2). \end{aligned}$$

We know that

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}^n) = g^{-1}(y_t)\,p(\mathbf{x}_t|\mathbf{y}_{1:t-1}^n)\,p(\mathbf{y}_t^n|\mathbf{x}_t, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^n)$$

where $g(y_t)$ represents the normalization terms not involving $\mathbf{x}_t$. Therefore

$$p(\mathbf{y}_t^n|\mathbf{x}_t, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}^n) = g(y_t)\frac{p(\mathbf{x}_t|\mathbf{y}_{1:t}^n)}{p(\mathbf{x}_t|\mathbf{y}_{1:t-1}^n)}. \tag{9}$$

From (8) and (9) we get

$$\begin{aligned} p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}^1, \mathbf{y}_{1:t}^2) \propto \frac{p(\mathbf{x}_t|\mathbf{y}_{1:t}^1)}{p(\mathbf{x}_t|\mathbf{y}_{1:t-1}^1)} \quad \frac{p(\mathbf{x}_t|\mathbf{y}_{1:t}^2)}{p(\mathbf{x}_t|\mathbf{y}_{1:t-1}^2)} \quad p(\mathbf{x}_t|\mathbf{x}_{t-1}) \\ \times\ p(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}^1, \mathbf{y}_{1:t-1}^2). \end{aligned} \tag{10}$$

Note that the distributions $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ and $p(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}^1, \mathbf{y}_{1:t-1}^2)$ are known to the FC. The former is obtained from the state equation and the latter is the GPD at time instant $t-1$. The two sensors transmit to the FC information about $p(\mathbf{x}_t|\mathbf{y}_{1:t}^1)$ and $p(\mathbf{x}_t|\mathbf{y}_{1:t}^2)$, whereas $p(\mathbf{x}_t|\mathbf{y}_{1:t-1}^n), n = 1, 2,$ can

[3] Note that this is a theoretical study. It is not possible in practice for the sensors to transmit to the FC the required information. The theory explained in this section will be applied to obtain practical methods later.

in principle be obtained from $p(\mathbf{x}_{t-1}|\mathbf{y}^n_{1:t-1})$ and $p(\mathbf{x}_t|\mathbf{x}_{t-1})$, which are known at the FC.

Generalizing (10) for N sensors, we have

$$p(\mathbf{x}_{0:t}|\mathbf{y}^{1:N}_{1:t}) \propto p(\mathbf{x}_{0:t-1}|\mathbf{y}^{1:N}_{1:t-1})\; p(\mathbf{x}_t|\mathbf{x}_{t-1}) \times \prod_{n=1}^{N} \frac{p(\mathbf{x}_t\,|\,\mathbf{y}^n_{1:t})}{p(\mathbf{x}_t\,|\,\mathbf{y}^n_{1:t-1})} \tag{11}$$

which is the optimal recursive fusion equation.

4.2 Fusion with feedback

As before, we first consider fusion of posterior distributions from two sensors but with combined posterior feedback from the FC. At time instant t, the FC feeds back to the sensors the GPD $p(\mathbf{x}_{0:t-1}|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t-1})$. Then the posterior of the first sensor is formed according to

$$p(\mathbf{x}_{0:t}|\mathbf{y}^1_{1:t}, \mathbf{y}^2_{1:t-1}) \propto p(\mathbf{y}^1_t|\mathbf{x}_t) \times p(\mathbf{x}_t|\mathbf{x}_{t-1})\, p(\mathbf{x}_{0:t-1}|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t-1}). \tag{12}$$

The posterior of the second sensor is obtained analogously.

The FC receives the posteriors $p(\mathbf{x}_{0:t}|\mathbf{y}^1_{1:t}, \mathbf{y}^2_{1:t-1})$ and $p(\mathbf{x}_{0:t}|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t})$ and fuses them by

$$p(\mathbf{x}_{0:t}|\mathbf{y}^1_{1:t}, \mathbf{y}^2_{1:t}) \propto \frac{p(\mathbf{x}_t|\mathbf{y}^1_{1:t}, \mathbf{y}^2_{1:t-1})}{p(\mathbf{x}_t|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t-1})} \frac{p(\mathbf{x}_t|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t})}{p(\mathbf{x}_t|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t-1})} \times p(\mathbf{x}_t|\mathbf{x}_{t-1})\, p(\mathbf{x}_{0:t-1}|\mathbf{y}^1_{1:t-1}, \mathbf{y}^2_{1:t-1}). \tag{13}$$

This expression was derived along the same lines of reasoning as (10).

When we generalize (13) for N sensors, the fusion rule becomes

$$p(\mathbf{x}_{0:t}|\mathbf{y}^{1:N}_{1:t}) \propto p(\mathbf{x}_{0:t-1}|\mathbf{y}^{1:N}_{1:t-1}) p(\mathbf{x}_t|\mathbf{x}_{t-1}) \times \prod_{n=1}^{N} \frac{p(\mathbf{x}_t|\mathbf{y}^1_{1:t-1}, \cdots \mathbf{y}^n_{1:t} \cdots \mathbf{y}^N_{1:t-1})}{p(\mathbf{x}_t|\mathbf{y}^1_{1:t-1}, \cdots \mathbf{y}^n_{1:t-1} \cdots \mathbf{y}^N_{1:t-1})}. \tag{14}$$

5 Fusion by using random measures

The derived fusion rules from the previous section are of little practical value. In our case where we use random measures to represent our knowledge about the evolving state, the situation is even worse. The transmission of the random measures would require sending a large number of particle values and weights/costs. This altogether would be much more demanding in communication resources than the transmission of the actual measurements to the FC and would beat the whole purpose of using SMC at the sensors. We propose schemes that alleviate the sending of complete measures and by sending summaries of the random measures. First we describe strategies for summarizing the random measures constructed by SPF and then for those of CRPF.

5.1 Fusion of summaries of SPF random measures

In the SPF framework the random measures approximate distributions. If these distributions are unimodal, we propose that their random measures are summarized by Gaussians.[4] This type of approximation has already been used in the framework of Gaussian particle filtering [14]. There the individual posterior and predictive distributions are approximated by Gaussian distributions, i.e.,

$$\begin{aligned} p(\mathbf{x}_t\,|\,\mathbf{y}^n_{1:t-1}) &\simeq \mathcal{N}(\tilde{\boldsymbol{\mu}}^n_t, \tilde{\boldsymbol{\Sigma}}^n_t) \\ p(\mathbf{x}_t\,|\,\mathbf{y}^n_{1:t}) &\simeq \mathcal{N}(\hat{\boldsymbol{\mu}}^n_t, \hat{\boldsymbol{\Sigma}}^n_t) \\ p(\mathbf{x}_t\,|\,\mathbf{y}^{1:N}_{1:t}) &\simeq \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t). \end{aligned} \tag{15}$$

With these approximations and from (11) we have

$$\begin{aligned} p(\mathbf{x}_t\,|\,\mathbf{y}^{1:N}_{1:t}) &\propto \prod_{n=1}^{N} \frac{\mathcal{N}(\hat{\boldsymbol{\mu}}^n_t, \hat{\boldsymbol{\Sigma}}^n_t)}{\mathcal{N}(\tilde{\boldsymbol{\mu}}^n_t, \tilde{\boldsymbol{\Sigma}}^n_t)}\; p(\mathbf{x}_t\,|\,\mathbf{x}_{t-1})\; p(\mathbf{x}_{0:t-1}\,|\,\mathbf{y}^{1:N}_{1:t-1}) \\ &\propto \frac{\mathcal{N}(\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t)}{\mathcal{N}(\tilde{\boldsymbol{\mu}}_t, \tilde{\boldsymbol{\Sigma}}_t)} p(\mathbf{x}_t\,|\,\mathbf{x}_{t-1}) p(\mathbf{x}_{0:t-1}\,|\,\mathbf{y}^{1:N}_{1:t-1}) \end{aligned}$$

where

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_t^{-1} &= \hat{\boldsymbol{\Sigma}}_t^{1^{-1}} + \hat{\boldsymbol{\Sigma}}_t^{2^{-1}} + \cdots + \hat{\boldsymbol{\Sigma}}_t^{N^{-1}} \\ \hat{\boldsymbol{\mu}}_t &= \hat{\boldsymbol{\Sigma}}_t \left(\hat{\boldsymbol{\Sigma}}_t^{1^{-1}} \hat{\boldsymbol{\mu}}^1_t + \hat{\boldsymbol{\Sigma}}_t^{2^{-1}} \hat{\boldsymbol{\mu}}^2_t + \cdots + \hat{\boldsymbol{\Sigma}}_t^{N^{-1}} \hat{\boldsymbol{\mu}}^N_t \right) \\ \tilde{\boldsymbol{\Sigma}}_t^{-1} &= \tilde{\boldsymbol{\Sigma}}_t^{1^{-1}} + \tilde{\boldsymbol{\Sigma}}_t^{2^{-1}} + \cdots + \tilde{\boldsymbol{\Sigma}}_t^{N^{-1}} \\ \tilde{\boldsymbol{\mu}}_t &= \tilde{\boldsymbol{\Sigma}}_t \left(\tilde{\boldsymbol{\Sigma}}_t^{1^{-1}} \tilde{\boldsymbol{\mu}}^1_t + \tilde{\boldsymbol{\Sigma}}_t^{2^{-1}} \tilde{\boldsymbol{\mu}}^2_t + \cdots + \tilde{\boldsymbol{\Sigma}}_t^{N^{-1}} \tilde{\boldsymbol{\mu}}^N_t \right). \end{aligned} \tag{16}$$

The parameters of the Gaussians can readily be computed from the SPF random measure. For example, if the approximation of $p(\mathbf{x}_t \mid \mathbf{y}^n_{1:t})$ is given by $\sum_{m=1}^{M} w_t^{(m),n} \delta(\mathbf{x}_t - \mathbf{x}_t^{(m),n}) \approx \mathcal{N}(\boldsymbol{\mu}^n_t, \boldsymbol{\Sigma}^n_t)$, then

$$\begin{aligned} \boldsymbol{\mu}^n_t &= \sum_{m=1}^{M} w_t^{(m),n} \mathbf{x}_t^{(m),n} \\ \boldsymbol{\Sigma}^n_t &= \sum_{m=1}^{M} w_t^{(m),n} (\mathbf{x}_t^{(m),n} - \boldsymbol{\mu}^n_t)(\mathbf{x}_t^{(m),n} - \boldsymbol{\mu}^n_t)^\top. \end{aligned} \tag{17}$$

We define transmission length ($\mathcal{T}_L$), as the number of real numbers transmitted to the FC by each sensor. Therefore with this fusion scheme we have $\mathcal{T}_L = \frac{d(d+3)}{2}$ where d is the dimension of the state.

[4] In Appendix A, we provide conditions for the validity of the Gaussian assumptions.

When the posterior densities have more than one mode, their representation by a single Gaussian density can be inaccurate. In such situations, one possibility is to summarize the random measure by mixture-Gaussians, i.e.,

$$p\left(\mathbf{x}_t \,|\, \mathbf{y}^n_{1:t-1}\right) \simeq \sum_{l=1}^{L} \tilde{\pi}_{t,l} \mathcal{N}\left(\tilde{\boldsymbol{\mu}}^n_{t,l}, \tilde{\boldsymbol{\Sigma}}^n_{t,l}\right)$$
$$p\left(\mathbf{x}_t \,|\, \mathbf{y}^n_{1:t}\right) \simeq \sum_{l=1}^{L} \hat{\pi}_{t,l} \mathcal{N}\left(\hat{\boldsymbol{\mu}}^n_{t,l}, \hat{\boldsymbol{\Sigma}}^n_{t,l}\right), \tag{18}$$

where L is the number of mixands.

From (11) and (18) we have

$$p(\mathbf{x}_t \,|\, \mathbf{y}^{1:N}_{1:t}) \propto \prod_{n=1}^{N} \frac{\sum_{l=1}^{L} \tilde{\pi}_{t,l} \mathcal{N}(\tilde{\boldsymbol{\mu}}^n_{t,l}, \tilde{\boldsymbol{\Sigma}}^n_{t,l})}{\sum_{l=1}^{L} \hat{\pi}_{t,l} \mathcal{N}(\hat{\boldsymbol{\mu}}^n_{t,l}, \hat{\boldsymbol{\Sigma}}^n_{t,l})} \; p(\mathbf{x}_t \,|\, \mathbf{x}_{t-1}) \times p(\mathbf{x}_{0:t-1} \,|\, \mathbf{y}^{1:N}_{1:t-1}). \tag{19}$$

Thus the FC obtains the GPD through (19). The parameters of the mixture Gaussians can be obtained by the expectation-maximization (EM) algorithm which is an iterative two-step method [15]. Here, the posterior distribution is approximated by a Gaussian mixture, therefore as above, the transmission length, $\mathcal{T}_L = L\frac{d(d+3)}{2} + L - 1$.

5.2 Fusion of summaries of CRPF random measures

Recall that in the CRPF framework, each sensor maintains a random measure with costs associated to the particles. We propose that the random measure of the nth sensor, ζ^n_t, is summarized by one of the following approaches:

1. CRPF-Pdf: summary-based on a probability distribution
2. CRPF-Par: summary-based on the best particle.

5.2.1 CRPF-Pdf: Summary based on a probability distribution function

In this method, the sensors convert the costs of the particles into probability masses. Once the conversion is accomplished, the summarization is carried out in the same way as described in the previous section. For e.g., one way of converting the costs is to use

$$\pi^{(m),n}_{c,t} \propto \frac{1}{\left(c^{(m),n}_t - \min\left(c^{(m),n}_t\right) + \frac{1}{M}\right)^2}. \tag{20}$$

When there is a feedback, the FC sends back to the sensors the Gaussian constructed from the individual Gaussians. The sensors use the received Gaussian to generate particles and assign to each of them zero costs. At each sensor, the CRPF is then implemented in the usual way until the sensors receive the next feedback from the FC. Here, too, as in the SPF method when the posterior is approximated by a single Gaussian, we have $\mathcal{T}_L = \frac{d(d+3)}{2}$.

5.2.2 CRPF-Par: Summary based on the best particle

This method is the simplest of all. Here each sensor transmits to the FC the particle that has the minimal cost and the FC computes the mean of all these samples. If the sensors also transmit their minimum costs, the corresponding "best" estimate can be a weighted estimate. When the sensors have good estimates and the regions of uncertainty are small, these methods of fusing single sensor estimates are efficient. However when the sensors have large regions of uncertainty, the resulting fused estimate may not be accurate. With this method, we have a transmission length of $\mathcal{T}_L = d + 1$.

Another possibility of using the best particles and their costs is to convert the costs to a pmf and from it construct a Gaussian that can be sent back to the sensors as a feedback.

6 Simulations and results

6.1 Bearings only target tracking

In many radar and sonar applications, target tracking is performed using only bearing measurements. Here the sensors operate in a passive mode and measure the direction of arrival of the signal emitted by the target. We consider the sensor network shown in Fig. 2 where the three sensors are denoted by small circles and are placed at positions $(-50, -10)$, $(75, 10)$ and $(80, -50)$. The bearing measurements of the target relative to the nth sensor location are mathematically modeled as

$$y^n_t = \arctan\left(\frac{x_{2,t} - l^n_y}{x_{1,t} - l^n_x}\right) + v^n_t, \tag{21}$$

where $\{l^n_x, l^n_y\}$ are the coordinates of the sensor [16], and v^n_t is zero mean white Gaussian noise with variance σ^2_v.

The initial state $\mathbf{x}_0$ of the target was drawn from $\mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ with $\boldsymbol{\mu}_0 = [0,\ 0,\ 0.1,\ 0.0]^\top$ and $\boldsymbol{\Sigma}_0 = \text{diag}(10,\ 10,\ 0.1,\ 0.1)$.[5] The target's dynamics were modeled using (3) with the covariance matrix of the zero mean Gaussian state process noise $\mathbf{C}_u = \text{diag}(0.05,\ 0.02)$. The target's trajectory was simulated for $N_t = 100$ time instants with a sampling period $T_s = 1$ s. The standard deviation of the measurement noise σ_v was set to 0.05. In computing the RMSEs of the unknowns, we used 100 trajectory runs.

In Fig. 2a we show a target trajectory and its estimates using the proposed SPF method (labeled as *SPF-Dist*, meaning distributed). For comparison purposes we also included

[5] The symbol diag($\mathbf{x}$) represents a diagonal matrix formed with vector $\mathbf{x}$ as its diagonal elements.

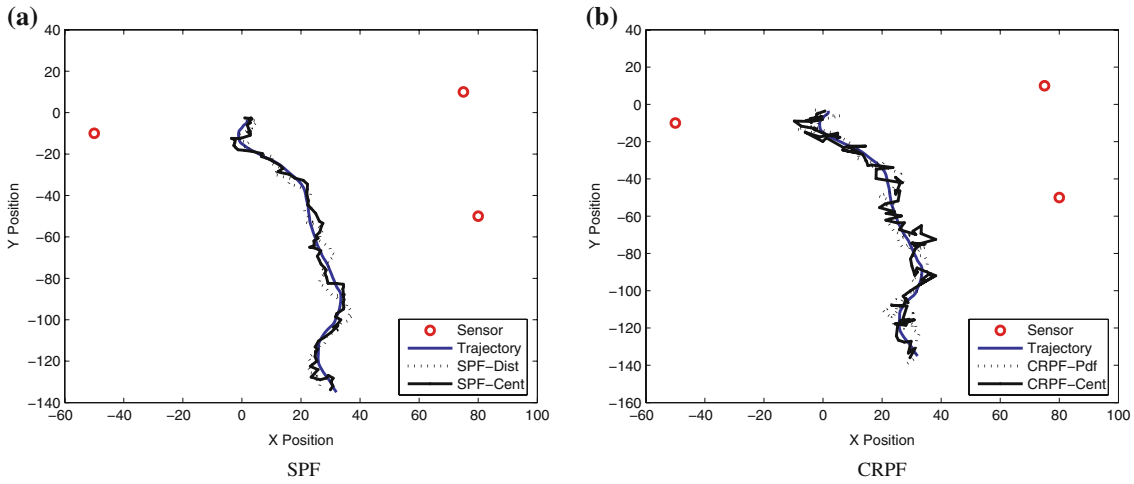

Fig. 2 The multisensor network and estimates of the target trajectory

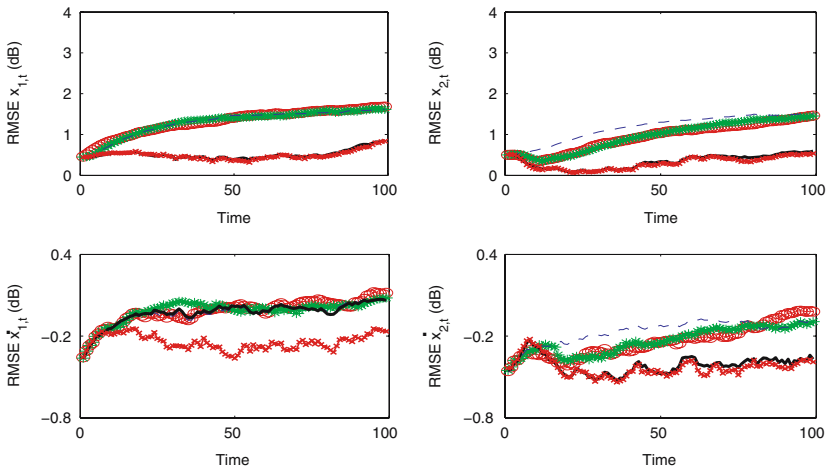

Fig. 3 RMSEs (dB) of the target dynamics by the sensors and FC using SPF *without* feedback. -o- sensor 1, -*- sensor 2, — sensor 3, – FC SPF-Dist, -x FC SPF-Cent

the performance of the system where the sensors sent their measurements to the FC. The FC used the measurements to run a SPF-based algorithm (labeled as *SPF-Cent*, meaning centralized). In Fig. 2b, we show analogous results for the CRPF-Pdf and the centralized CRPF (labeled as *CRPF-Cent*) methods operating on the full measurements. We chose the risk and cost functions outlined in (4) and (7) with $q = 2$. This method is thus analogous to a least squares fitting approach where decision on the particles is made using the residuals.

In Figs. 3 and 4, we plot the RMSEs obtained by the sensors and the FC using SPF methods with and without feedback. Each of the sensors maintained a random measure with $M = 1,000$ samples. The SPF methods utilized the prior for propagation of the particles. In both scenarios, the sensors transmitted to the FC the summaries of their random measures at every time instant. In the feedback scenario, the FC also transmitted to the sensors the global summary which was used to re-initialize their filters. It can be observed that with feedback, the RMSEs of each of the individual sensors is smaller than in situations without any feedback. The overall improvement in the RMSEs of the FC estimates, however, is small.

In Fig. 5a, we plot the RMSEs of the position when the SPF algorithms approximated the random measure with a single Gaussian density (GM-1) and with a Gaussian mixture with two mixands (GM-2). We use the χ^2 divergence[6] measure to compare the approximation of the posterior by a Gaussian distribution using SMC and EKF methods. In Fig. 5b, we plot the χ^2 divergence measure due to approximation of the sampled density with a Gaussian distribution at sensor 1.

[6] The definition and computation of this measure are described in Appendix B.

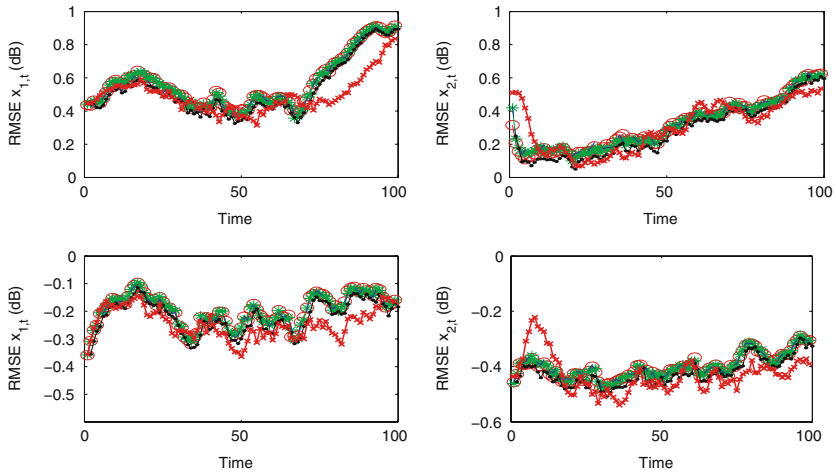

Fig. 4 RMSEs of the target dynamics by the sensors and FC using SPF *with* feedback. -o- sensor 1, -*- sensor 2, — sensor 3, – FC SPF-Dist, -x FC SPF-Cent

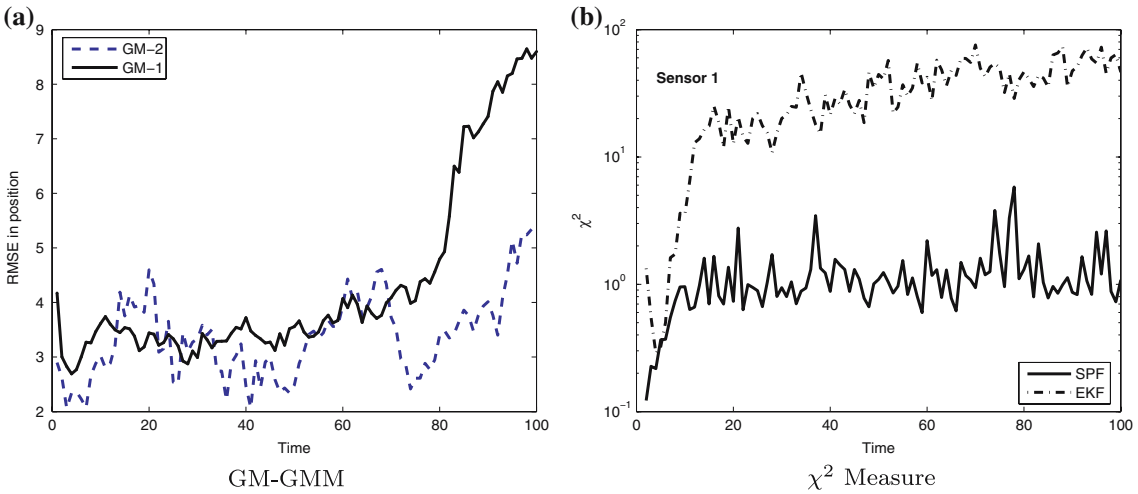

Fig. 5 (**a**) Comparison of RMSEs with a single Gaussian approximation and with a Gaussian mixture approximation (**b**) χ^2-Divergence for SPF and EKF based posterior approximations at sensor 1

Table 1 Bias (m) location estimates of the targets at different times

Time (s)	Sen 1	Sen 2	Sen 3	FC
20	–2.83	–0.46	–0.23	0.19
40	–0.77	–2.18	–2.92	–0.47
60	1.61	–5.63	–5.04	–0.49
80	2.70	–8.23	–4.22	–0.07
100	5.94	–7.16	–0.12	0.81

Table 2 Standard deviation (m) of location estimates of the targets at different times

Time (s)	Sen 1	Sen 2	Sen 3	FC
20	10.96	11.40	10.61	3.62
40	19.83	24.17	22.53	2.68
60	32.76	31.03	30.99	2.60
80	40.82	43.12	35.80	2.99
100	50.86	50.61	43.28	7.03

Clearly the performance of the SPF-based approximation is better than the Gaussian approximation with the EKF by an order of magnitude. The plots are similar for the approximations made at other sensors. In Fig. 6, we plot the mean error and the 3σ confidence interval depicting the spread in error in estimating the target dynamics at the FC. We summarize the results in Tables 1 and 2.

Clearly we can see through these plots that the spreads in errors are similar for the SPF-based summary approach and the SPF method with complete measurements. However the EKF exhibits divergence as can be seen from the high spread of errors. The results in the table show that the errors increase with time, which is due to the target

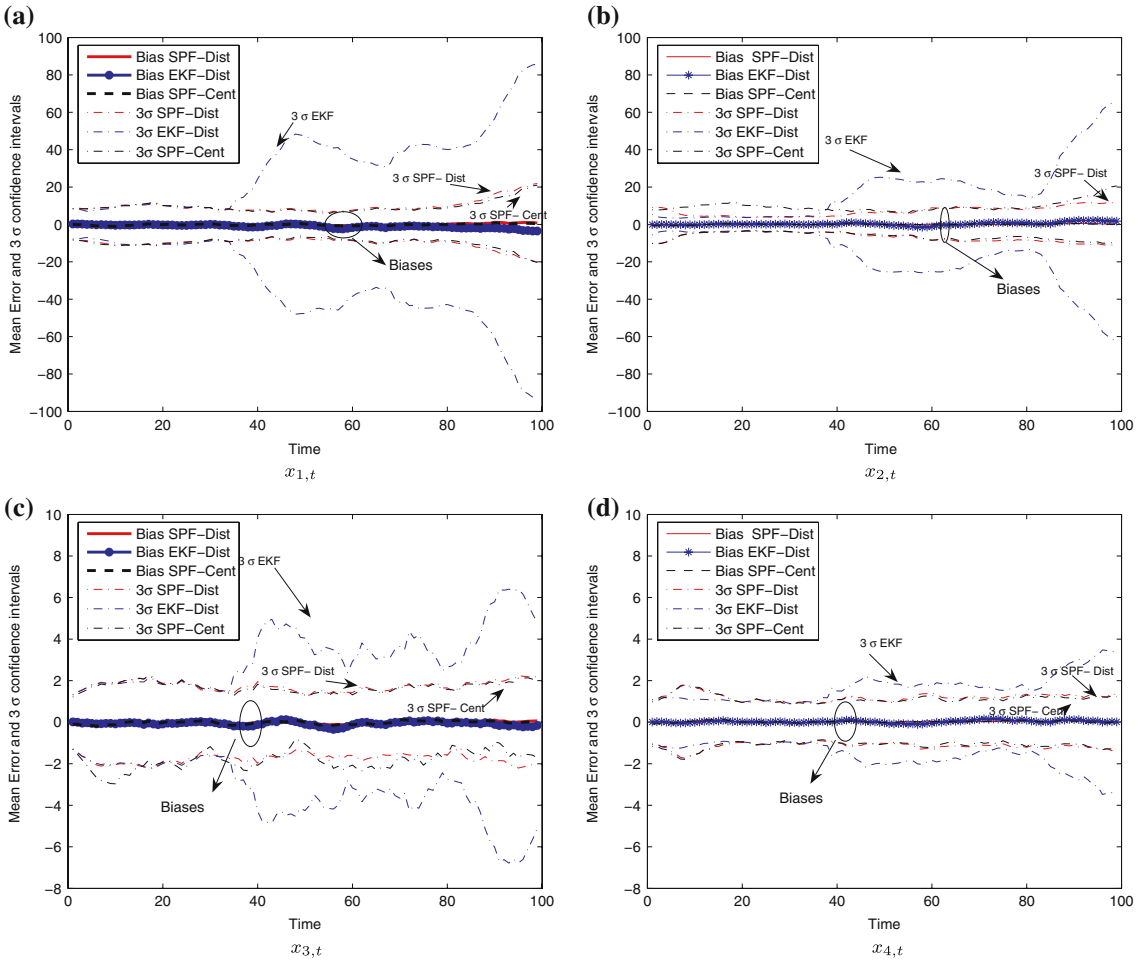

Fig. 6 Spread of errors in estimating the state at the FC

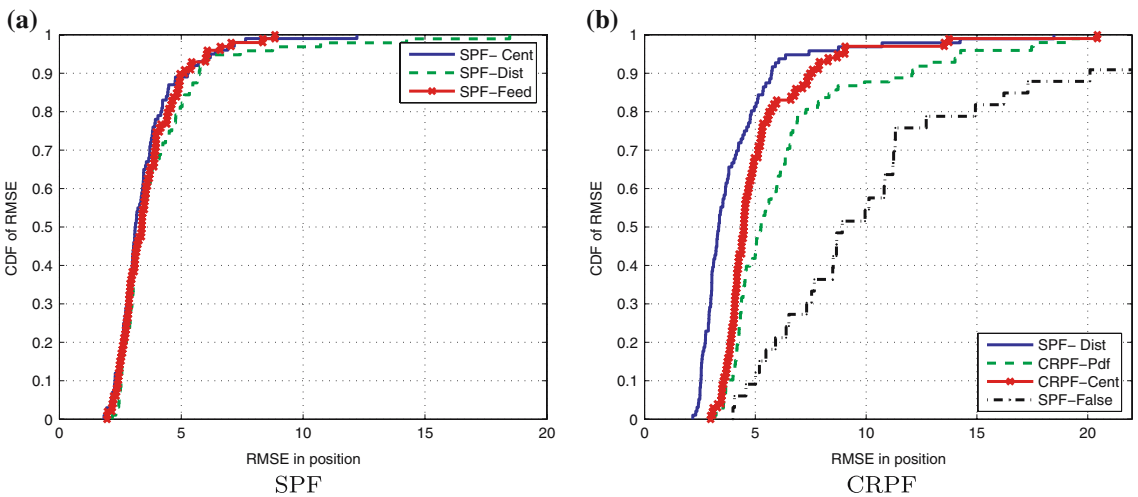

Fig. 7 Comparison of RMSEs with centralized and proposed SPF and CRPF methods

leaving the sensor field. We also use the cumulative distribution function (cdf) of the total RMSE accumulated over time as our metric for analyzing the performance of the proposed and the centralized algorithms. It can be seen from Fig. 7a that there is a small loss of performance with the proposed method over the centralized methods. We also plot

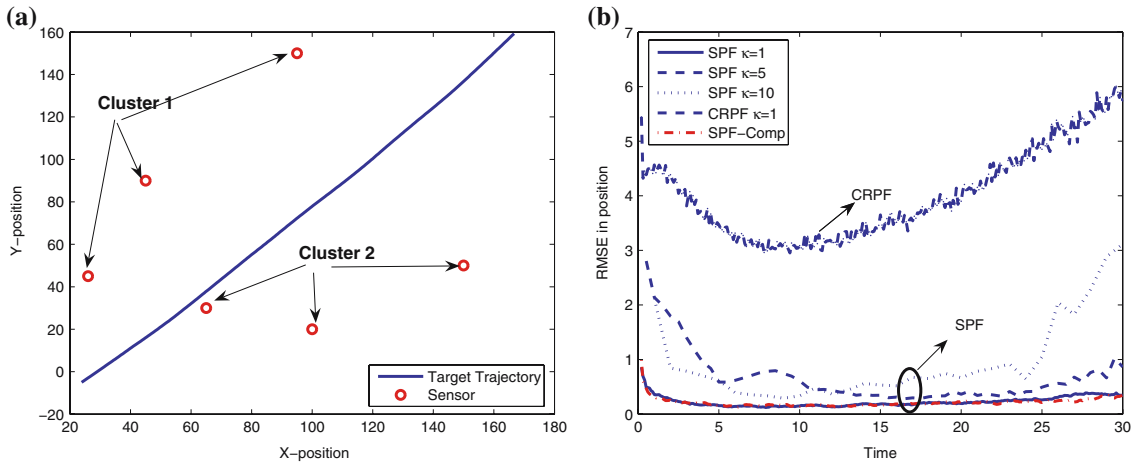

Fig. 8 (**a**) A hierarchical sensor network (**b**) RMSEs of the target position by using SPF and CRPF for various κ

in Fig. 7b the cdf of the RMSEs obtained using the CRPF methods. The performance of the CRPF-Pdf method is comparable with that of the SPF-Dist method which makes full assumptions of the noise processes. Moreover, the performance of the proposed CRPF-Dist method is very similar to the CRPF-Cent. We also simulated the case when the SPF methods make wrong assumptions about the noise process. In Fig. 7b, the filter associated with the SPF-False method assumed $\mathbf{C}_u = \text{diag}(0.005,\ 0.002)$ and $\sigma_v = 0.01$. With a single Gaussian approximation we have, the probability P that the RMSE in position is less than or equal to $10m$, i.e., $P(\text{RMSE}_{\text{pos}} \leq 10m) = 0.96$ while with a Gaussian mixture $(L = 2)$, $P(\text{RMSE}_{\text{pos}} \leq 10m)) \approx 1.0$; and with CRPF-Pdf, $P(\text{RMSE}_{\text{pos}} \leq 10m) = 0.86$. However when the probability distributions of the noise processes are unknown, for the SPF with incorrect noise statistics, $P(\text{RMSE}_{\text{pos}} \leq 10m) = 0.57$. Clearly the performance of the SPF method under wrong noise assumptions is poor, motivating the use of other methods which make less assumptions of the noise processes when their distributions are unknown.

6.2 Target tracking in a hierarchical sensor network

An important application of data fusion is target tracking in hierarchical sensor networks (HSN), where sensors form clusters and transmit their measurements to a specialized node known as a leader node (LN). This LN is a specialized node which has greater computational and communication capabilities than the sensors. Upon obtaining the measurements from the sensors, the LNs estimate the posterior density of the target's dynamics. The summaries of the random measures are then transmitted to the FC which combines them to obtain a joint summary. In this experiment, we implemented our proposed ideas for target tracking in the HSN shown in Fig. 8a. There are six sensors placed at $(26, 45)$, $(45, 90)$, $(65, 30)$, $(95, 150)$, $(100, 20)$, $(150, 50)$ which form two clusters. The sensors collected three different measurements, the bearing, the power of the signal emitted by the target and the relative velocity of the target [17]. These three signal modalities can be mathematically written as

$$y^n_{1,t} = \arctan\left(\frac{x_{2,t} - l^n_y}{x_{1,t} - l^n_x}\right) + v^n_{1,t} \tag{22}$$

$$y^n_{2,t} = \Psi - 5\log_{10}\left((x_{1,t} - l^n_x)^2 + (x_{2,t} - l^n_y)^2\right) + v^n_{2,t} \tag{23}$$

$$y^n_{3,t} = v^n_{r,t} + v^n_{3,t} \tag{24}$$

where $\Psi = -50\,\text{dB}$ is the signal strength within a known reference distance, $v^n_{r,t}$ is relative velocity of the target relative to the nth sensor, and $v^n_{1,t} \sim N(\mu_{v_1}, \sigma^2_{v_1})$, $v^n_{2,t} \sim N(\mu_{v_2}, \sigma^2_{v_2})$, $v^n_{3,t} \sim N(\mu_{v_3}, \sigma^2_{v_3})$, are zero mean Gaussian measurement noise processes. We assumed the variance of the noise to be the same across all the sensors. In our simulations we used $\sigma_{v_1} = 0.01$, $\sigma_{v_2} = 1$, and $\sigma_{v_3} = 1$. The initial target dynamics were drawn from $\mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ with $\boldsymbol{\mu}_0 = [0,\ 0,\ 5,\ 5]^\top$ and $\boldsymbol{\Sigma}_0 = \text{diag}\,(10,\ 10,\ 0.01,\ 0.01)$. The target trajectories were obtained as earlier, with the noise covariance matrix and $\mathbf{C}_u = \text{diag}(0.1,\ 0.2)$. The sampling period was $T_s = 0.1$ s and the length of each trajectory was $N_t = 300$, i.e., the target was observed for 30 s.

Here we also study the scenarios when the LNs transmit the summaries of their random measures to the FC for every $\kappa = 1, 5$, and 10 time instants. The RMSEs were computed over 50 different trajectory runs. In Fig. 8b, the RMSEs of the proposed SPF method at the FC are shown. Here, the LNs' approximate their random measure using single Gaussians. There is a very small loss in performance with the LNs transmitting the summaries to the FC every $\kappa = 10$ time instants, thus considerably saving power and communication resources. In Fig. 9b, we display the bias and the spread

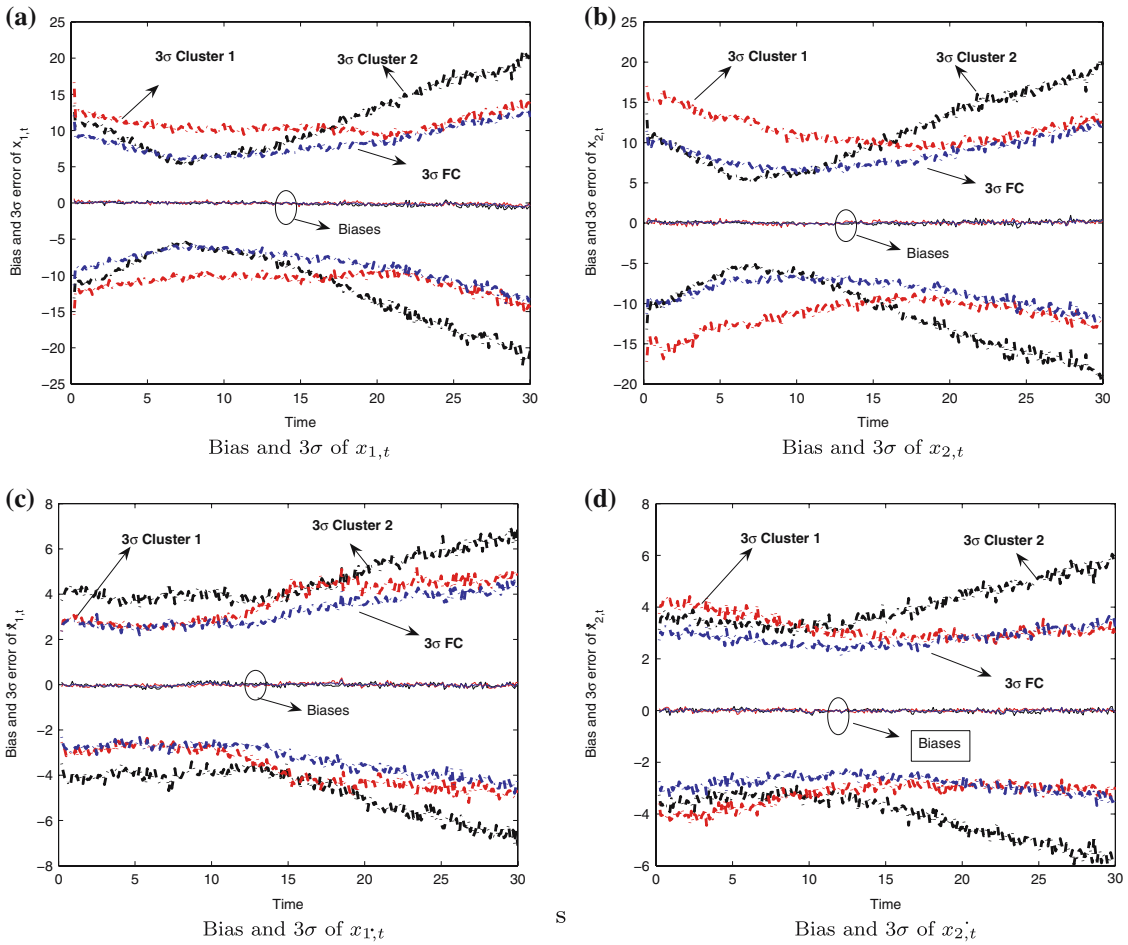

Fig. 9 Spread of errors in estimating the state at FC Bias and 3σ of **a** $x_{1,t}$; **b** $x_{2,t}$; **c** $\dot{x}_{1,t}$; **d** $\dot{x}_{2,t}$

of the errors with the CRPF-par method. Clearly, it can be seen that the estimate by the FC is for most of the time better than the estimates of the LNs.

7 Conclusion

In this paper SPF and CRPF fusion algorithms for target tracking have been presented for various scenarios, including those of known and unknown noise probability distributions. The sensors implement these SMC methods and periodically transmit summaries of their results instead of the complete random measures. Therefore, low communication and power requirements are needed. In SPF methods, the summaries are represented by mean and covariance matrices. When the posterior distribution is unimodal, these random measures are adequately approximated with a single Gaussian which has lower transmission requirements. With CRPF methods, the summaries are represented either with a parametric distribution or by using the best particles of the random measure.

We have demonstrated the performance of the proposed algorithms through two target tracking examples in a flat and hierarchical sensor network. Comparisons of the proposed methods with the standard centralized approach is performed and their good performance is illustrated by way of simulations. Approximations of random measures with a single Gaussian have slightly larger RMSE than approximations with Gaussian mixtures. The CRPF-based summaries have moderate error statistics but outperform the SPF with incorrect noise statistics. We have also investigated the effect of feedback from the FC and observed that feedback from the FC improves the individual sensor tracking performance.

Some of the open issues not addressed in this paper is determining the number of components required for an adequate representation of the sample-based distribution. Clearly, this becomes a model order selection problem with simultaneous estimation of model order and the parameters of the model.

Appendix A. Conditions for the validity of the Gaussian assumptions

Let $p(\boldsymbol{\theta})$ be a smooth density with $\boldsymbol{\theta} \in \mathbb{R}^n$. Expanding the logarithm of the density around its mean $\hat{\boldsymbol{\theta}}$ using Taylors series and neglecting the higher order derivative terms (> 2) we have

$$\begin{aligned}\ln(p(\boldsymbol{\theta})) = &-\frac{1}{2}(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}) + \mathbf{g}(\hat{\boldsymbol{\theta}})^\top(\hat{\boldsymbol{\theta}}) \\ &+\mathbf{g}(\hat{\boldsymbol{\theta}})^\top \boldsymbol{\Sigma}\mathbf{g}(\hat{\boldsymbol{\theta}}) + \ln(p(\hat{\boldsymbol{\theta}}) + \epsilon(\boldsymbol{\theta})\end{aligned}$$

where we express, $\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} + \boldsymbol{\Sigma}\mathbf{g}(\hat{\boldsymbol{\theta}})$, $\mathbf{g}(\hat{\boldsymbol{\theta}}) = \frac{\partial \ln(p(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \mid_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ and $\boldsymbol{\Sigma}^{-1} = -\frac{\partial^2 \ln(p_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} \mid_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$. Thus we have

$$p(\boldsymbol{\theta}) \propto \exp\left[-\frac{1}{2}(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})\right] \propto \mathcal{N}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\Sigma}). \quad (25)$$

Similarly we have for a product of N densities

$$p_N(\boldsymbol{\theta}) = \prod_n p_n(\boldsymbol{\theta}) \propto \prod_n \mathcal{N}_n(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}_n, \boldsymbol{\Sigma}_n) \propto \mathcal{N}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}_f, \boldsymbol{\Sigma}_f) \quad (26)$$

$$\begin{aligned}\boldsymbol{\Sigma}_f^{-1} &= \boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1} + \cdots + \boldsymbol{\Sigma}_N^{-1} \\ \tilde{\boldsymbol{\theta}}_f &= \boldsymbol{\Sigma}_f(\boldsymbol{\Sigma}_1^{-1}\tilde{\boldsymbol{\theta}}_1 + \cdots + \boldsymbol{\Sigma}_N^{-1}\tilde{\boldsymbol{\theta}}_N) \\ \tilde{\boldsymbol{\theta}}_f &= \boldsymbol{\Sigma}(\boldsymbol{\Sigma}_1^{-1}\hat{\boldsymbol{\theta}}_1 + \cdots + \boldsymbol{\Sigma}_N^{-1}\hat{\boldsymbol{\theta}}_N) + \boldsymbol{\Sigma}(\mathbf{g}(\hat{\boldsymbol{\theta}}_1) \\ &\quad +\mathbf{g}(\hat{\boldsymbol{\theta}}_2) + \cdots + \mathbf{g}(\hat{\boldsymbol{\theta}}_N)) \\ \boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\theta}}_f &= \boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\theta}}_f + (\mathbf{g}(\hat{\boldsymbol{\theta}}_1) + \mathbf{g}(\hat{\boldsymbol{\theta}}_2) + \cdots + \mathbf{g}(\hat{\boldsymbol{\theta}}_N)) \quad (27) \\ &= \boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\theta}}_f + \Delta\mathbf{g}(\hat{\boldsymbol{\theta}}).\end{aligned}$$

A ratio of two distributions can be expressed as

$$\frac{p_1(\boldsymbol{\theta})}{p_2(\boldsymbol{\theta})} \propto \mathcal{N}_1(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}_s, \boldsymbol{\Sigma}_s), \quad (28)$$

where $\boldsymbol{\Sigma}_s^{-1} = \boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1}$ and $\boldsymbol{\Sigma}_s^{-1}\tilde{\boldsymbol{\theta}}_s = \boldsymbol{\Sigma}_1^{-1}\tilde{\boldsymbol{\theta}}_1 - \boldsymbol{\Sigma}_2^{-1}\tilde{\boldsymbol{\theta}}_2$. Thus we have the following form for a ratio of a product of densities using (26) and (28)

$$\frac{\prod_k p_{a,k}(\boldsymbol{\theta})}{\prod_k p_{b,k}(\boldsymbol{\theta})} \propto \frac{\mathcal{N}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}_a, \boldsymbol{\Sigma}_a)}{\mathcal{N}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}_b, \boldsymbol{\Sigma}_b)} \propto \mathcal{N}(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}_{ab}, \boldsymbol{\Sigma}_{ab}) \quad (29)$$

$$\begin{aligned}\boldsymbol{\Sigma}_{ab}^{-1} &= \boldsymbol{\Sigma}_a^{-1} - \boldsymbol{\Sigma}_b^{-1} \\ \boldsymbol{\Sigma}_{ab}^{-1}\tilde{\boldsymbol{\theta}}_{ab} &= \boldsymbol{\Sigma}_a^{-1}\tilde{\boldsymbol{\theta}}_a - \boldsymbol{\Sigma}_b^{-1}\tilde{\boldsymbol{\theta}}_b = \boldsymbol{\Sigma}_a^{-1}\hat{\boldsymbol{\theta}}_a - \boldsymbol{\Sigma}_b^{-1}\hat{\boldsymbol{\theta}}_b \\ &\quad +\Delta\mathbf{g}(\hat{\boldsymbol{\theta}}_a) - \Delta\mathbf{g}(\hat{\boldsymbol{\theta}}_b) \\ &= \boldsymbol{\Sigma}_{ab}^{-1}\hat{\boldsymbol{\theta}}_{ab} + \Delta\mathbf{g}(\hat{\boldsymbol{\theta}}_{ab}).\end{aligned}$$

When these distributions are approximated by a Gaussian distribution around the true mean $\hat{\boldsymbol{\theta}}$ of the distribution, instead of $\tilde{\boldsymbol{\theta}}$ it can be shown that the ratio of these products is offset by an exponential product of $\exp[-2\Delta\mathbf{g}(\hat{\boldsymbol{\theta}}_{ab})^\top \boldsymbol{\theta}]$.

Appendix B. Divergence of the distributions

In [18], the author proposes the use of χ^2 and harmonic divergence to asses the accuracy of posterior approximations. The χ^2 divergence, defined as

$$\chi^2 = \int \frac{p^2(\mathbf{x}_t \mid y_{1:t})}{q(\mathbf{x}_t \mid y_{1:t})} d\mathbf{x}_t \quad (30)$$

measures the difference between the posterior distribution $p(\cdot)$ and its approximating density $q(\cdot)$. We have

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = cp(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = c\tilde{p}(\mathbf{x}_t \mid \mathbf{y}_{1:t}),$$

where c is a normalization constant. With a slight abuse of notation the χ^2 distance can now be expressed as follows

$$\chi^2 = \int \frac{p^2}{q} d\mathbf{x}_t = \int \frac{p^2}{q^2} \quad q \quad d\mathbf{x}_t$$

whose estimate can be obtained as

$$\hat{\chi}^2 = M \frac{\sum_{m=1}^M \left(\frac{\tilde{p}(\mathbf{x}_t^{(m)} \mid \mathbf{y}_{1:t})}{q(\mathbf{x}_t^{(m)} \mid y_{1:t})}\right)^2}{\left(\sum_{m=1}^M \frac{\tilde{p}(\mathbf{x}_t^{(m)} \mid \mathbf{y}_{1:t})}{q(\mathbf{x}_t^{(m)} \mid y_{1:t})}\right)^2}, \quad (31)$$

where $\mathbf{x}_t^{(m)} \sim q(\mathbf{x}_t \mid \mathbf{y}_{1:t})$.

References

1. Varshney, P.K.: Multisensor data fusion. Electron. Commun. Eng. J. **9**(6), 245–253 (1997)
2. Luo, R.C., Yih, C.C., Su, K.L.: Mutlisensor fusion integration: approaches applications, and future research direction. IEEE Sensors J. **2**(2), 107–119 (2002)
3. Anderson, D.O., Moore, J.B.: Optimal Filtering. Prentice-Hall, Inc. Englewood Cliffs (1979)
4. Bar-Shalom, Y., Fortmann, T.E.: Tracking and Data Association. Academic, New York (1988)
5. Chang, K.C., Chong, C.Y., Bar-Shalom, Y.: Distributed Estimation in Distributed Sensor Networks in Stochastic Large Scale Systems. M. Dekker (1992)
6. Chang, K.C., Saha, R.K., Bar-Shalom, Y.: On optimal track-to-track fusion. IEEE Trans. Aeros. Electron. Systems **33**(4), 1271–1276 (1997)
7. Doucet, A., de Freitas, N., Gordon, N.: Sequential Monte Carlo Methods in Practice. Springer, Heidelberg, New York (2001)
8. Míguez, J., Bugallo, M.F., Djurić, P.M.: A new class of particle filters for random dynamical systems with unknown statistics. EURASIP J. Appl. Sig. Process. **15**, 2278–2294 (2004)
9. Sheng, X., Hu, Y.-H., Ramanathan, P.: Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network. In: Fourth International Symposium on Information Processing in Sensor Networks, (IPSN 2005) vol. 15, April 2005, 181–188
10. Coates, M.: Distributed particle filters for sensor networks. In: Proceedings of the third International Symposium on Information Processing in Sensor Networks (IPSN-04), ACM Press, New York: April 26–27, 2004, pp. 99–107

11. Rosencrantz, M., Gordon, G.J., Thrun, S.: Decentralized sensor fusion with distributed particle filters. In: UAI '03, Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence, August 7–10, Acapulco, Mexico 493–500 (2003)
12. Challa, S., Palaniswami, M., Shilton, A.: Distributed data fusion using support vector machines. In: Proceedings of the Fifth International Conference on Information FUSION 2002, vol. 2, pp. 881–885 (2002)
13. Rey, W.J.J.: Introduction to Robust and Quasi–Robust Statistical Methods. Springer, Heidelberg, New York (1983)
14. Kotecha, J.H., Djurić, P.M.: Gaussian particle filtering, IEEE Trans. Sig. Process. **51**(10), 2592–2601 (2003)
15. McLachlan, G.J., Krishnan, T.: The EM Algorithm and Extensions. Wiley, New York (1997)
16. Gordon, N.J., Salmond, D.J., Smith, A.F.M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings-F **140**(2), 107–113 (1993)
17. Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R., Nordlund, P.: Particle filters for positioning, navigation and tracking. IEEE Trans. Sig. Process. **50**(2), 425–436 (2002)
18. Gilardoni, G.L.: Accuracy of posterior approximations via χ^2 and harmonic divergences. J. Statist. Plann. Inf. **128**(1), 475–487 (2005)