Signal Processing 90 (2010) 2731–2748

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

An ant stochastic decision based particle filter and its convergence

Benlian Xu [a,*], Jihong Zhu [b], Huigang Xu [a]

[a] School of Electric & Automatic Engineering, Changshu Institute of Technology, 215500 Changshu, China
[b] School of Automation, Nanjing University of Science & Technology, 210094 Nanjing, China

ARTICLE INFO

Article history:
Received 17 November 2009
Received in revised form
20 March 2010
Accepted 24 March 2010
Available online 27 March 2010

Keywords:
Particle filter
Ant colony optimization algorithm
Target tracking
Convergence analysis

ABSTRACT

Particle filter (PF) is a kind of flexible and powerful sequential Monte-Carlo technique designed to solve the optimal nonlinear parameter estimation numerically, and the degradation of particles in generic PF occurs when it is applied to the model switching dynamic system. To avoid this phenomenon, an ant stochastic decision based particle filter is proposed to encapsulate model switching information through dividing probabilistically particles into two model operations, and then a well defined re-sampling scheme is introduced to gain a better overlap with the true density function. To show the theoretic consistency with the generic PF, its basic convergence result is presented as well. Finally, we compare the performance of our proposed algorithm with that of other estimators (e.g., PF and moving ant estimator), and simulation results demonstrate its superior robustness of parameter estimation for switching dynamic system.

1. Introduction

In the recursive Bayesian framework, an unknown state can be derived from its posterior distribution. If the observations are received sequentially in time, the posterior distribution is then updated or evolves in time as well using measurement likelihood function and state transition function. When the posterior distribution has an analytical form, it is easy to implement the Bayesian recursion in a polite manner. However, such case only applies to few models, such as Gaussian and linear state space model. In other words, the posterior distribution generally has no analytical expression for large and practical engineering applications.

Particle filter (PF), also known as Sequential Monte-Carlo (SMC), is a kind of flexible and powerful sequential Monte-Carlo technique designed to solve the optimal nonlinear parameter estimation *numerically*, and it has received tremendous popularity in recent years over a wide range of applications [1–5]. The basic idea behind the PF is that the state density distribution is approximated by a group of "weighted" samples, and these samples are propagated and updated once new measurements are available. Using these "weighted" samples, the posterior state density distribution can be approximated by standard Monte-Carlo integration techniques.

It is acknowledged that a successful implementation of PF resorts to two aspects: (1) how to select appropriately samples (re-sampling), i.e., how to avoid the phenomenon of degeneration in which a number of samples are removed from the sample set due to the lower importance weights. Specially, *for a model switching dynamic system*, the phenomenon of particle degeneration is much more severe than other cases; (2) how to design an appropriate proposal distribution to facilitate easy sampling and further to achieve a large overlap with the true state density function.

So far, there are three widely used re-sampling techniques, namely, sampling-importance re-sampling (SIR), residual re-sampling, and minimum variance sampling [6]. Of the three strategies, minimum variance sampling has the smallest variance on the number of

* Corresponding author. Tel.: +08613913646026.
E-mail address: xu_benlian@yahoo.com.cn (B. Xu).

0165-1684/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2010.03.024

"children" of each particle ($N_i, i=1,2,\ldots,N$), while SIR has the largest one. However, they share the same computational complexity $O(N)$. Additionally, it is found that a specific choice of re-sampling scheme does not significantly affect the performance of parameter estimation [6]. Thus, we resort to the design of proposal distribution, and a commonly accepted choice is that it takes the same form as the assumed state evolution function due to its easy-sampling-implementation. However, such technique does not take advantage of posterior information such as new measurements, and if the assumed state evolution function deviates from the true one, it will lead to low likelihood value or weight and further result in divergence of estimation. To remedy this, many variants of PF have been proposed [7–9], such as the extended H_∞ filter based PF [7], the PF with functional auxiliary sampling density Kullback–Leibler (FASD-KL) [8], the threshold-based PF [9], and so on. Besides, evolutionary algorithms, such as Genetic Algorithm (GA), are incorporated into the generic PF, and some promising results are reported as well [10]. Inspired from ant stochastic behavior in foraging, we propose, in our work, an ant stochastic decision based particle filter, in which each particle evolves either of two proposal ways to accommodate model variations, and then particles are selected (re-sampling) according to ant empiricism acquired from available measurement, which results in a smaller particle variance than SIR, finally its convergence is investigated theoretically for a general use.

The remainder of this paper is organized as follows. In Section 2, we first briefly review the background on the particle filter, and then an ant stochastic decision based particle filter is described in detail. Section 3 gives the convergence proof of our proposed algorithm. Numerical simulation is carried out and results are analyzed in Section 4. Finally, conclusions are drawn in Section 5.

2. Ant stochastic behavior based particle filter

2.1. Background on particle filter

The objective of nonlinear filtering is to estimate the state according to the dynamic model

$$\boldsymbol{x}_{t+1}=f(\boldsymbol{x}_t,\boldsymbol{w}_t)$$

$$\boldsymbol{z}_t=h(\boldsymbol{x}_t,\boldsymbol{v}_t) \tag{1}$$

where $\boldsymbol{x}_t\in\mathbb{R}^{n_x}$ denotes the state at time t, $\boldsymbol{z}_t\in\mathbb{R}^{n_z}$ is the corresponding measurement, $\boldsymbol{w}_t$ and $\boldsymbol{v}_t$ represent the process and measurement noise, respectively. In the above formula, the system model is characterized by the mapping $f(\cdot)(\mathbb{R}^{n_x}\times\mathbb{R}^{n_w}\mapsto\mathbb{R}^{n_x})$, usually assumed to follow the first order Markov density function $p_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ on the state space $\chi\subseteq\mathbb{R}^{n_x}$, and the observation function $h(\cdot)(\mathbb{R}^{n_x}\times\mathbb{R}^{n_v}\mapsto\mathbb{R}^{n_z})$, usually modeled as a likelihood function $p_t(\boldsymbol{z}_t|\boldsymbol{x}_t)$ on the observation space $Z\subseteq\mathbb{R}^{n_z}$, respectively.

Given an initial density function $\pi_0(\boldsymbol{x}_0)$ of state, the posterior filtering density at time t can be calculated by the Bayesian recursion

$$\pi_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})=\int p_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{x})\pi_{t-1|t-1}(\boldsymbol{x}|\boldsymbol{z}_{1:t-1})\,d\boldsymbol{x} \tag{2}$$

$$\pi_{t|t}(\boldsymbol{x}_t|\boldsymbol{z}_{1:t})=\frac{p_t(\boldsymbol{z}_t|\boldsymbol{x}_t)\pi_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})}{\int p_t(\boldsymbol{z}_t|\boldsymbol{x})\pi_{t|t-1}(\boldsymbol{x}|\boldsymbol{z}_{1:t-1})\,d\boldsymbol{x}} \tag{3}$$

where $\boldsymbol{z}_{1:t}=\{\boldsymbol{z}_1,\boldsymbol{z}_2,\ldots,\boldsymbol{z}_t\}$ denotes the accumulated observations up to time t. Once the posterior probabilistic density function $\pi_{t|t}(\boldsymbol{x}_t|\boldsymbol{z}_{1:t})$ is available, the corresponding estimate of state $\boldsymbol{x}_t$ can be achieved by the maximum a posterior (MAP) estimator.

To implement the Bayesian recursion, a direct and easy alternative is the PF or Sequential Monte-Carlo technique. The PF algorithm is initialized with a set of un-weighted particles and then undergoes prediction step, update step and re-sampling step sequentially [9].

Initialization: One draws N samples equally from the initial density function $\pi_{0|0}$, and yields $\{x_0^{(i)},N^{-1}\}_{i=1}^N$. Let $\pi_{t-1|t-1}^N$ be a measure, so we define

$$\pi_{t-1|t-1}(d\boldsymbol{x}_{t-1}|\boldsymbol{z}_{1:t-1})\sim\pi_{t-1|t-1}^N(d\boldsymbol{x}_{t-1}|\boldsymbol{z}_{1:t-1})\triangleq\frac{1}{N}\sum_{i=1}^N\delta_{x_{t-1}^{(i)}}(d\boldsymbol{x}_{t-1}) \tag{4}$$

where $\pi_{t-1|t-1}^N$ denotes the empirical distribution close to $\pi_{t-1|t-1}$, and $\delta_{\boldsymbol{x}_{t-1}^{(i)}}(\cdot)$ denotes the delta-Dirac function centered at $x_{t-1}^{(i)}$. If we set $t=1$, then Eq. (4) corresponds to the initial step.

Prediction step: This step is carried out for each particle to obtain predicted particle $\overline{x}_t^{(i)}$ using the proposal density function

$$\overline{\boldsymbol{x}}_t^{(i)}\sim q_t(\cdot|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t)\quad i=1,2,\ldots,N \tag{5}$$

Meanwhile, the individual weight for each particle is predicted and evaluated by the law

$$w_{t|t-1}^{(i)}=\frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{q_t(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t)}w_{t-1}^{(i)} \tag{6}$$

Note that if we take $q_t(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t)=p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})$, then $w_{t|t-1}^{(i)}=w_{t-1}^{(i)}$. Let $\overline{\pi}_{t|t-1}^N$ be the measure, following Eq. (4), the empirical one-step predicted distribution is obtained

$$\pi_{t|t-1}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})\sim\overline{\pi}_{t|t-1}^N(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})\triangleq\sum_{i=1}^N w_{t|t-1}^{(i)}\delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t) \tag{7}$$

Update step: Once new measurement $\boldsymbol{z}_t$ is available, and if Eq. (7) is substituted into Eq. (3), we have the Monte-Carlo approximation of $\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t})$

$$\begin{aligned}\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t})&\sim\overline{\pi}_{t|t}^N(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t})\\&\triangleq\frac{p_t(\boldsymbol{z}_t|\boldsymbol{x}_t)\overline{\pi}_{t|t-1}^N(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})}{\int p_t(\boldsymbol{z}_t|\boldsymbol{x})\overline{\pi}_{t|t-1}^N(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})d\boldsymbol{x}}\\&=\frac{\sum_{i=1}^N\boldsymbol{w}_{t|t-1}^{(i)}p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})\delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t)}{\sum_{i=1}^N w_{t|t-1}^{(i)}p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})}\\&=\sum_{i=1}^N\left(\frac{w_{t|t-1}^{(i)}p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})}{\sum_{i=1}^N w_{t|t-1}^{(i)}p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})}\right)\delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t)\end{aligned} \tag{8}$$

The last term in Eq. (8) can be further written as a simplified form in terms of updated weight, thus Eq. (8) is reduced to

$$\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \sim \overline{\pi}^N_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \triangleq \sum_{i=1}^{N} w_t^{(i)} \delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t),$$

$$w_t^{(i)} = \frac{p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})}{\sum_{i=1}^{N} w_{t|t-1}^{(i)} p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})} w_{t|t-1}^{(i)} \tag{9}$$

Re-sampling step: Finally, the re-sampling step is done to obtain a set of un-weighted particles $\{x_t^{(i)}, N^{-1}\}_{i=1}^{N}$, and these particles constitute the empirical distribution close to $\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t})$

$$\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \sim \pi^N_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \triangleq \frac{1}{N}\sum_{i=1}^{N} \delta_{\boldsymbol{x}_t^{(i)}}(d\boldsymbol{x}_t) \tag{10}$$

So, the final estimate for $\mathbf{x}_t$ at time t can be written as

$$\hat{\boldsymbol{x}}_t = E(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \sum_{i=1}^{N} N^{-1} \boldsymbol{x}_t^{(i)} \tag{11}$$

Eqs. (4)–(10) are the numerically implementation of Bayesian recursion (Eqs. (2) and (3)), and this constitutes one cycle of generic PF algorithm. It can be observed that these particles $\overline{\boldsymbol{x}}_t^{(i)}$ obtained from proposal density function have a critical and direct effect on the update step and even the re-sampling step, since these following steps are done fully based on $\overline{\boldsymbol{x}}_t^{(i)}(i=1,2,\ldots,N)$. In addition, these particles $x_t^{(i)}$ obtained from the re-sampling step in turn affect the generation of particles in Eq. (5) at next time. Therefore, our proposed algorithm focuses on variations of the two steps with the help of ants, which differ conspicuously from the counterparts in PF.

2.2. Ant stochastic behavior based algorithm

In this section, the ant stochastic selection behavior, acquired from ant colony optimization [11–13], is incorporated into the above PF (called generic PF later) to improve the effectiveness and diversity of particles in different dynamic systems. The reason why we use the ant algorithm lies in:

(1) Ant probabilistic selection accommodates part of particles undergoing another Markov transition process, i.e., allows for various uncertainties such as model switching in time and noise level changes in time.
(2) The incorporation of ants into PF could ease the burden of re-sampling step, i.e., it can guide some particles toward interested region in advance. As a result, the degradation phenomenon of particles in re-sampling step is alleviated in some sense.

Since our algorithm is a variant of particle filter, it has the same steps as the PF, namely, prediction step, update step and re-sampling step. However, particles in prediction step may follow either the original transitional density function or a new one, and such selection is based on ants' stochastic decision. While in the re-sampling step, a direct or empirical re-sampling scheme is investigated, which guarantees a smaller variance of the number of particles and further results in a better overlap with the true density function.

Prediction step: To accommodate the possibility of variation during model evolution process, we consider two proposal density functions denoted by $q_t^{(1)}(\cdot|\cdot,\boldsymbol{z}_t)$ and $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$. For simplicity, for particle $\boldsymbol{x}_{t-1}^{(i)}$ we take $q_t^{(1)}(\cdot|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t) = p_{t|t-1}(\boldsymbol{x}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})$, which means that the state evolves in the pre-assumed model and is independent of new measurement $\boldsymbol{z}_t$; while $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$ is related to new measurement $\boldsymbol{z}_t$ and called posterior proposal density functions, which will be detailed later in this section. Note that the two proposal density function can be designed in other forms depending on applications. For instance, in the target tracking field, we employ Singer model [14] with different noise levels to adapt to model change such as from the non-maneuvering motion to maneuvering one, so $q_t^{(1)}(\cdot|\cdot,\boldsymbol{z}_t)$ and $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$ may take the Markov transition density functions corresponding to small and large process noise, respectively.

For a set of weighted particles $\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N}$, they are pre-divided by the ant's stochastic decision into two subgroups denoted by Ω_1 and Ω_2 according to the following law

$$\begin{cases} x_{t-1}^{(i)} \to \Omega_1 & \text{if } q \le q_0 \\ x_{t-1}^{(i)} \to \Omega_2 & \text{otherwise} \end{cases} \tag{12}$$

where q is a random number uniformly distributed in the range of [0,1], q_0 is a threshold which takes the value between 0 and 1. Parameter q_0 plays a role of threshold, and it determines the ratio of particles assigned into the two groups.

To improve the efficiency, we use N ants in parallel, each ant corresponds to a particle, but they use the same criterion q_0. It can be observed that the expectation of the number of particles in Ω_1 and Ω_2 is Nq_0 and $N(1-q_0)$, respectively. In addition, the setting of q_0 may affect the resulting performance, and its sensitiveness will be investigated in Section 4.

Therefore, the one-step prediction of each particle is done according to

$$\overline{x}_t^{(i)} \sim \begin{cases} q_t^{(1)}(\cdot|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t) & \text{if } x_{t-1}^{(i)} \in \Omega_1 \\ q_t^{(2)}(\cdot|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t) & \text{if } \boldsymbol{x}_{t-1}^{(i)} \in \Omega_2 \end{cases} \tag{13}$$

As discussed earlier, for those particles in Ω_1 we set $q_t^{(1)}(\cdot|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t) = p_{t|t-1}(\boldsymbol{x}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})$. Furthermore, if $p_{t|t-1}(\boldsymbol{x}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})$ is assumed to follow the Gaussian distribution, then it yields $\overline{\boldsymbol{x}}_t^{(i)} \sim p_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}^{(i)}) = \mathbb{N}(\boldsymbol{x}; f(\boldsymbol{x}_{t-1}^{(i)}), \mathbf{Q}_t)$ with mean $f(\boldsymbol{x}_{t-1}^{(i)})$ and variance $\mathbf{Q}_t$. Suppose that there are N_{t1}[1] particles in Ω_1 for a given iteration, so its corresponding

[1] Note that N_{t1} is time-dependent, i.e., the number of particle in space Ω_1 is time-varying, but the expectation of N_{t1} is invariant and equals to N_{q0}. Similarly, the above explanation applies to N_{t2}.

predicted weights

$$w_{t|t-1}^{(i)}=\frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{q_t^{(1)}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)},\boldsymbol{z}_t)}w_{t-1}^{(i)}=\frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}w_{t-1}^{(i)}=w_{t-1}^{(i)},$$
$$i=1,2,\ldots,N_{t1} \tag{14}$$

For the remaining N_{t2} (where $N_{t2}=N-N_{t1}$) particles in Ω_2, the posterior proposal density functions $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$ is designed as below. To take advantage of posterior information on new measurement $\boldsymbol{z}_t$, for particle $\boldsymbol{x}_{t-1}^{(i)}(i=N_{t1}+1,N_{t1}+2,\ldots,N)$, we choose the region close to $\boldsymbol{f}(\boldsymbol{x}_{t-1}^{(j)})$ as its intended destination, where $\boldsymbol{x}_{t-1}^{(j)}$ is determined by

$$\boldsymbol{x}_{t-1}^{(j)}=\arg\max_{\boldsymbol{x}_{t-1}^{(j)}\in\Omega_2}(p_t(\boldsymbol{z}_t|\boldsymbol{f}(\boldsymbol{x}_{t-1}^{(j)}))) \tag{15}$$

Now the state of particle $\boldsymbol{x}_{t-1}^{(i)}$ is perturbed and changed to $\boldsymbol{y}\triangleq\xi\boldsymbol{x}_{t-1}^{(j)}$ which is close to $\boldsymbol{x}_{t-1}^{(j)}$, where ξ is assumed to be a Gaussian stochastic variable with mean unity and variance σ^2. Apparently, the standard deviation σ determines the degree to which $\xi\boldsymbol{x}_{t-1}^{(j)}$ is close to $\boldsymbol{x}_{t-1}^{(j)}$. As a result, the posterior proposal density functions $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$ can be written as

$$q_t^{(2)}(\boldsymbol{x}_t|\cdot,\boldsymbol{z}_t)=q_t^{(2)}(\boldsymbol{x}_t|\boldsymbol{z}_t)=\int p_{t|t-1}(\boldsymbol{x}_t|\xi\boldsymbol{x}_{t-1}^{(j)})p(\xi\boldsymbol{x}_{t-1}^{(j)}|\boldsymbol{z}_t)d(\xi\boldsymbol{x}_{t-1}^{(j)})$$
$$=\int p_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{y})p(\boldsymbol{y})\,d\boldsymbol{y} \tag{16}$$

Since ξ is a Gaussian random variable with mean unity and variance σ^2, then $p(\boldsymbol{y}=\xi\boldsymbol{x}_{t-1}^{(j)})$ is also a Gaussian density function with $p(\boldsymbol{y}=\xi\boldsymbol{x}_{t-1}^{(j)})=\mathbb{N}(\boldsymbol{y};\boldsymbol{x}_{t-1}^{(j)},\boldsymbol{Q}_{1t})$, where $\boldsymbol{Q}_{1t}=\sigma^2 diag((\boldsymbol{x}_{t-1}^{(j,1)})^2,\ldots,(\boldsymbol{x}_{t-1}^{(j,n_x)})^2)$ with $\boldsymbol{x}_{t-1}^{(j,i)}$ denoting the ith component of vector $\boldsymbol{x}_{t-1}^{(j)}$. And if $p_{t|t-1}(\boldsymbol{x}_t|\xi\boldsymbol{x}_{t-1}^{(j)})$ is assumed to be a Gaussian function, thus it can be formulated as $p_{t|t-1}(\boldsymbol{x}_t|\xi\boldsymbol{x}_{t-1}^{(j)})=\mathbb{N}(\boldsymbol{x}_t;\boldsymbol{f}(\xi\boldsymbol{x}_{t-1}^{(j)}),\boldsymbol{Q}_t)$. As a result, the integral form in Eq. (16) can be written as

$$\int p_{t|t-1}(\boldsymbol{x}_t|\xi\boldsymbol{x}_{t-1}^{(j)})p(\xi\boldsymbol{x}_{t-1}^{(j)})d(\xi\boldsymbol{x}_{t-1}^{(j)})=\int\mathbb{N}(\boldsymbol{x}_t;\boldsymbol{f}(\boldsymbol{y}),\boldsymbol{Q}_t)\mathbb{N}(\boldsymbol{y};\boldsymbol{x}_{t-1}^{(j)},\boldsymbol{Q}_{1t})\,d\boldsymbol{y} \tag{17}$$

Lemma 1. *Assume that there are three Gaussian density functions denoted by* $p_1(\boldsymbol{v})=\mathbb{N}(\boldsymbol{v};\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x},\boldsymbol{R})$, $p_2(\boldsymbol{x})=\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})$, *and* $p_3(\boldsymbol{z})=\mathbb{N}(\boldsymbol{z};\boldsymbol{H}\boldsymbol{m},\boldsymbol{R}+\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^T)$, *respectively. And if* $\boldsymbol{R}$ *and* $\boldsymbol{P}$ *are positive definite, then we have*

$$\frac{p_1(\boldsymbol{v})p_2(\boldsymbol{x})}{p_3(\boldsymbol{z})}=\frac{\mathbb{N}(\boldsymbol{v};\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x},\boldsymbol{R})\cdot\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})}{\mathbb{N}(\boldsymbol{z};\boldsymbol{H}\boldsymbol{m},\boldsymbol{R}+\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^T)}$$
$$=\frac{\mathbb{N}(\boldsymbol{z};\boldsymbol{H}\boldsymbol{x},\boldsymbol{R})\cdot\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})}{\mathbb{N}(\boldsymbol{z};\boldsymbol{H}\boldsymbol{m},\boldsymbol{R}+\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^T)}=c\mathbb{N}(\boldsymbol{x};\overline{\boldsymbol{m}},\overline{\boldsymbol{P}}) \tag{18}$$

with $\overline{\boldsymbol{m}}=\boldsymbol{m}+\boldsymbol{P}\boldsymbol{H}^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{H}\boldsymbol{m})$, $\overline{\boldsymbol{P}}=\boldsymbol{P}-\boldsymbol{P}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^T+\boldsymbol{R})^{-1}\boldsymbol{H}\boldsymbol{P}$, *and* c *is a constant and equals to* $c=(|\overline{\boldsymbol{P}}|^{1/2}|\boldsymbol{R}+\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^T|^{1/2})/|\boldsymbol{R}|^{1/2}|\boldsymbol{P}|^{1/2}$. (*See the proof in Appendix A*).

Lemma 2. *Given the following matrixes or vectors* $\boldsymbol{F},\boldsymbol{Q},\boldsymbol{m}$ *and* $\boldsymbol{P}$, *each with appropriate dimension, and if* $\boldsymbol{Q}$ *and* $\boldsymbol{P}$ *are positive definite, we have*

$$\int\mathbb{N}(\boldsymbol{x};\boldsymbol{F}\boldsymbol{y},\boldsymbol{Q})\mathbb{N}(\boldsymbol{y};\boldsymbol{m},\boldsymbol{P})\,d\boldsymbol{y}=\tilde{c}\,\mathbb{N}(\boldsymbol{x};\boldsymbol{F}\boldsymbol{m},\boldsymbol{Q}+\boldsymbol{F}\boldsymbol{P}\boldsymbol{F}^T) \tag{19a}$$

Proof. Using the result of Lemma 1, the right side of Eq. (19a) can be rewritten as

$$\int\mathbb{N}(\boldsymbol{x};\boldsymbol{F}\boldsymbol{y},\boldsymbol{Q})\mathbb{N}(\boldsymbol{y};\boldsymbol{m},\boldsymbol{P})\,d\boldsymbol{y}$$
$$=\int\mathbb{N}(\boldsymbol{x};\boldsymbol{F}\boldsymbol{m},\boldsymbol{Q}+\boldsymbol{F}\boldsymbol{P}\boldsymbol{F}^T)\cdot\tilde{c}\,\mathbb{N}(\boldsymbol{y};\overline{\boldsymbol{m}}',\overline{\boldsymbol{P}}')\,d\boldsymbol{y}$$
$$=\tilde{c}\,\mathbb{N}(\boldsymbol{x};\boldsymbol{F}\boldsymbol{m},\boldsymbol{Q}+\boldsymbol{F}\boldsymbol{P}\boldsymbol{F}^T)\int\mathbb{N}(\boldsymbol{y};\overline{\boldsymbol{m}}',\overline{\boldsymbol{P}}')\,d\boldsymbol{y}$$
$$=\tilde{c}\,\mathbb{N}(\boldsymbol{x};\boldsymbol{F}\boldsymbol{m},\boldsymbol{Q}+\boldsymbol{F}\boldsymbol{P}\boldsymbol{F}^T) \tag{19b}$$

where $\tilde{c}$ is a constant. Using Lemma 2, Eq. (16) or (17) can be formulated by

$$q_t^{(2)}(\boldsymbol{x}_t|\cdot,\boldsymbol{z}_t)=\int p_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{y})p(\boldsymbol{y})\,d\boldsymbol{y}$$
$$=\int\mathbb{N}(\boldsymbol{x}_t;\boldsymbol{f}(\boldsymbol{y}),\boldsymbol{Q}_t)\mathbb{N}(\boldsymbol{y};\boldsymbol{x}_{t-1}^{(j)},\boldsymbol{Q}_{1t})\,d\boldsymbol{y}$$
$$=\hat{c}\,\mathbb{N}(\boldsymbol{x}_t;\boldsymbol{F}\boldsymbol{x}_{t-1}^{(j)},\boldsymbol{Q}_t+\boldsymbol{F}\boldsymbol{Q}_{1t}\boldsymbol{F}^T) \tag{20}$$

where

$$\boldsymbol{F}=\frac{\partial\boldsymbol{f}}{\partial\boldsymbol{y}}\Big|_{\boldsymbol{y}=\xi\boldsymbol{x}_{t-1}^{(j)}}$$

and $\hat{c}$ is a constant which is equal to

$$\hat{c}=\left(\frac{|\boldsymbol{Q}_t+\boldsymbol{F}\boldsymbol{Q}_{1t}\boldsymbol{F}^T||\boldsymbol{Q}_{1t}-\boldsymbol{Q}_{1t}\boldsymbol{F}^T(\boldsymbol{Q}_t+\boldsymbol{F}\boldsymbol{Q}_{1t}\boldsymbol{F}^T)^{-1}\boldsymbol{F}\boldsymbol{Q}_{1t}|}{|\boldsymbol{Q}_t||\boldsymbol{Q}_{1t}|}\right)^{1/2}$$

So far, we have given the analytic formulations of the two proposal density functions, which are characterized by the easy-to-sampling, i.e., both in the form of Gaussian functions.

So, following Eq. (14), for those particles in Ω_2,

$$w_{t|t-1}^{(i)}=\frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{q_t^{(2)}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)},z_t)}w_{t-1}^{(i)},i=\overbrace{N_{t1}+1,N_{t1}+2,\ldots,N}^{N_{t2}} \tag{21}$$

Let $\overline{\pi}_{t|t-1}^{N_{t1}}$ and $\overline{\pi}_{t|t-1}^{N_{t2}}$ be the measures, the respective empirical one-step predicted distributions are obtained

$$\overline{\pi}_{t|t-1}^{N_{t1}}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})\triangleq\sum_{i=1}^{N_{t1}}w_{t|t-1}^{(i)}\delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t)$$

$$\overline{\pi}_{t|t-1}^{N_{t2}}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})\triangleq\sum_{i=N_{t1}+1}^{N}w_{t|t-1}^{(i)}\delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t) \tag{22}$$

Moreover, we let

$$\pi_{t|t-1}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})\sim\overline{\pi}_{t|t-1}^{N=N_{t1}+N_{t2}}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})$$
$$=\overline{\pi}_{t|t-1}^{N_{t1}}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1})+\overline{\pi}_{t|t-1}^{N_{t2}}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1}) \tag{23}$$

Update step: This step is done in the similar way as the generic PF, and when the new measurement arrives, the weight of each particle is updated

$$w_t^{(i)}=\frac{p_t(z_t|\overline{\boldsymbol{x}}_t^{(i)})}{\sum_{i=1}^{N}w_{t|t-1}^{(i)}p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})}w_{t|t-1}^{(i)} \tag{24}$$

Let $\overline{\pi}^N_{t|t}$ be the measure, we have

$$\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \sim \overline{\pi}^N_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \triangleq \sum_{i=1}^{N} w_t^{(i)} \delta_{\overline{\boldsymbol{x}}_t^{(i)}}(d\boldsymbol{x}_t) \tag{25}$$

Re-sampling step: As discussed earlier, sampling-importance re-sampling, residual re-sampling, and minimum variance sampling are the currently widely used re-sampling schemes. Although each has different variances on the number of "children" of each particle, these schemes have no significant effects on the performance of PF. In our algorithm, the ant's stochastic behavior will be utilized to develop an easy-to-implement re-sampling scheme, which could yield a smaller variance on the number of "children" of each particle. Besides this, to the best of our knowledge, the ant's stochastic behavior has been investigated to develop various parameter estimators, such as moving ant estimator in [15] and ACO-estimator in [16]. The former is a recursive estimation technique suitable for real-time target tracking, while the latter is based on the time-consuming batch processing technique.

Next, we will describe our re-sampling scheme inspired from the ant's stochastic probabilistic behavior. As shown in Fig. 1, an ant samples sequentially particles starting from the first particle to the last particle, i.e., from the right to the left. Note that the ant will face N selections or operations, and for each selection, there are two cases of being sampled for a given particle i, namely, the ant is now either located at particle i or at other particle j ($j \neq i$).

Case 1: Assume that an ant is located at particle i, for the remaining $N-1$ particles, the ant first selects particle j as its potential location according to

$$\overline{P}(j) = \frac{w_t^{(j)}}{\sum_{k=1,k\neq i}^{N} w_t^{(k)}} \tag{26}$$

Then the ant will make a decision whether it move towards particle j or not. However, it decides to stay its original position with probability

$$P(i) = \frac{w_t^{(i)}}{w_t^{(i)} + w_t^{(j)}} \tag{27}$$

Eqs. (26) and (27) describe the decision probability of particle i being selected or sampled for a given sampling, and the resulting probability is denoted by

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{w_t^{(i)} + w_t^{(j)}} \frac{w_t^{(j)}}{\sum_{k=1,k\neq i}^{N} w_t^{(k)}} \tag{28}$$

Case 2: If the ant is currently located at particle j ($j \neq i$), then the ant selects particle i as its potential location according to

$$\overline{P}(i) = \frac{w_t^{(i)}}{\sum_{k=1,k\neq j}^{N} w_t^{(k)}} \tag{29}$$

Thus, particle i is to be chosen or sampled by

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{w_t^{(i)} + w_t^{(j)}} \frac{w_t^{(i)}}{\sum_{k=1,k\neq i}^{N} w_t^{(k)}} \tag{30}$$

Without loss of generality, we now detail the expectation of particle i being sampled during N selection tries of ant, and the discussion on other remaining particles performs in a similar way.

This stochastic decision process of ant behaves similarly to the event that particle i is re-sampled in SIR. However, the probability of particle i being selected is less than the original one done in SIR, i.e., $\tilde{w}_t^{(i)} \leq w_t^{(i)}$ (See the proof in Appendix B), which further results in a smaller variance of particles $\mathrm{var}(N_i)$, in the form of $\mathrm{var}(N_i) = Nw_t^{(i)}(1-w_t^{(i)})$, than that in SIR. The direct benefit is that the obtained empirical distribution has a better overlap with the true one due to $\tilde{w}_t^{(i)} \leq w_t^{(i)}$. Furthermore, it is observed that the computational complexity remains invariant and is $O(N)$.

When all particles move towards their individual positions under ants' guidance, we obtain a set of un-weighted particles $\{x_t^{(i)}, N^{-1}\}_{i=1}^{N}$, and these particles constitute the empirical distribution

$$\pi_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \sim \pi^N_{t|t}(d\boldsymbol{x}_t|\boldsymbol{z}_{1:t}) \triangleq \frac{1}{N}\sum_{i=1}^{N} \delta_{\boldsymbol{x}_t^{(i)}}(d\boldsymbol{x}_t) \tag{31}$$

So far, we have presented a whole cycle of ant stochastic decision based particle filter, and its corresponding main steps are summarized in Table 1.

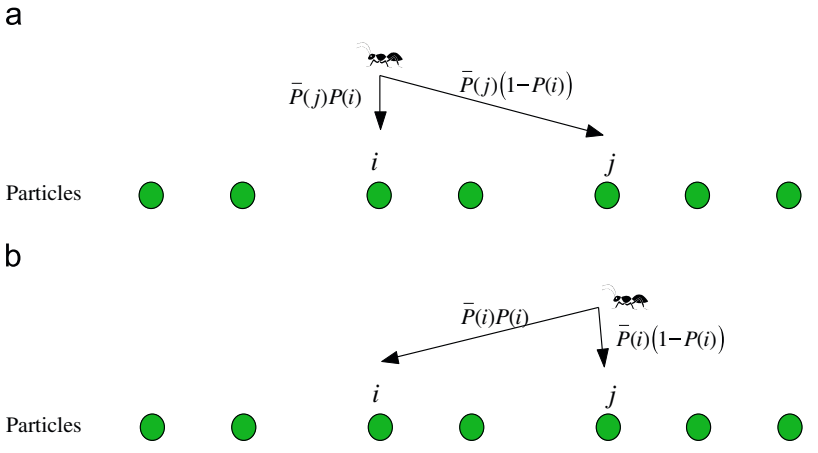

Fig. 1. The ant decision process. (a) Case 1 and case 2.

Table 1
The framework of our ant stochastic decision based PF.

Step (1) Initialization: sample N particles $\{x_0^{(i)}, N^{-1}\}_{i=1}^{N}$ from $\pi_0(\boldsymbol{x}_0)$;

Step (2) Prediction: first, divide N particles into two groups Ω_1 and Ω_2, and then obtain one-step prediction of each particle according to

$$\overline{\boldsymbol{x}}_t^{(i)} \sim \begin{cases} q_t^{(1)}(\cdot|\boldsymbol{x}_{t-1}^{(i)}, \boldsymbol{z}_t) & \text{if } \boldsymbol{x}_{t-1}^{(i)} \in \Omega_1 \\ q_t^{(2)}(\cdot|\boldsymbol{x}_{t-1}^{(i)}, \boldsymbol{z}_t) & \text{if } \boldsymbol{x}_{t-1}^{(i)} \in \Omega_2 \end{cases}$$

Step (3) Update: according to obtained one-step prediction of weight, using the available measurement to update the weight of each particle

$$w_{t|t-1}^{(i)} = \frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{q_t^{(1)}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)}, \boldsymbol{z}_t)} w_{t-1}^{(i)} = \frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})} w_{t-1}^{(i)} = w_{t-1}^{(i)} \quad \text{if } \boldsymbol{x}_{t-1}^{(i)} \in \Omega_1$$

$$w_{t|t-1}^{(i)} = \frac{p_{t|t-1}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{q_t^{(2)}(\overline{\boldsymbol{x}}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)}, \boldsymbol{z}_t)} w_{t-1}^{(i)}, \quad \text{if } \boldsymbol{x}_{t-1}^{(i)} \in \Omega_2$$

$$w_t^{(i)} = \frac{p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})}{\sum_{i=1}^{N} w_{t|t-1}^{(i)} p_t(\boldsymbol{z}_t|\overline{\boldsymbol{x}}_t^{(i)})} w_{t|t-1}^{(i)} \quad i = 1,2,\ldots,N$$

Step (4) Re-sample: particle i is to be chosen or sampled by ant stochastic behavior

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{w_t^{(i)} + w_t^{(j)}} \frac{w_t^{(j)}}{\sum_{k=1,k\neq i}^{N} w_t^{(k)}} \quad \text{for case 1}$$

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{w_t^{(i)} + w_t^{(j)}} \frac{w_t^{(i)}}{\sum_{k=1,k\neq i}^{N} w_t^{(k)}} \quad \text{for case 2}$$

Step (5) Repeat from Step (2) until the last sampling time.

3. Convergence analysis

In this section, to verify the effectiveness of our algorithm in theory, we will discuss its convergence in terms of the average mean-square error $E[(\langle \pi_{t|t}^N, \varphi \rangle - \langle \pi_{t|t}, \varphi \rangle)^2]$ for any function $\varphi \in B(\mathbb{R}^{n_x})$, where $B(\mathbb{R}^{n_x})$ is the set of bounded Borel measurable function on $\mathbb{R}^{n_x}$, and $\langle \pi_{t|t}, \varphi \rangle$ denotes the inner product and has the form

$$\langle \pi_{t|t}, \varphi \rangle = \int \pi_{t|t}(\boldsymbol{x}_t|\boldsymbol{z}_{1:t})\varphi(\boldsymbol{x}_t)\, d\boldsymbol{x}_t \tag{32}$$

Inspired from literature [17,18], we will prove the convergence of our proposed algorithm. Let $\{\mu^N\}_{N=1}^{\infty}$ be a sequence of probability measures, we say that μ^N converges to μ as $N \to \infty$ for $\forall \varphi \in B(\mathbb{R}^{n_x})$, and if the following conditions are satisfied

(1) The state transition function $p_{t|t-1}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ is a bounded one and Feller;
(2) The likelihood function $p_t(\boldsymbol{z}_t|\boldsymbol{x}_t)$ is bounded;
(3) Assume that the importance sampling ratios concerning $q_t^{(1)}$ and $q_t^{(2)}$ are bounded, namely, $\|p_{t|t-1}/q_t^{(1)}\| \leq C1$, $\|p_{t|t-1}/q_t^{(2)}\| \leq C2$, where $C1$ and $C2$ are real constant values.

Then we have three lemmas as below:

Lemma 3. *Suppose that for any $\varphi \in B(\mathbb{R}^{n_x})$, if $E[(\langle \pi_{t-1|t-1}^N, \varphi \rangle - \langle \pi_{t-1|t-1}, \varphi \rangle)^2] \leq (a_{t-1|t-1}/N)\|\varphi\|^2$ holds, there exists a constant $a_{t|t-1}$ such that*

$$E[(\langle \overline{\pi}_{t|t-1}^N, \varphi \rangle - \langle \pi_{t|t-1}, \varphi \rangle)^2] \leq \frac{a_{t|t-1}}{N}||\varphi||^2 \tag{33}$$

Lemma 4. *Suppose that for any $\varphi \in B(\mathbb{R}^{n_x})$, if $E[(\langle \overline{\pi}_{t|t-1}^N, \varphi \rangle - \langle \pi_{t|t-1}, \varphi \rangle)^2] \leq (a_{t|t-1}/N)\|\varphi\|^2$ holds, then there exists a constant $\overline{a}_{t|t}$ such that*

$$E[(\langle \overline{\pi}_{t|t}^N, \varphi \rangle - \langle \pi_{t|t}, \varphi \rangle)^2] \leq (\overline{a}_{t|t}/N)\|\varphi\|^2 \tag{34}$$

Lemma 5. *Suppose that for any $\varphi \in B(\mathbb{R}^{n_x})$, if $E[(\langle \overline{\pi}_{t|t}^N, \varphi \rangle - \langle \pi_{t|t}, \varphi \rangle)^2] \leq (\overline{a}_{t|t}/N)\|\varphi\|^2$, then there exists a constant $a_{t|t}$ such that*

$$E[(\langle \pi_{t|t}^N, \varphi \rangle - \langle \pi_{t|t}, \varphi \rangle)^2] \leq (a_{t|t}/N)\|\varphi\|^2 \tag{35}$$

where $\|\varphi\|$ denotes the suprenum norm of φ on $B(\mathbb{R}^{n_x})$. For the proof of each lemma, readers are referred to Appendix C.

4. Simulation and results

In this section, the tracking ability of our proposed algorithm is investigated through various tracking examples including linear and nonlinear cases of system motion model, and corresponding performance comparison is done among the generic PF, moving ant estimator [15], and our proposed algorithm. Moreover, all examples are conducted on a DELL 6 GHz processor with 1.99 GB RAM.

Consider a parameter estimation in the field of bearings-only tracking (BOT) of a bistatic system, namely, we assume that two observers are utilized to measure the target's bearing at each sampling time, and the state of observer $s(s=1,2)$ is denoted by $\boldsymbol{X}_t^{os} = [x_t^{os}, 0, y_t^{os}, 0]^T$, thus the measurement equation is described as:

$$\boldsymbol{\beta}_t^M = \begin{bmatrix} \beta_t^{M,1} \\ \beta_t^{M,1} \end{bmatrix} = \boldsymbol{h}(\boldsymbol{x}_t, \boldsymbol{v}_t) = \begin{pmatrix} \tan^{-1}\left(\dfrac{x_t - x_t^{o1}}{y_t - y_t^{o1}}\right) \\ \tan^{-1}\left(\dfrac{x_t - x_t^{o2}}{y_t - y_t^{o2}}\right) \end{pmatrix} + \boldsymbol{v}_t \tag{36}$$

where $\boldsymbol{\beta}_t^M$ denotes the bearing measurement vector, and $\boldsymbol{v}_t$ is the measurement Gaussian white noise, assumed to be zero mean with a 2×2 covariance $\boldsymbol{R}$.

4.1. Linear example

Suppose that the target moves in a straight line, so its motion model can be formulated as

$$\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{w}_t) = \boldsymbol{F}\boldsymbol{x}_{t-1} + \boldsymbol{w}_t \tag{37}$$

where $\boldsymbol{F}$ denotes the state transition matrix and equals to

$$\boldsymbol{F} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with sampling interval T, and $\boldsymbol{w}_t$ is the system process Gaussian white noise assumed to be zero mean with 4×4 covariance $\boldsymbol{Q}(t)$.

- If the two observers are fixed at (0,0) and $(-2\,\text{km},0)$, respectively, in a surveillance region, each having an

identical measurement noise 0.02^0, and the sampling interval is set to $T=1$ s, The target makes a uniform rectilinear motion, lasting for 100 samples, with a given initial state $\boldsymbol{x}_0=[0\,\text{m},45\,\text{m/s},14816\,\text{m},-30\,\text{m/s}]^T$ and process noise covariance $\boldsymbol{Q}(t)=\ diag(5^2,5^2,5^2,5^2)$.

- Concerning the PF and our ant based PF, we use $N=200$ particles or ants, $q_0=0.9$ and $\sigma=20$.

Given the initial estimate of state and covariance, i.e., $\hat{\boldsymbol{x}}_0=[0.1\,\text{km},50\,\text{m/s},16\,\text{km},-20\,\text{m/s}]^T$ and $\hat{\boldsymbol{P}}_0=diag(1000,$

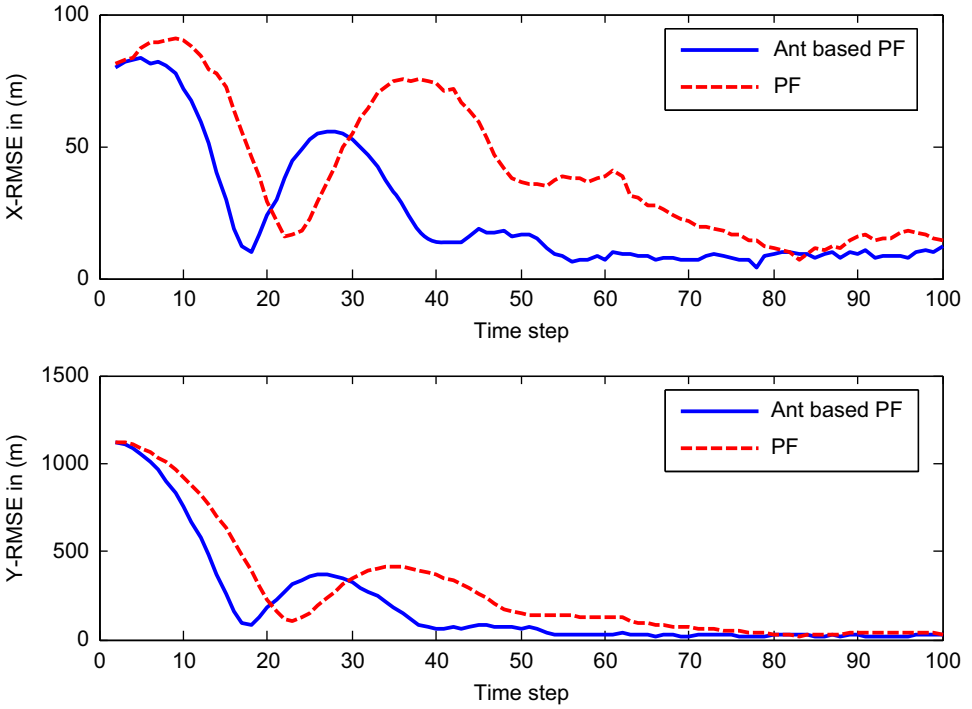

Fig. 2. The obtained RMSE curves by PF and ant based PF.

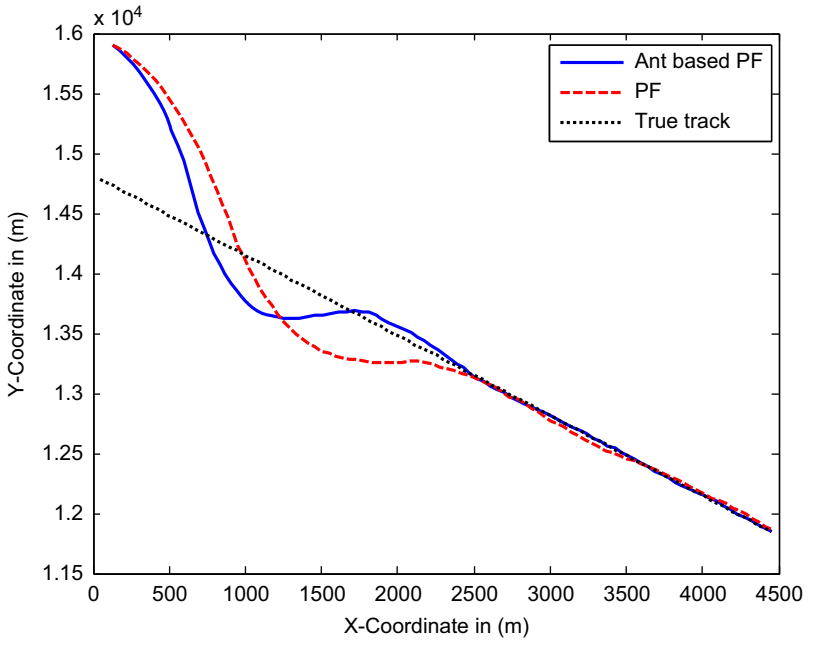

Fig. 3. The estimated tracks by PF and ant based PF.

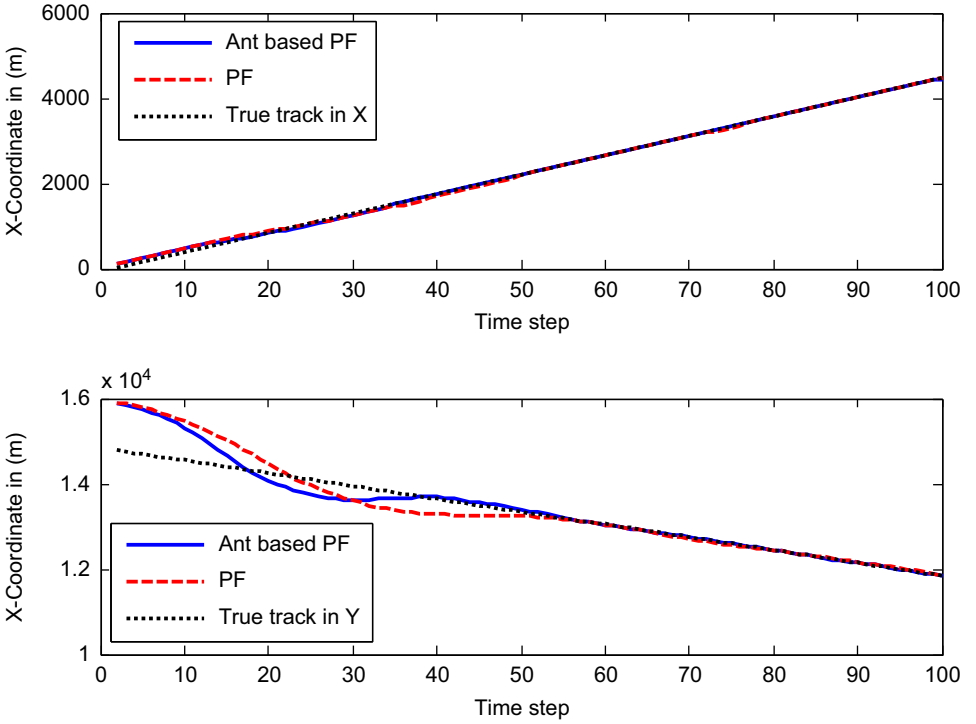

Fig. 4. The estimated positions in x and y directions by PF and ant based PF.

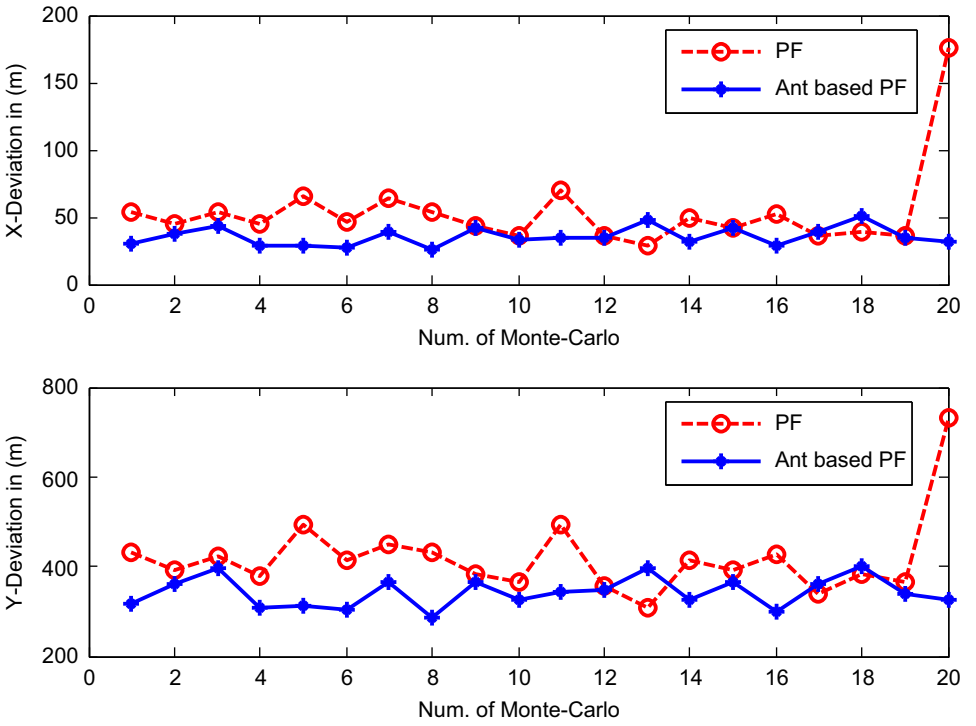

Fig. 5. The obtained deviation in x and y directions by PF and ant based PF.

50,1000,50), it yields the following results based on the average of 20 Monte-Carlo runs.

Fig. 2 plots the root mean squared error (RMSE) curves, and it can be observed that our ant based PF achieves smaller errors than the generic PF after the 30th sampling. Fig. 3 presents a direct performance comparison in terms of the obtained tracks, and our proposed algorithm fits the true track better than the generic PF. Additionally, Fig. 4 illustrates the individual estimated positions in the x and y directions, respectively, and the above conclusions applies to this case as well. To show the enjoyable performance of our algorithm, we give the deviation

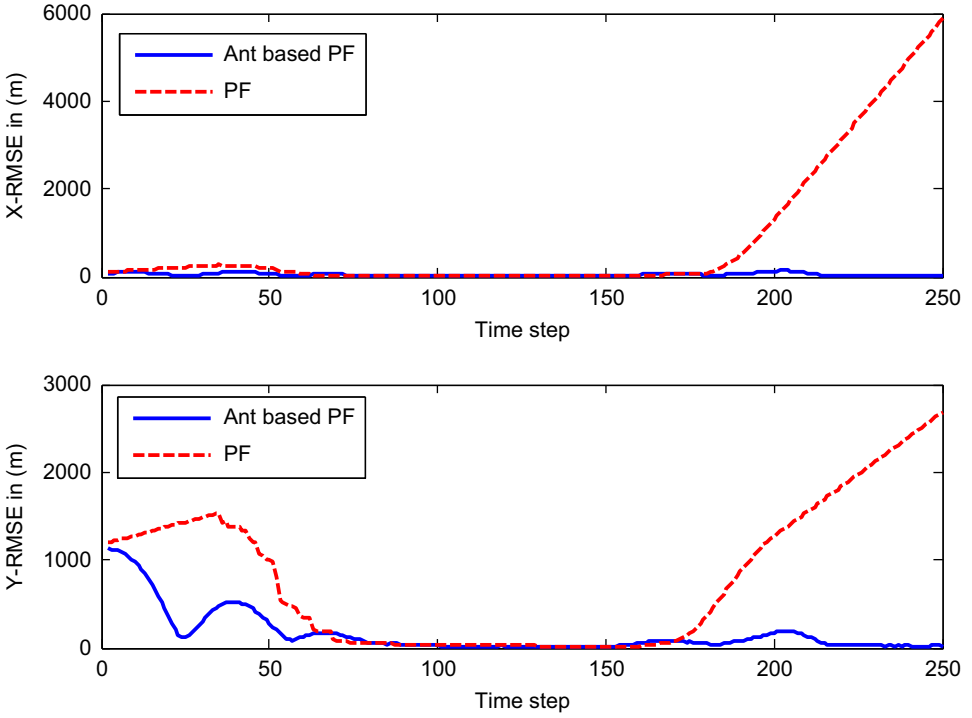

Fig. 6. The obtained RMSE curves by PF and ant based PF.

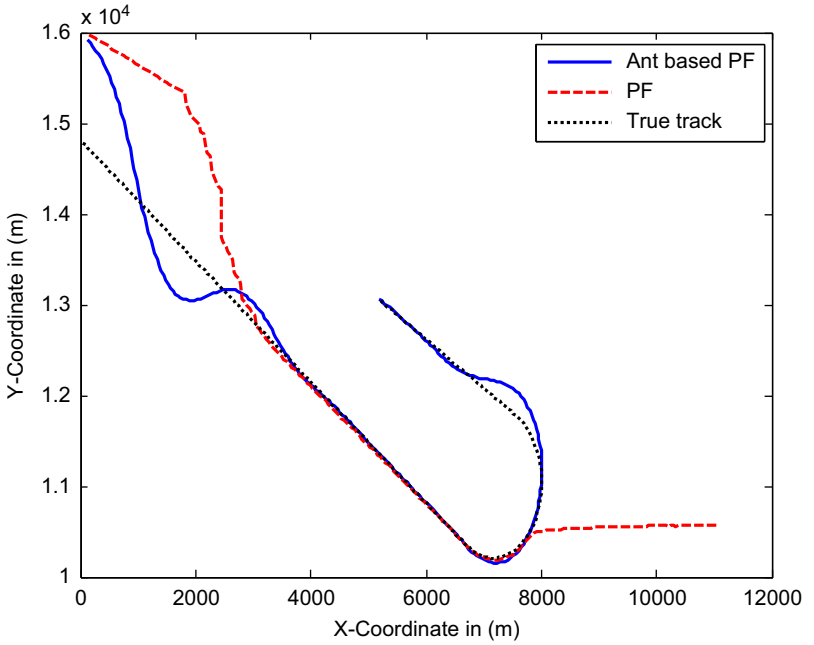

Fig. 7. The estimated tracks by PF and ant based PF.

values over 20 Monte-Carlo runs in the x and y directions, respectively, as shown in Fig. 5. Note that the deviation is defined as $Dev^{(i)} = std(\tilde{\boldsymbol{X}}^{(i)})$ with $\tilde{\boldsymbol{X}}^{(i)} = [\hat{\boldsymbol{x}}_1^{(i)} - \boldsymbol{x}_1, \hat{\boldsymbol{x}}_2^{(i)} - \boldsymbol{x}_2, \ldots]^T$, where superscript i denotes the ith Monte-Carlo run, $\hat{\boldsymbol{x}}_t^{(i)}$ represents the estimated state at time t at the ith Monte-Carlo run, and $\boldsymbol{x}_t$ is the true state at time t. It can be observed that our ant based PF results in a smaller deviation at most runs.

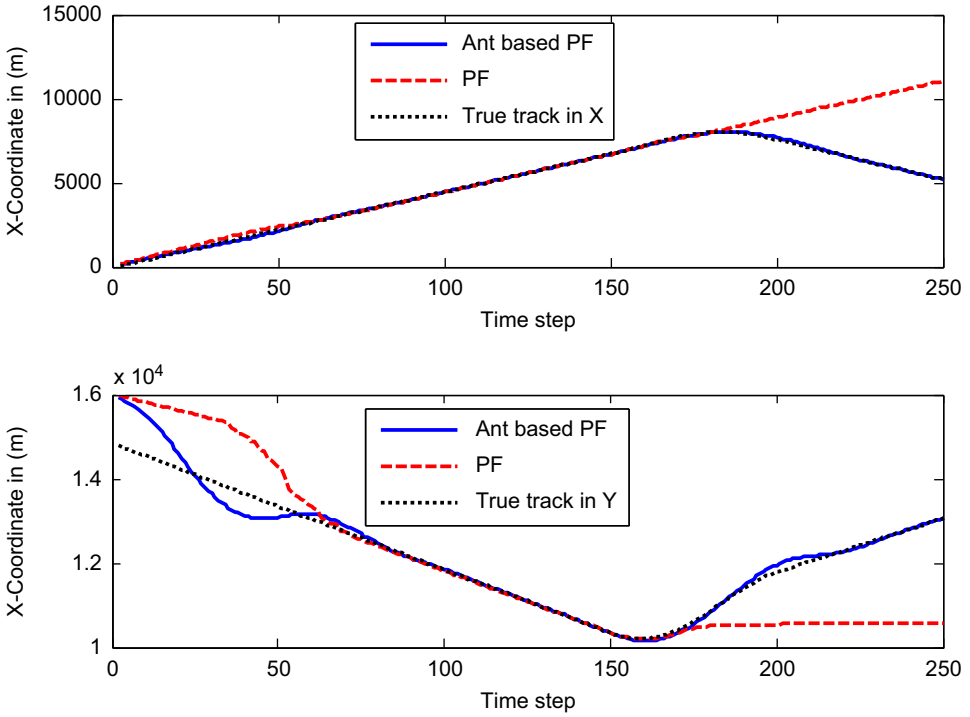

Fig. 8. The estimated positions in x and y directions by PF and ant based PF.

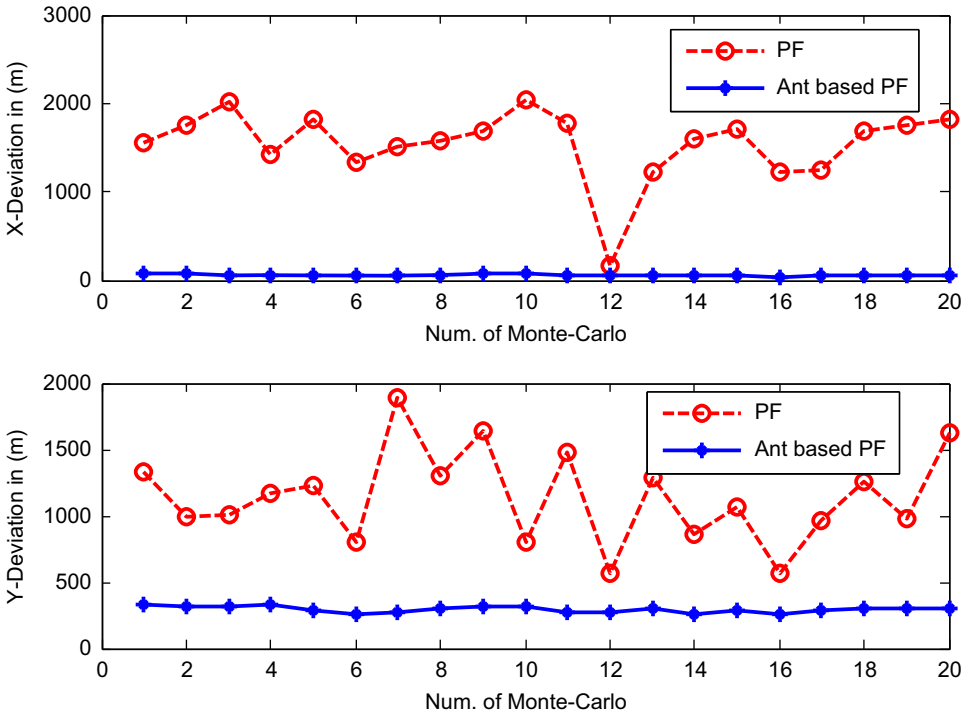

Fig. 9. The obtained deviation in x and y directions by PF and ant based PF.

4.2. Non-linear example

We will extend our algorithm to the non-linear case and investigate its resulting performance, which is our main purpose to develop the ant based PF.

Suppose that the target follows the dynamic model as Eq. (37) during the first 150 samples phase with the same initial conditions as the linear example, and then the target executes a constant rate turn with $a_n = 3.5\,\mathrm{m/s^2}$, which lasts 50 samples. Note that during the turning

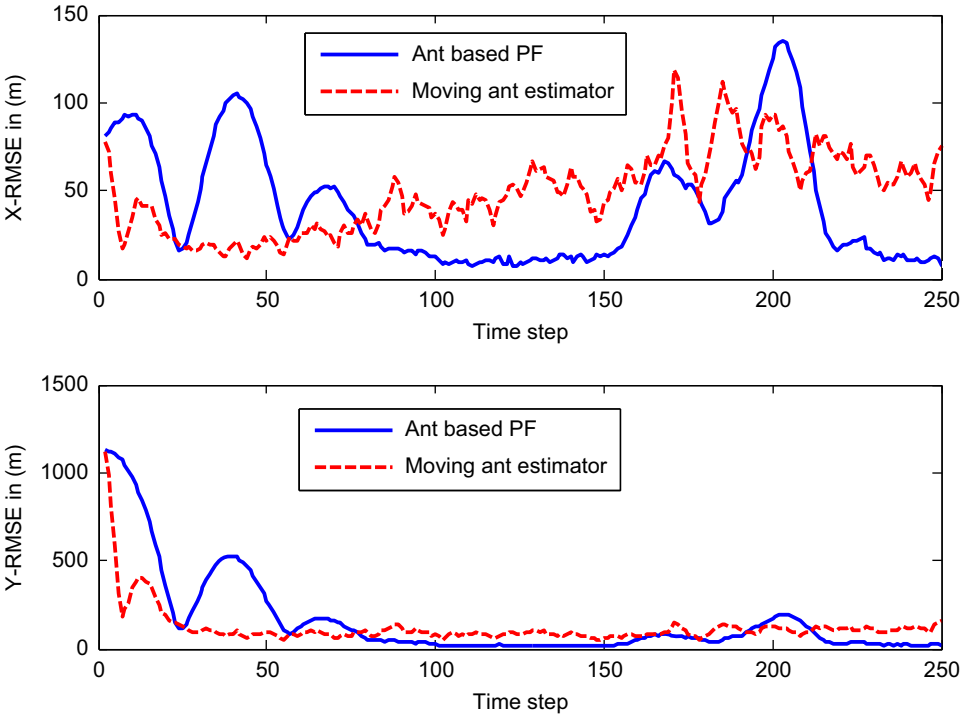

Fig. 10. The obtained RMSE curves by moving ant estimator and ant based PF.

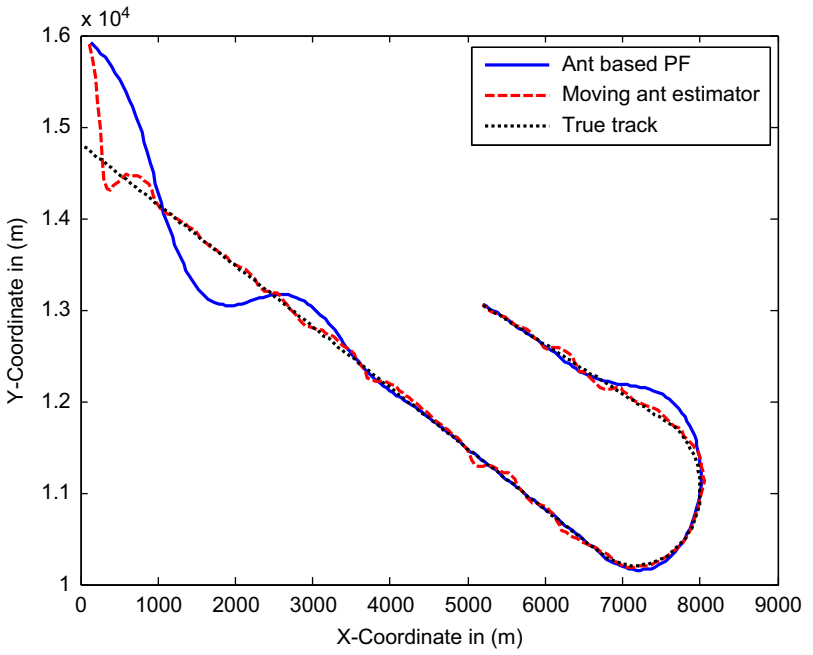

Fig. 11. The estimated tracks by moving ant estimator and ant based PF.

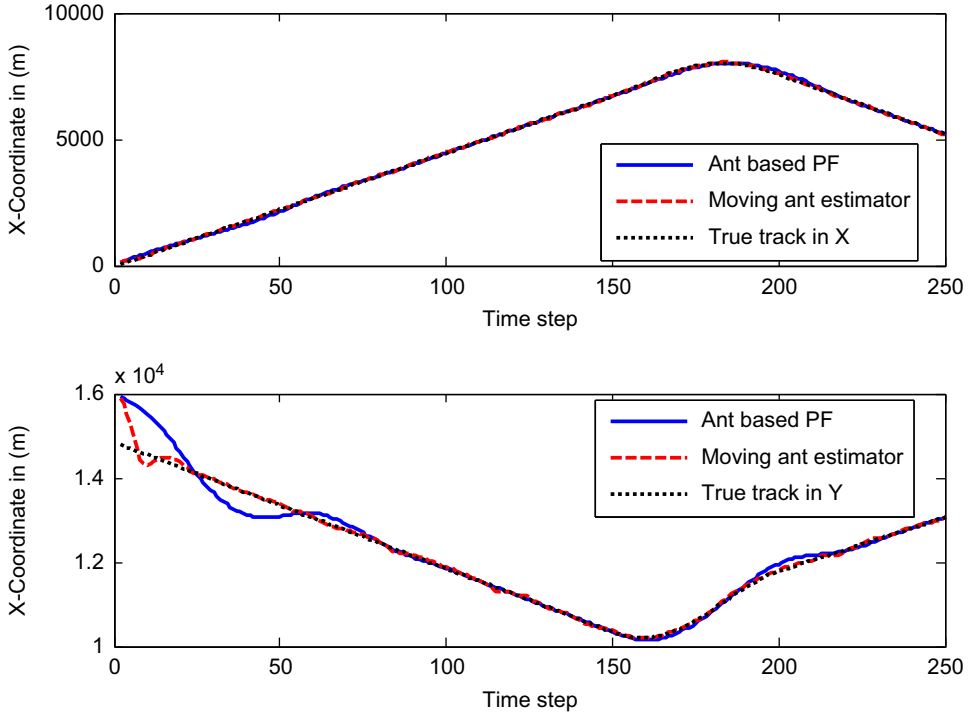

Fig. 12. The estimated positions in x and y directions by moving ant estimator and ant based PF.

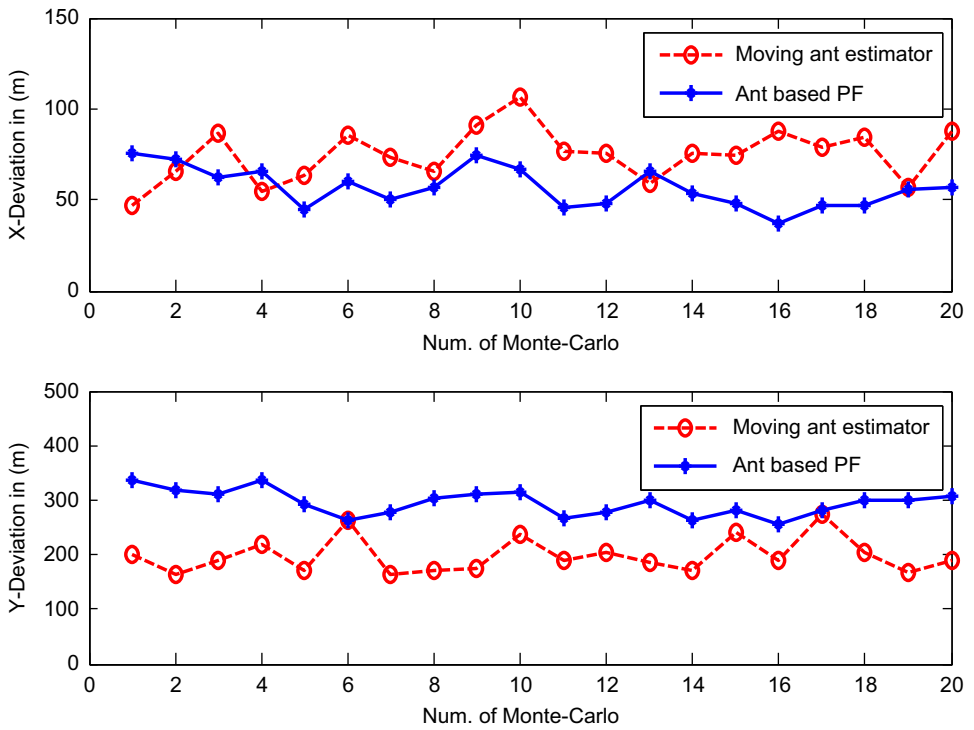

Fig. 13. The obtained deviation in x and y directions by moving ant estimator and ant based PF.

period, we adopt another state transition matrix represented by

$$\overline{F}=\begin{bmatrix} 1 & \frac{\sin(\omega_t T)}{\omega_t} & 0 & -\frac{1-\cos(\omega_t T)}{\omega_t} \\ 0 & \cos(\omega_t T) & 0 & -\sin(\omega_t T) \\ 0 & \frac{1-\cos(\omega_t T)}{\omega_t} & 1 & \frac{\sin(\omega_t T)}{\omega_t} \\ 0 & \sin(\omega_t T) & 0 & \cos(\omega_t T) \end{bmatrix} \tag{38}$$

with turn rate ω_t. Finally, after the maneuver the target maintains a constant velocity lasting about 50 samples. As mentioned above, all initial conditions are the same as the linear example except for adopting a smaller process noise covariance $\overline{\mathbf{Q}}(t)=diag(2.5^2,2.5^2,2.5^2,2.5^2)$. As shown in Figs. 6–9, the generic PF fails to deal with the model switching case, i.e., the generic PF loses track, illustrated in Figs. 7 and 8, or leads to tracking error divergence depicted in Fig. 6. Furthermore, the obtained

Table 2
The deviation sensitivity analysis in x direction with respect to parameters σ and q_0.

Average deviation in x direction		σ				
		10	15	20	25	30
q_0	0.70	68.6679	66.1778	73.5333	69.8424	64.4496
	0.75	63.7259	71.3825	74.2402	64.5417	64.5383
	0.80	63.4821	66.2929	***56.7081***	62.4278	59.8730
	0.85	***57.5554***	***57.1549***	65.5156	***56.6699***	58.8815
	0.90	59.3388	60.8777	65.9728	***56.3477***	65.2788

Table 3
The deviation sensitivity analysis in y direction with respect to parameters σ and q_0.

Average deviation in y direction		σ				
		10	15	20	25	30
q_0	0.70	301.0434	308.6211	315.9457	320.0349	306.0266
	0.75	304.0371	310.4306	308.7123	***295.8866***	309.9126
	0.80	302.7003	304.1744	294.9048	302.4232	307.8475
	0.85	303.0106	299.3038	308.1803	***290.2453***	***291.2302***
	0.90	296.5718	300.4598	***290.8575***	293.***0848***	304.3757

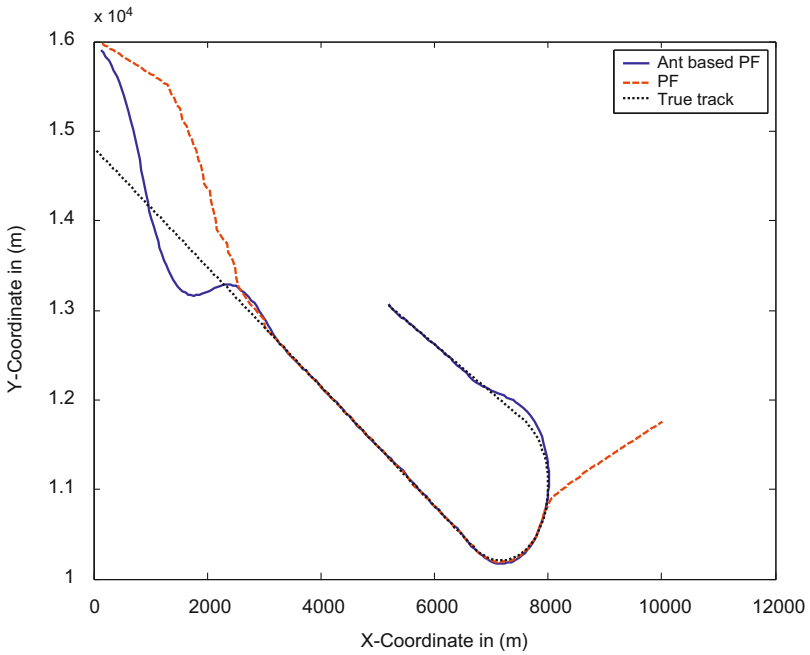

Fig. 14. The estimated tracks by PF and ant based PF.

deviation values of the generic PF is obviously larger than those in our proposed ant based PF, as indicated in Fig. 9.

Figs. 10–13 present the performance comparison of two ant-related estimators: moving ant estimator proposed in [15] and ant based PF proposed in this paper. In the initial phase of tracking, the moving ant estimator has a smaller RMSE than that of our proposed algorithm, but, as the tracking evolves, the two algorithms almost enjoy nice tracking performance, as illustrated in Figs. 10–12. As depicted in Fig. 13, compared to the moving ant estimator, the ant based PF has a smaller deviation in the x direction, while has a larger one in the y direction, thus we can say those two ant-related estimators have nearly identical performance in our discussed benchmark problem.

As mentioned in Section 2.2, there are two parameters introduced in our ant based PF, namely, σ and q_0. Parameter σ determines the deviation of particles generated according to $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$ from particle $\boldsymbol{f}(\boldsymbol{x}_{t-1}^{(j)})$, and a larger σ generally increases the distribution range of particles, but it in turn reduces the speed of convergence. While parameter q_0 determines the number of particles which are generated from the proposal density function $q_t^{(2)}(\cdot|\cdot,\boldsymbol{z}_t)$, and a smaller q_0 means a large number of particles that are involved in Ω_2. So we may decrease slightly the value of q_0 for model switching system to increase the possibility of catching true state. Tables 2 and 3 give the statistic results of average deviations in x and y directions for the above non-linear example, and the five smallest values are selected in a Italic and Bold style. It is observed that a smart choice of σ and q_0 lies in the approximate range of [20,25] and [0.85, 0.90], respectively.

Figs. 14 and 15 plot the tracking results of the non-linear example when N=400. It is seen that, as more particles are employed (e.g., from N=200 to 400), the resulting performances of both algorithms are improved slightly (compared with Figs. 7 and 8). However, the generic PF can not catch the motion of maneuvering target. Table 4 further illustrates the comparison of computation time required by the generic PF and our proposed algorithm. Although our proposed algorithm needs more time than the generic PF, it is suitable for real-time target tracking since the average execution time per step is less than the sample interval (T=1 s).

Table 4
Average required time (per time step) for different recursive techniques.

	Number of particles or ants		
	100	200	400
Our algorithm	0.0406s	0.1378s	0.4615s
Generic PF	0.0353s	0.0688s	0.1379s

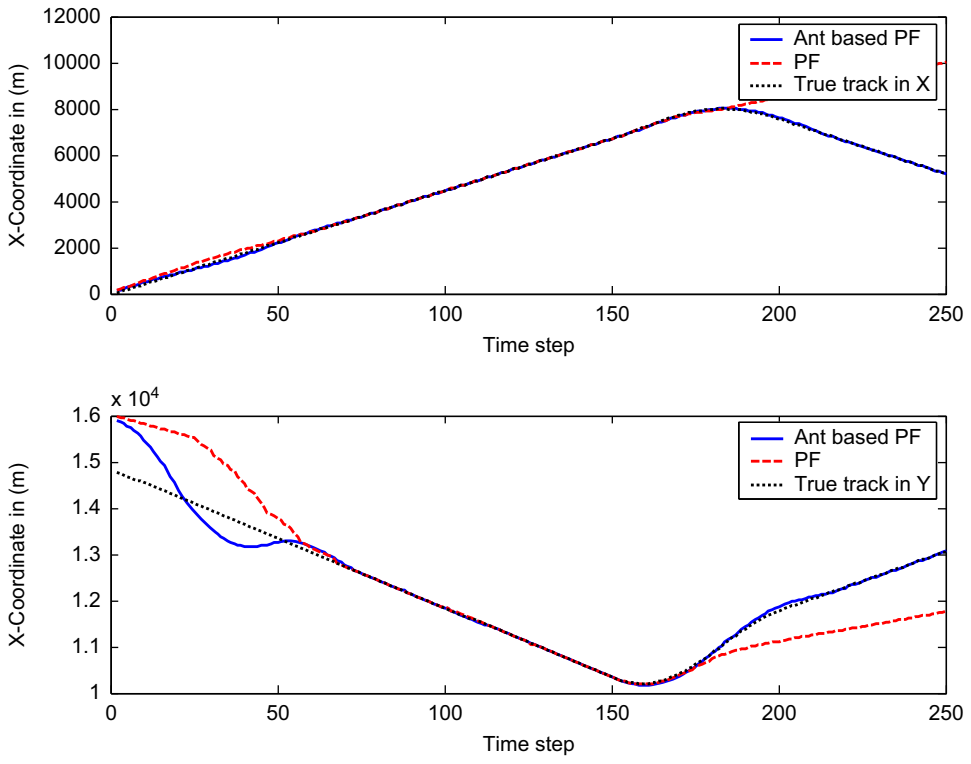

Fig. 15. The estimated positions in x and y directions.

5. Conclusions

In this work, we propose an intelligent and ant-based particle filter, in which two theoretic proposal density functions are developed to encapsulate model switching information by dividing probabilistically particles into respective model operations. On the basis of it, a nice but simple re-sampling scheme, which is developed from ant stochastic behavior as well, is proposed to achieve a better overlap with the true density function. Numerical simulations are carried out and the performance comparison is done among various estimators. It is found, firstly, that our proposed algorithm performs better than the generic PF both in a linear example and in a non-linear example, and the generic PF fails to estimate the state of non-linear system when the system begins to switch between models. Secondly, our proposed algorithm is competitive with another ant-related estimator (moving ant estimator) through a benchmark nonlinear example. Finally, the appropriate values taken by σ and q_0 are suggested to guarantee the tracking accuracy of our algorithm.

Acknowledgments

This work is supported by National Natural Science Foundation of China (No. 60804068) and by natural science fundamental research program of higher education colleges in Jiangsu province (No. 07KJB510001). Especially, I would thank professor Ba-Ngu Vo for his constructive suggestions when I was in the University of Melbourne.

Appendix A

Given $p_1(\boldsymbol{v})=\mathbb{N}(\boldsymbol{v};\boldsymbol{z}-\boldsymbol{Hx},\boldsymbol{R})$, $p_2(\boldsymbol{x})=\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})$, and $p_3(\boldsymbol{z})=\mathbb{N}(\boldsymbol{z};\boldsymbol{Hm},\boldsymbol{R}+\boldsymbol{HPH}^T)$, one computes the following ratio as

$$\frac{p_1(\boldsymbol{v})p_2(\boldsymbol{x})}{p_3(\boldsymbol{z})}=\frac{\mathbb{N}(\boldsymbol{v};\boldsymbol{z}-\boldsymbol{Hx},\boldsymbol{R})\cdot\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})}{\mathbb{N}(\boldsymbol{z};\boldsymbol{Hm},\boldsymbol{R}+\boldsymbol{HPH}^T)}=\frac{\mathbb{N}(\boldsymbol{z};\boldsymbol{Hx},\boldsymbol{R})\cdot\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})}{\mathbb{N}(\boldsymbol{z};\boldsymbol{Hm},\boldsymbol{R}+\boldsymbol{HPH}^T)}$$
$$=\frac{|2\pi\boldsymbol{R}|^{-1/2}|2\pi\boldsymbol{P}|^{-1/2}}{|2\pi(\boldsymbol{R}+\boldsymbol{HPH}^T)|^{-1/2}}e^{-1/2((\boldsymbol{z}-\boldsymbol{Hx})^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{Hx})+(\boldsymbol{x}-\boldsymbol{m})^T\boldsymbol{P}^{-1}(\boldsymbol{x}-\boldsymbol{m})-(\boldsymbol{z}-\boldsymbol{Hm})^T(\boldsymbol{R}+\boldsymbol{HPH}^T)^{-1}(\boldsymbol{z}-\boldsymbol{Hm}))} \tag{A.1}$$

And moreover, the exponential term in Eq. (A.1) can be written as a square form, so it simplifies to

$$(\boldsymbol{z}-\boldsymbol{Hx})^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{Hx})+(\boldsymbol{x}-\boldsymbol{m})^T\boldsymbol{P}^{-1}(\boldsymbol{x}-\boldsymbol{m})-(\boldsymbol{z}-\boldsymbol{Hm})^T(\boldsymbol{R}+\boldsymbol{HPH}^T)^{-1}(\boldsymbol{z}-\boldsymbol{Hm})$$
$$=(\boldsymbol{x}-\boldsymbol{m}-\boldsymbol{PH}^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{Hm}))^T(\boldsymbol{P}^{-1}+\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})(\boldsymbol{x}-\boldsymbol{m}-\boldsymbol{PH}^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{Hm}))\triangleq(\boldsymbol{x}-\overline{\boldsymbol{m}})^T\overline{\boldsymbol{P}}^{-1}(\boldsymbol{x}-\overline{\boldsymbol{m}}) \tag{A.2}$$

where $\overline{\boldsymbol{m}}=\boldsymbol{m}+\boldsymbol{PH}^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{Hm})$, $\overline{\boldsymbol{P}}^{-1}=\boldsymbol{P}^{-1}+\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H}$.

In addition, according to matrix inversion lemma

$$(\boldsymbol{A}+\boldsymbol{BCB}^T)^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}+\boldsymbol{C}^{-1})^{-1}\boldsymbol{B}^T\boldsymbol{A}^{-1} \tag{A.3}$$

We obtain

$$\overline{\boldsymbol{P}}=\boldsymbol{P}-\boldsymbol{PH}^T(\boldsymbol{HPH}^T+\boldsymbol{R})^{-1}\boldsymbol{HP} \tag{A.4}$$

So, Eq. (A.1) is written as a Gaussian function

$$\frac{p_1(\boldsymbol{v})p_2(\boldsymbol{x})}{p_3(\boldsymbol{z})}=\frac{\mathbb{N}(\boldsymbol{z};\boldsymbol{Hx},\boldsymbol{R})\cdot\mathbb{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{P})}{\mathbb{N}(\boldsymbol{z};\boldsymbol{Hm},\boldsymbol{R}+\boldsymbol{HPH}^T)}=\frac{|\overline{\boldsymbol{P}}|^{1/2}|\boldsymbol{R}+\boldsymbol{HPH}^T|^{1/2}}{|\boldsymbol{R}|^{1/2}|\boldsymbol{P}|^{1/2}}\mathbb{N}(\boldsymbol{x};\overline{\boldsymbol{m}},\overline{\boldsymbol{P}})\triangleq\boldsymbol{c}\mathbb{N}(\boldsymbol{x};\overline{\boldsymbol{m}},\overline{\boldsymbol{P}}) \tag{A.5}$$

with $c=|\overline{\boldsymbol{P}}|^{1/2}|\boldsymbol{R}+\boldsymbol{HPH}^T|^{1/2}/|\boldsymbol{R}|^{1/2}|\boldsymbol{P}|^{1/2}$.

Appendix B

Case 1:

According to Eq. (24), we have $\sum_{i=1}^{N}w_t^{(i)}=1$, and we denote

$$\hat{w}_t^{(j)}\triangleq\frac{w_t^{(j)}}{\sum_{k=1,k\neq i}^{N}w_t^{(k)}}=\frac{w_t^{(j)}}{1-w_t^{(i)}}$$

so the following inequality is required to hold

$$\tilde{w}_t^{(i)}=\frac{w_t^{(i)}}{w_t^{(i)}+w_t^{(j)}}\,\frac{w_t^{(j)}}{\sum_{k=1,k\neq i}^{N}w_t^{(k)}}\leq w_t^{(i)} \tag{B.1}$$

Rewriting (B.1) into a simplified form, we have

$$\frac{1}{w_t^{(i)}+w_t^{(j)}}\hat{w}_t^{(j)} \le 1 \tag{B.2}$$

(by simple mathematical operation)

$$\begin{aligned}\hat{w}_t^{(j)} &\le w_t^{(i)}+w_t^{(j)} \leftrightarrow \frac{w_t^{(j)}}{1-w_t^{(i)}} - w_t^{(j)} \le w_t^{(i)} \leftrightarrow w_t^{(j)} w_t^{(j)} \le w_t^{(i)} - (w_t^{(i)})^2 \leftrightarrow w_t^{(j)} w_t^{(j)} \le 1 \cdot w_t^{(i)} - (w_t^{(i)})^2 \leftrightarrow w_t^{(j)} w_t^{(j)} \\ &\le w_t^{(i)} \sum_{i=1}^{N} w_t^{(i)} - (w_t^{(i)})^2 \leftrightarrow w_t^{(j)} w_t^{(j)} \le w_t^{(i)} (w_t^{(1)} + \cdots + w_t^{(i)} + \cdots + w_t^{(j)} + \cdots w_t^{(N)}) - (w_t^{(i)})^2 \leftrightarrow 0 \\ &\le w_t^{(j)} (w_t^{(1)} + \cdots + w_t^{(i-1)} + w_t^{(i+1)} + \cdots + w_t^{(j-1)} + w_t^{(j+1)} + \cdots w_t^{(N)})\end{aligned} \tag{B.3}$$

Case 2: We now prove the inequality

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{w_t^{(i)}+w_t^{(j)}} \frac{w_t^{(i)}}{\sum_{k=1,k\neq j}^{N} w_t^{(k)}} \le w_t^{(i)}$$

holds.

Following the same step as in *Case* 1, the inequality proof can be simplified as

$$\frac{1}{w_t^{(i)}+w_t^{(j)}}\hat{w}_t^{(i)} \le 1 \tag{B.4}$$

where $\hat{w}_t^{(i)} = w_t^{(i)} / \sum_{k=1,k\neq j}^{N} w_t^{(k)}$.

$$\begin{aligned}\hat{w}_t^{(i)} &\le w_t^{(i)}+w_t^{(j)} \leftrightarrow \frac{w_t^{(i)}}{1-w_t^{(j)}} - w_t^{(i)} \le w_t^{(j)} \leftrightarrow w_t^{(i)} - w_t^{(i)} + w_t^{(i)} w_t^{(j)} \le w_t^{(j)} - (w_t^{(j)})^2 \leftrightarrow w_t^{(i)} w_t^{(j)} \le 1 \cdot w_t^{(j)} - (w_t^{(j)})^2 \leftrightarrow w_t^{(i)} w_t^{(j)} \\ &\le w_t^{(j)} \sum_{i=1}^{N} w_t^{(i)} - (w_t^{(j)})^2 \leftrightarrow w_t^{(i)} w_t^{(j)} \le w_t^{(j)} (w_t^{(1)} + \cdots + w_t^{(i)} + \cdots + w_t^{(j)} + \cdots w_t^{(N)}) - (w_t^{(j)})^2 \leftrightarrow 0 \\ &\le w_t^{(j)} (w_t^{(1)} + \cdots + w_t^{(i-1)} + w_t^{(i+1)} + \cdots + w_t^{(j-1)} + w_t^{(j+1)} + \cdots w_t^{(N)})\end{aligned} \tag{B.5}$$

Appendix C

Proof of Lemma 3. Suppose that one-step predicted distribution $\pi_{t|t-1}$ can be decomposed into $\pi'_{t|t-1}$ and $\pi''_{t|t-1}$, i.e. , $\pi_{t|t-1} = \pi'_{t|t-1} + \pi''_{t|t-1}$, so the prior condition $E[(\langle \pi^N_{t-1|t-1}, \varphi \rangle - \langle \pi_{t-1|t-1}, \varphi \rangle)^2] \le (a_{t-1|t-1}/N)\|\varphi\|^2$ can be described by $E[(\langle \pi^{N_{t1}}_{t-1|t-1}, \varphi \rangle - \langle \pi'_{t-1|t-1}, \varphi \rangle)^2] \le (\boldsymbol{a}'_{t-1|t-1}/N)\|\varphi\|^2$ and $E[(\langle \pi^{N_{t2}}_{t-1|t-1}, \varphi \rangle - \langle \pi''_{t-1|t-1}, \varphi \rangle)^2] \le (a''_{t-1|t-1}/N)\|\varphi\|^2$, respectively.

According to Eq. (23), we obtain

$$\begin{aligned}E[(\langle \overline{\pi}^N_{t|t-1}, \varphi \rangle - \langle \pi_{t|t-1}, \varphi \rangle)^2]^{1/2} &= E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1} + \overline{\pi}^{N_{t2}}_{t|t-1}, \varphi \rangle - \langle \pi'_{t|t-1} + \pi''_{t|t-1}, \varphi \rangle)^2]^{1/2} = E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - \langle \pi'_{t|t-1}, \varphi \rangle \\ &\quad + \langle \overline{\pi}^{N_{t2}}_{t|t-1}, \varphi \rangle - \langle \pi''_{t|t-1}, \varphi \rangle)^2]^{1/2}\end{aligned} \tag{C.1}$$

From the Minkowski's inequality, the above equation satisfies the following relation

$$E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - \langle \pi'_{t|t-1}, \varphi \rangle + \langle \overline{\pi}^{N_{t2}}_{t|t-1}, \varphi \rangle - \langle \pi''_{t|t-1}, \varphi \rangle)^2]^{1/2} \le E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - \langle \pi'_{t|t-1}, \varphi \rangle)^2]^{1/2} + E[(\langle \overline{\pi}^{N_{t2}}_{t|t-1}, \varphi \rangle - \langle \pi''_{t|t-1}, \varphi \rangle)^2]^{1/2} \tag{C.2}$$

For the first term at the right side of (C.2), it will be found to be a bounded function.

Using the mathematical inequality, we have

$$|\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - \langle \pi'_{t|t-1}, \varphi \rangle| \le |\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - \langle \pi^{N_{t1}}_{t-1|t-1}, p_{t|t-1}\varphi \rangle| + |\langle \pi^{N_{t1}}_{t-1|t-1}, p_{t|t-1}\varphi \rangle - \langle \pi'_{t-1|t-1}, p_{t|t-1}\varphi \rangle| \tag{C.3}$$

Let $\mathcal{G}_{t-1}$ be the σ-algebra generated by particles $\{x^{(i)}_{t-1}\}_{i=1}^{N_{t1}}$, then

$$\langle \pi^{N_{t1}}_{t-1|t-1}, p_{t|t-1}\varphi \rangle = E[\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle | \mathcal{G}_{t-1}] \tag{C.4}$$

According to the first term at the right side of (C.3), and independence of particles, we obtain

$$\begin{aligned}&E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - \langle \pi^{N_{t1}}_{t-1|t-1}, p_{t|t-1}\varphi \rangle)^2 | \mathcal{G}_{t-1}] \\ &\quad = E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle - E[\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle | \mathcal{G}_{t-1}])^2 | \mathcal{G}_{t-1}] \\ &\quad = E[(\langle \overline{\pi}^{N_{t1}}_{t|t-1}, \varphi \rangle)^2 | \mathcal{G}_{t-1}] - (\langle \pi^{N_{t1}}_{t-1|t-1}, p_{t|t-1}\varphi \rangle)^2\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N_{t1}} E[(w_{t|t-1}^{(i)}\varphi)^2|\mathcal{G}_{t-1}] - \sum_{i=1}^{N_{t1}} \left(\frac{1}{N} p_{t|t-1}\varphi\right)^2 \\
&= \sum_{i=1}^{N_{t1}} \left(E[(w_{t|t-1}^{(i)}\varphi)^2|\mathcal{G}_{t-1}] - \left(\frac{1}{N} p_{t|t-1}\varphi\right)^2\right) \\
&= \sum_{i=1}^{N_{t1}} \left(E\left[\left(\frac{p_{t|t-1}}{q_t^{(1)}}\frac{1}{N}\varphi\right)^2|\mathcal{G}_{t-1}\right] - \left(\frac{1}{N} p_{t|t-1}\varphi\right)^2\right) \\
&\le \frac{1}{N}\|\varphi\|^2\left(\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(1)}}\right\|^2\right)
\end{aligned} \tag{C.5}
$$

From (C.3) and Minkowski's inequality, it yields

$$
\begin{aligned}
&E[(\langle \overline{\pi}_{t|t-1}^{N_{t1}}, \varphi\rangle - \langle \pi'_{t|t-1}, \varphi\rangle)^2]^{1/2} \\
&\quad \le E[(\langle \overline{\pi}_{t|t-1}^{N_{t1}}, \varphi\rangle - \langle \pi_{t-1|t-1}^{N_{t1}}, p_{t|t-1}\varphi\rangle)^2]^{1/2} + E[(\langle \pi_{t-1|t-1}^{N_{t1}}, p_{t|t-1}\varphi\rangle - \langle \pi'_{t-1|t-1}, p_{t|t-1}\varphi\rangle)^2]^{1/2} \\
&\quad \le \frac{1}{\sqrt{N}}\|\varphi\|\left(\sqrt{\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(1)}}\right\|^2} + \sqrt{a'_{t-1|t-1}}\right)
\end{aligned} \tag{C.6}
$$

Similarly, for the second term at the right side of (C.2), we have

$$
E[(\langle \overline{\pi}_{t|t-1}^{N_{t2}}, \varphi\rangle - \langle \pi''_{t|t-1}, \varphi\rangle)^2]^{1/2} \le \frac{1}{\sqrt{N}}\|\varphi\|\left(\sqrt{\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(2)}}\right\|^2} + \sqrt{a''_{t-1|t-1}}\right) \tag{C.7}
$$

Thus we obtain

$$
\begin{aligned}
E[(\langle \overline{\pi}_{t|t-1}^{N}, \varphi\rangle - \langle \pi_{t|t-1}, \varphi\rangle)^2]^{1/2} &= E[(\langle \overline{\pi}_{t|t-1}^{N_{t1}}, \varphi\rangle - \langle \pi'_{t|t-1}, \varphi\rangle + \langle \overline{\pi}_{t|t-1}^{N_{t2}}, \varphi\rangle - \langle \pi''_{t|t-1}, \varphi\rangle)^2]^{1/2} \\
&\le \frac{1}{\sqrt{N}}\|\varphi\|\left(\sqrt{\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(1)}}\right\|^2} + \sqrt{a'_{t-1|t-1}} + \sqrt{\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(2)}}\right\|^2} + \sqrt{a''_{t-1|t-1}}\right)
\end{aligned} \tag{C.8}
$$

Consequently, it yields

$$
E[(\langle \overline{\pi}_{t|t-1}^{N}, \varphi\rangle - \langle \pi_{t|t-1}, \varphi\rangle)^2] \le \frac{a_{t|t-1}}{N}\|\varphi\|^2 \tag{C.9}
$$

where

$$
a_{t|t-1} \triangleq \left(\sqrt{\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(1)}}\right\|^2} + \sqrt{a'_{t-1|t-1}} + \sqrt{\|p_{t|t-1}\|^2 + \left\|\frac{p_{t|t-1}}{q_t^{(2)}}\right\|^2} + \sqrt{a''_{t-1|t-1}}\right)^2.
$$

Proof of Lemma 4. We consider

$$
\langle \overline{\pi}_{t|t}^{N}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle = \frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle} - \frac{\langle \pi_{t|t-1}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle} = \frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle} - \frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle} + \frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle} - \frac{\langle \pi_{t|t-1}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle} \tag{C.10}
$$

Note that in the above Eq. (C.10) the first two terms can be written as

$$
\left|\frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle} - \frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle}\right| = \frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle|\langle \pi_{t|t-1}, p_t\rangle - \langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle|}{\langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle\langle \pi_{t|t-1}, p_t\rangle} \le \frac{\|\varphi\|}{\langle \pi_{t|t-1}, p_t\rangle}|\langle \pi_{t|t-1}, p_t\rangle - \langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle| \tag{C.11}
$$

And the last two terms in Eq. (C.10) satisfy the inequality

$$
\left|\frac{\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle} - \frac{\langle \pi_{t|t-1}, p_t\varphi\rangle}{\langle \pi_{t|t-1}, p_t\rangle}\right| \le \frac{|\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle - \langle \pi_{t|t-1}, p_t\varphi\rangle|}{\langle \pi_{t|t-1}, p_t\rangle} \tag{C.12}
$$

Consequently, using Minkowski's inequality, we have

$$
\begin{aligned}
&E[(\langle \overline{\pi}_{t|t}^{N}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle)^2]^{1/2} \\
&\quad \le \frac{\|\varphi\|}{\langle \pi_{t|t-1}, p_t\rangle} E[|\langle \pi_{t|t-1}, p_t\rangle - \langle \overline{\pi}_{t|t-1}^{N}, p_t\rangle|^2]^{1/2} + \frac{1}{\langle \pi_{t|t-1}, p_t\rangle} E[|\langle \overline{\pi}_{t|t-1}^{N}, p_t\varphi\rangle - \langle \pi_{t|t-1}, p_t\varphi\rangle|^2]^{1/2} \\
&\quad \le \frac{\|\varphi\|}{\langle \pi_{t|t-1}, p_t\rangle}\left(\|p_t\|\sqrt{\frac{a_{t|t-1}}{N}}\right) + \frac{1}{\langle \pi_{t|t-1}, p_t\rangle}\left(\|p_t\varphi\|\sqrt{\frac{a_{t|t-1}}{N}}\right)
\end{aligned}
$$

$$\leq \frac{1}{\sqrt{N}}\|\varphi\|\left(\frac{2}{\langle \pi_{t|t-1}, p_t\rangle}\|p_t\|\sqrt{a_{t|t-1}}\right) \triangleq \frac{1}{\sqrt{N}}\|\varphi\|\sqrt{\overline{a}_{t|t}} \tag{C.13}$$

where

$$\overline{a}_{t|t} = \left(\frac{2}{\langle \pi_{t|t-1}, p_t\rangle}\|p_t\|\sqrt{a_{t|t-1}}\right)^2.$$

Proof of Lemma 5. We consider

$$\langle \pi^N_{t|t}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle = \langle \pi^N_{t|t}, \varphi\rangle - \langle \overline{\pi}^N_{t|t}, \varphi\rangle + \langle \overline{\pi}^N_{t|t}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle \tag{C.14}$$

Using Minkowski's inequality, it yields

$$E[(\langle \pi^N_{t|t}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle)^2]^{1/2} \leq E[(\langle \pi^N_{t|t}, \varphi\rangle - \langle \overline{\pi}^N_{t|t}, \varphi\rangle)^2]^{1/2} + E[(\langle \overline{\pi}^N_{t|t}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle)^2]^{1/2} \tag{C.15}$$

Let $\mathcal{J}_t$ be the be the σ-algebra generated by particles $\{\overline{x}^{(i)}_t\}_{i=1}^N$, according to the assumption of independence of particles and Lemma 6, we have

$$E[(\langle \pi^N_{t|t}, \varphi\rangle - \langle \overline{\pi}^N_{t|t}, \varphi\rangle)^2|\mathcal{J}_t] = E[(\langle \pi^N_{t|t}, \varphi\rangle - E[\langle \pi^N_{t|t}, \varphi\rangle|\mathcal{J}_t])^2|\mathcal{J}_t] \leq 2^2 E[|\langle \pi^N_{t|t}, \varphi\rangle|^2] \leq 2^2 \sum_{i=1}^N E\left[\left(\frac{1}{N}\varphi\right)^2\right] \leq \frac{2^2}{N}\|\varphi\|^2 \tag{C.16}$$

Using the conclusion in Lemma 4, (C.15) yields

$$E[(\langle \pi^N_{t|t}, \varphi\rangle - \langle \pi_{t|t}, \varphi\rangle)^2]^{1/2} \leq \frac{2}{\sqrt{N}}\|\varphi\| + \frac{1}{\sqrt{N}}\|\varphi\|\sqrt{\overline{a}_{t|t}} = \frac{1}{\sqrt{N}}\|\varphi\|(2+\sqrt{\overline{a}_{t|t}}) = \frac{1}{\sqrt{N}}\|\varphi\|\sqrt{a_{t|t}} \tag{C.17}$$

Lemma 6. *If $E|\eta|^p < \infty$, then, for any $p > 1$, we have $E|\eta - E\eta|^p \leq 2^p E|\eta|^p$.*

Proof. According to Jensen inequality, for $p > 1$, $(E|\eta|)^p \leq E|\eta|^p$ holds. Using Minkowski's inequality

$$(E|\eta - E\eta|^p)^{1/p} \leq (E|\eta^p|)^{1/p} + |E|\eta|| \leq 2(E|\eta|^p)^{1/p} \tag{C.18}$$

References

[1] Y. Yu, Q. Cheng, Particle filters for maneuvering target tracking problem, Signal Processing 86 (1) (2006) 195–203.
[2] M. Costagli, E. Engin Kuruoglu, Image separation using particle filters, Digital Signal Processing 17 (5) (2007) 935–946.
[3] F. Mustière, M. Bouchard, M. Bolić, Low-cost modifications of Rao–Blackwellized particle filters for improved speech denoising, Signal Processing 88 (11) (2008) 2678–2692.
[4] B.-N. Vo, S. Singh, A. Doucet, Sequential Monte Carlo methods for multi-target filtering with random finite sets, IEEE Transactions on Aerospace & Electronic Systems 41 (4) (2005) 1224–1245.
[5] H.-Y. Cheng, J.-N. Hwang, Adaptive particle sampling and adaptive appearance for multiple video object tracking, Signal Processing 89 (9) (2009) 1844–1849.
[6] R. van der Merwe, A. Doucet, N. de Freitas et al., The unscented particle filter, Technical Report CUED/F-INFENG/TR380 (2000) 1–46.
[7] K. Nishiyama, Fast and effective generation of the proposal distribution for particle filters, Signal Processing 85 (12) (2005) 2412–2417.
[8] M. Šimandl, O. Straka, Functional sampling density design for particle filters, Signal Processing 88 (11) (2008) 2784–2789.
[9] X.-L. Hu, T.B. Schön, L. Ljung, A basic convergence result for particle filtering, IEEE Transactions on Signal Processing 56 (4) (2008) 1337–1348.
[10] S. Park, J.P. Hwang, E. Kim, H.-J. Kang, A new evolutionary particle filter for the prevention of sample impoverishment, IEEE Transactions on Evolutionary Computation 13 (4) (2009) 801–809.
[11] M. Dorigo, V. Maniezzo, A. Colorni, The ant system: optimization by a colony of cooperating agents, IEEE Transactions on System, Man, and Cybernetics—part B 26 (1) (1996) 29–42.
[12] J.E. Bell, P.R. McMullen, Ant colony optimization techniques for the vehicle routing problem, Advanced Engineering Informatics 18 (1) (2004) 41–48.
[13] B. Xu, Q. Chen, Z. Wang, Track initiation with ant colony optimization, Communications in Nonlinear Science and Numerical Simulation 14 (9–10) (2009) 3629–3644.
[14] Y. Bar-Shalom, X.R. Li, T. Kirubarajan, Estimation with Applications to Tracking and Navigation, John Wiley & Sons, INC, New York, 2001.
[15] B. Xu, Q. Chen, X. Wang, J. Zhu, A novel estimator with moving ants, Simulation Modelling Practice and Theory 17 (10) (2009) 1663–1677.
[16] L. Nolle, On a novel ACO-estimator and its application to the target motion analysis problem, Knowledge-Based Systems 21 (3) (2008) 225–231.
[17] D. Crisan, A. Doucet, A survey of convergence results on particle filtering methods for practitioners, Transactions on Signal Processing 50 (3) (2002) 736–746.
[18] D.E. Clark, J. Bell, Convergence results for the particle PHD filter, IEEE Transactions on Signal Processing 54 (7) (2006) 2652–2661.