
Proceedings Paper:
Li, X., Liu, Y., Mihaylova, L. orcid.org/0000-0001-5856-2223 et al. (3 more authors) (2020) Enhanced multiple model GPB2 filtering using variational inference. In: 2019 22th International Conference on Information Fusion (FUSION). 22nd International Conference on Information Fusion, 02-05 Jul 2019, Ottawa, Canada. IEEE. ISBN: 9781728118406.

Enhanced Multiple Model GPB2 Filtering Using Variational Inference

Xi Li[1†], Yi Liu[2†], Lyudmila Mihaylova[3], Le Yang[4*], Steve Weddell[4], Fucheng Guo[1]
1. State Key Lab of Complex Electromagnetic Environment Effects on Electronics and Information Systems, National University of Defense Technology, Changsha China
2. School of Internet of Things (IoT) Engineering, Jiangnan University, Wuxi China
3. Department of Automatic Control and System Engineering, University of Sheffield, Sheffield UK
4. Department of Electrical and Computer Engineering, University of Canterbury, Christchurch NZ
†: Equal contributors, *: corresponding author: le.yang@canterbury.ac.nz.

***Abstract*—Multiple model filtering has been widely used to handle uncertainties in system dynamics and noise characteristics in state estimation problems. The generalized pseudo-Bayesian filter of order 2 (GPB2) is a suboptimal multiple model state estimator. It achieves computational tractability via approximating each model-matched state posterior, which is a Gaussian mixture, with a single Gaussian density. This paper illustrates from the viewpoint of variational inference that this approximation affects the performance of GPB2 through the model probability update stage. An enhanced GPB2 algorithm is proposed. It takes into account the above approximation by applying a correction factor that is dependent on the Kullback-Leibler divergence (KLD) of the Gaussian mixture and single Gaussian density. A control variate-based Monte Carlo method for evaluating the KLD is developed. The upper and lower bounds for the desired KLD are derived to correct the Monte Carlo KLD result if it falls out of bounds. Simulations show that the enhanced GPB2 algorithm outperforms the original GPB2 and interacting multiple model (IMM) methods in maneuvering target tracking tasks.**

I. Introduction

Estimating the state of a dynamic system from noisy measurements up to the current time, also referred to as *filtering* [1], has been extensively studied due to its diverse applications. Many filtering algorithms become available [1], [2]. Normally, they are designed on the basis of a state-space model with a process equation describing system state evolution and a measurement equation relating the measurement to the state. For example, the celebrated Kalman filter (KF) [1] was developed for linear state-space models with white Gaussian noise.

The performance of state estimation algorithms relies heavily on that the *assumed* state-space model matches the actual system. However, model uncertainties arise when the system has high complexity and is difficult to be modeled accurately [3]. Uncertainties can also come from the process and/or measurement noise statistics if they are not known [4]. Another common source for model uncertainties is the abrupt change in the system dynamics due to e.g., target maneuvering [5].

The presence of model uncertainties may degrade the state estimation performance greatly. To address this problem, robust techniques such as the risk sensitive filter and H_∞ filter have been proposed (see [6], [7] and references therein). They are designed to avoid large errors for a wide range of model parameters. On the other hand, multiple model approaches employ a set of candidate models to tackle model uncertainties [8]. The state estimate is obtained via performing model-matched filtering and combining the filtering results. We shall consider multiple model filtering in this work.

The optimal Bayesian solution to multiple model state estimation is computationally intractable due to the exponentially increasing number of hypotheses [1], [2]. Suboptimal techniques have thus been developed. The interacting multiple model (IMM) estimator and the generalized pseudo-Bayesian filters of orders 1 and 2 (GPB1 and GPB2) [1], [2] are among the most well-known methods. They constrain the number of hypotheses through approximating at each time instant every Gaussian mixture with a single Gaussian density via moment matching. IMM performs this approximation *before* the model-matched filtering, while GPB1 and GPB2 apply it *after* the model-matched filtering. Among the three methods, GPB2 has the highest complexity because it considers model transition between two consecutive time instants. But it performs as good as and sometimes better than IMM and GBP1 [9], [10].

In this paper, we focus on mitigating the effect of approximating the Gaussian mixture with the moment-matched single Gaussian density to improve the performance of GPB2. The error caused by this approximation in measurement update has been analyzed in [10], and the error covariance was approximately derived. A mixed IMM-GPB2 algorithm was built, where the more time-consuming GPB2 is performed only if the error covariance is above a threshold [10]. This method has an accuracy close to GPB2 but with reduced complexity.

Different from [10], this work examines the effect of the approximation error from the viewpoint of variational inference [11]. For this purpose, the development of the GPB2 algorithm is re-visited by following the derivation of the optimal Bayesian filter for the multiple model state estimation problem. We find that the approximation affects the filtering performance through the model probability update stage in GPB2. It is also shown that the approximation can be accounted for by introducing a correction factor, which depends on the Kullback-Leibler divergence (KLD) between the Gaussian mixture and the moment-matched single Gaussian density.

An enhanced GPB2 algorithm is then proposed that applies the aforementioned correction factor in its model identity

probability updating. We modify the control-variate based Monte Carlo method originally proposed in [12] for evaluating the KLD between two Gaussian mixtures to compute our analytically intractable KLD at low cost. To monitor the quality of the Monte Carlo result, the upper and lower bounds for the desired KLD are established by following [13], [14]. If the Monte Carlo result does not belong to the interval specified by the corresponding upper and lower bounds, we replace it with the mean of the two bounds. Simulation results show that the new GPB2 algorithm, which uses only 10^3 samples for the KLD evaluation, outperforms standard IMM and GPB2 in a linear and a nonlinear maneuvering target tracking problems.

The proposed enhanced GPB2 algorithm is different from the re-weighted IMM (RIMM) filter [15] that is based on expectation maximization (EM). Within RIMM, the corrections are applied in both the mixing step and model identity probability updating. In [16], the variational inference was applied to derive a new multiple model filtering method. It utilizes the mean field approximation and assumes that the joint posterior of the state and model identity can factorize. The enhanced GBP2, on the other hand, takes into account the effect of the moment matching-based approximation in its model probability updating only. It is still an approximation of the optimal Bayesian filter and does not require the posteriors of the state and model identity to be independent to each other as in [16].

The rest of this paper is organized as follows. Section II re-examines the derivation of the standard GPB2. Section III gives the enhanced GPB2 algorithm. Section IV shows the simulation results and Section V concludes the paper.

II. GPB2 Algorithm Re-visited

In this section, we re-visit the GPB2 algorithm by following the establishment of the optimal Bayesian filter for the multiple model state estimation problem [2]. The impact of approximating each Gaussian mixture with a single Gaussian density is analyzed from the viewpoint of variational inference.

The multiple model filtering problem in consideration is formulated as follows. There are M models and the model switching is *independent* of the system state transition. Let the model identity at time k be r_k. The evolution of r_k (i.e., the model switching) is a homogeneous Markovian jump process with a transition probability matrix $\mathbf{P}$ defined as

$$\mathbf{P}(i,j) = p(r_k = j | r_{k-1} = i) = p_{ij}, \tag{1}$$

where $i, j = 1, 2, ..., M$ and $\sum_{j=1}^{M} p_{ij} = 1$.

The state-space model under model r_k is described by the following model-conditioned state transition probability density function (PDF) and measurement likelihood function

$$\mathbf{x}_k | \mathbf{x}_{k-1}, r_k \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k), \tag{2a}$$

$$\mathbf{z}_k | \mathbf{x}_k, r_k \sim p(\mathbf{z}_k | \mathbf{x}_k, r_k). \tag{2b}$$

Here, $\mathbf{x}_k$ and $\mathbf{z}_k$ are the system state vector and measurement vector at time k, respectively. It is assumed that the process noise and measurement noise are independent and both white.

We are interested in finding the posterior of the state vector $\mathbf{x}_k$, denoted by $p(\mathbf{x}_k | \mathbf{Z}_{1:k})$, where $\mathbf{Z}_{1:k} = \{\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_k\}$ collects the measurements obtained up to time k.

A. Optimal Multiple Model Bayesian Filter

The optimal Bayesian filter evaluates $p(\mathbf{x}_k | \mathbf{Z}_{1:k})$ recursively. For this purpose, the state posterior is expressed as

$$p(\mathbf{x}_k | \mathbf{Z}_{1:k}) = \sum_{j=1}^{M} p(\mathbf{x}_k | r_k = j, \mathbf{Z}_{1:k}) \underbrace{p(r_k = j | \mathbf{Z}_{1:k})}_{u^j_{k|k}}, \tag{3}$$

where $p(\mathbf{x}_k | r_k = j, \mathbf{Z}_{1:k})$ is the model-matched posterior of $\mathbf{x}_k$ and $u^j_{k|k}$ is the model identity probability at time k.

Invoking Bayes' theorem yields

$$p(\mathbf{x}_k | r_k = j, \mathbf{Z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k, r_k = j) p(\mathbf{x}_k | r_k = j, \mathbf{Z}_{1:k-1})}{p(\mathbf{z}_k | r_k = j, \mathbf{Z}_{1:k-1})}, \tag{4}$$

$$u^j_{k|k} = \frac{p(\mathbf{z}_k | r_k = j, \mathbf{Z}_{1:k-1}) u^j_{k|k-1}}{p(\mathbf{z}_k | \mathbf{Z}_{1:k-1})}. \tag{5}$$

Here, $p(\mathbf{z}_k | \mathbf{x}_k, r_k = j)$ is the measurement likelihood function under model $r_k = j$ (see (2b)) and $u^j_{k|k-1}$ is the predictive model identity probability, which is equal to

$$u^j_{k|k-1} = p(r_k = j | \mathbf{Z}_{1:k-1}) = \sum_{i=1}^{M} p_{ij} u^i_{k-1|k-1}. \tag{6}$$

$u^i_{k-1|k-1}$, $i = 1, 2, ..., M$, are the model identity probabilities at the previous time instant $k-1$.

We evaluate (4) and (5) to establish the optimal multiple model Bayesian filter. Specifically, computing (4) requires the model-conditioned predictive PDF $p(\mathbf{x}_k | r_k = j, \mathbf{Z}_{1:k-1})$, which is

$$\begin{aligned} &p(\mathbf{x}_k | r_k = j, \mathbf{Z}_{1:k-1}) \\ &\quad = \sum_{i=1}^{M} p(\mathbf{x}_k | r_k = j, r_{k-1} = i, \mathbf{Z}_{1:k-1}) \lambda^{ij}_k. \end{aligned} \tag{7}$$

In (7), λ^{ij}_k is the *mixing probability* defined as

$$\lambda^{ij}_k = p(r_{k-1} = i | r_k = j, \mathbf{Z}_{1:k-1}) = \frac{p_{ij} u^i_{k-1|k-1}}{\sum_{i=1}^{M} p_{ij} u^i_{k-1|k-1}} \tag{8}$$

and $p(\mathbf{x}_k | r_k = j, r_{k-1} = i, \mathbf{Z}_{1:k-1})$ is given by the Chapman-Kolmogorov equation [1]

$$\begin{aligned} &p(\mathbf{x}_k | r_k = j, r_{k-1} = i, \mathbf{Z}_{1:k-1}) \\ &\quad = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k = j) p(\mathbf{x}_{k-1} | r_{k-1} = i, \mathbf{Z}_{1:k-1}) \mathrm{d}\mathbf{x}_{k-1}. \end{aligned} \tag{9}$$

Here, $p(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k = j)$ is the state transition PDF under model $r_k = j$ (see (2a)) and $p(\mathbf{x}_{k-1} | r_{k-1} = i, \mathbf{Z}_{1:k-1})$ is the model-matched posterior of the state vector at time $k-1$.

To evaluate (5), we need to compute the model-conditioned measurement likelihood $p(\mathbf{z}_k|r_k = j, \mathbf{Z}_{1:k-1})$, which is also the normalization factor in (4). It can be found via

$$p(\mathbf{z}_k|r_k = j, \mathbf{Z}_{1:k-1}) = \int p(\mathbf{z}_k|\mathbf{x}_k, r_k = j)p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k-1})\mathrm{d}\mathbf{x}_k. \tag{10}$$

This completes the derivation of the optimal Bayesian filter for the considered multiple model state estimation problem. To summarize, at time k, we start with $p(\mathbf{x}_{k-1}|r_{k-1} = i, \mathbf{Z}_{1:k-1})$ and $u^i_{k-1|k-1}$, and performs two processing steps:

1) **Model-matched filtering**: For $r_k = j$, $j = 1, 2, ..., M$, first evaluate (9) for $i = 1, 2, ..., M$. Then, combine the results using (7) and mixing probabilities λ^{ij}_k from (8) to obtain the model-conditioned predictive PDF. Finally, perform the model-matching filtering using (4) to produce the model-matched state posterior at the current time $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$.

2) **Model identity probability updating**: For $r_k = j$, $j = 1, 2, ..., M$, first find the predictive model identity probability $u^j_{k|k-1}$ from (6). Next, compute the model-conditioned measurement likelihood using (10) and put the result into (5) to generate $u^j_{k|k}$, the model identity probability at time k.

B. Standard GPB2

In the optimal Bayesian filter, the model-matched state posterior $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$ is always a mixture, due to the interaction among the models in (7). This would eventually lead to an exponentially increasing number of components in $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$. To address this drawback, the standard GPB2 assumes a *linear* multiple model state-space model [1], [2]. Suppose at time $k-1$, the model-matched state posterior $p(\mathbf{x}_{k-1}|r_{k-1} = i, \mathbf{Z}_{1:k-1})$ is a single Gaussian PDF. As such, evaluating (9) would produce a Gaussian density as well and the model-conditioned predictive PDF in (7) now becomes a Gaussian mixture with the functional form

$$p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k-1}) = \sum_{i=1}^{M} \lambda^{ij}_k \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{ij}_{k|k-1}, \boldsymbol{\Sigma}^{ij}_{k|k-1}), \tag{11}$$

where λ^{ij}_k are the mixing probabilities defined in (8). $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents a multivariate Gaussian density with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

At time k, the GPB2 algorithm first follows *exactly* the processing of the optimal Bayesian filter. Specifically, it substitutes $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k-1})$ in (11) into (4) and carries out the model-matched filtering, which involves applying the updating stage of a KF M times for a given j. The obtained model-matched state posterior would also be a Gaussian mixture, which can be expressed as

$$p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}) = \sum_{i=1}^{M} w^{ij}_k \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{ij}_{k|k}, \boldsymbol{\Sigma}^{ij}_{k|k}). \tag{12}$$

The weights w^{ij}_k are equal to $w^{ij}_k = \frac{1}{c^j_k}\lambda^{ij}_k \Lambda^{ij}_k$, where they satisfy $\sum_{i=1}^{M} w^{ij}_k = 1$ and

$$\Lambda^{ij}_k = \int p(\mathbf{z}_k|\mathbf{x}_k, r_k = j)\mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{ij}_{k|k-1}, \boldsymbol{\Sigma}^{ij}_{k|k-1})\mathrm{d}\mathbf{x}_k. \tag{13}$$

c^j_k is the normalization factor equal to

$$c^j_k = p(\mathbf{z}_k|r_k = j, \mathbf{Z}_{1:k-1}) = \sum_{i=1}^{M} \lambda^{ij}_k \Lambda^{ij}_k. \tag{14}$$

GPB2 finds the model identity probability at time k using

$$u^j_{k|k} \propto c^j_k \cdot u^j_{k|k-1}. \tag{15}$$

To prevent the number of components in $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$ from growing exponentially, GPB2 introduces an additional operation. Specifically, it approximates $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$ in (12) using a single Gaussian density such that

$$p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}) \approx q_j(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^j_{k|k}, \boldsymbol{\Sigma}^j_{k|k}). \tag{16}$$

The *approximated* posterior mean and covariance $\boldsymbol{\mu}^j_{k|k}$ and $\boldsymbol{\Sigma}^j_{k|k}$ are obtained via moment matching. They are equal to

$$\boldsymbol{\mu}^j_{k|k} = \sum_{i=1}^{M} w^{ij}_k \boldsymbol{\mu}^{ij}_{k|k}, \tag{17a}$$

$$\boldsymbol{\Sigma}^j_{k|k} = \sum_{i=1}^{M} w^{ij}_k \left(\boldsymbol{\Sigma}^{ij}_{k|k} + (\boldsymbol{\mu}^{ij}_{k|k} - \boldsymbol{\mu}^j_{k|k})(\boldsymbol{\mu}^{ij}_{k|k} - \boldsymbol{\mu}^j_{k|k})^T \right). \tag{17b}$$

It is worthwhile to point out that the standard GPB2 derived above can be generalized to *nonlinear* multiple model filtering scenarios (see e.g., [17]). They apply nonlinear Gaussian filters such as the deterministic/random-point-based filters [18]–[22] to evaluate the integrals in (9) and (13), and perform the model-matched filtering in (4).

C. Analysis from the Viewpoint of Variational Inference

As shown in the previous subsection, the standard GPB2 propagates the *approximated* model-matched state posteriors $q_j(\mathbf{x}_k)$ to the next time instant to achieve computational tractability. The model identity probabilities $u^j_{k|k}$ are also passed on but they are computed under the condition that the corresponding model-matched posteriors are Gaussian mixtures. To further illustrate this, we note from (15) that updating the model identity probabilities requires the model-conditioned measurement likelihood c^j_k but it is obtained in (14) *before* the approximation in (16). In other words, in the standard GPB2, the effect of replacing the model-matched state posteriors with their single Gaussian density approximations on the model identity probability updating is *not* taken into account.

From the perspective of variational inference, for the model-matched filtering problem (4), the logarithm of the associated measurement likelihood when using $q_j(\mathbf{x}_k)$ to approximate the true posterior $p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$ is [11]

$$L(q_j(\mathbf{x}_k)) = \int q_j(\mathbf{x}_k) \log \frac{p(\mathbf{z}_k, \mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k-1})}{q_j(\mathbf{x}_k)} \mathrm{d}\mathbf{x}_k. \tag{18}$$

As expected, if $q_j(\mathbf{x}_k) = p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})$, the model-conditioned measurement likelihood computed from (18), denoted for notation simplicity as

$$\tilde{c}^j_k = \exp(L(q_j(\mathbf{x}_k))) \tag{19}$$

would be equal to $p(\mathbf{z}_k|r_k=j,\mathbf{Z}_{1:k-1})=c_k^j$ in (14), because

$$\begin{aligned}&p(\mathbf{z}_k,\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k-1})\\&\quad=p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})p(\mathbf{z}_k|r_k=j,\mathbf{Z}_{1:k-1}).\end{aligned}$$

In general, $\tilde{c}_k^j$ is smaller than c_k^j, which can be verified via

$$\begin{aligned}L(q_j(\mathbf{x}_k))&=\int q_j(\mathbf{x}_k)\log\frac{p(\mathbf{z}_k,\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k-1})}{p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})}\mathrm{d}\mathbf{x}_k\\&\quad+\int q_j(\mathbf{x}_k)\log\frac{p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})}{q_j(\mathbf{x}_k)}\mathrm{d}\mathbf{x}_k\\&=\log(c_k^j)-D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})),\end{aligned}\tag{20}$$

where

$$\begin{aligned}&D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k}))\\&\quad=\int q_j(\mathbf{x}_k)\log\frac{q_j(\mathbf{x}_k)}{p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})}\mathrm{d}\mathbf{x}_k\end{aligned}\tag{21}$$

is the KLD between $q_j(\mathbf{x}_k)$ and $p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})$. As the KLD is non-negative [23], we have

$$L(q_j(\mathbf{x}_k))\le\log(c_k^j)\iff\tilde{c}_k^j\le c_k^j.\tag{22}$$

In literature, $L(q_j(\mathbf{x}_k))$ is sometimes referred to as the evidence lower bound (ELBO) [11]. Its applications include model selection [11] and particle filter (PF)-based online multi-output Gaussian process regression and learning [24], to name a few.

The above analysis reveals that propagating the *approximated* model-matched state posterior $q_j(\mathbf{x}_k)$, a single Gaussian density, instead of the true version $p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})$, a Gaussian mixture, leads to decreased model-conditioned measurement likelihood. More importantly, from (20), we have

$$\tilde{c}_k^j=c_k^j\cdot\underbrace{\exp(-D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k})))}_{\Gamma_k^j},\tag{23}$$

where Γ_k^j can be considered as the correction factor due to the approximation in (16).

III. Proposed Algorithm

A. Enhanced GPB2

We shall present an enhanced GPB2 algorithm for multiple model state estimation based on the insights obtained in Section II.C. The proposed method differs from the standard GPB2 only in the model identity probability updating. More specifically, the enhanced GPB2 algorithm finds the model identity probability using, after substituting (23),

$$u_{k|k}^j\propto\tilde{c}_k^j u_{k|k-1}^j=c_k^j\Gamma_k^j u_{k|k-1}^j.\tag{24}$$

In words, the new method considers the impact of approximating the true model-matched state posterior with a single Gaussian density by utilizing the ELBO of the corresponding measurement likelihood in the model identity probability updating.

Because c_k^j and $u_{k|k-1}^j$ are *already* calculated in the standard GPB2, the rest of this section will focus on how to find the correction factor Γ_k^j to complete the establishment of the enhanced GPB2 algorithm.

B. Control Variate-based Monte Carlo for Computing Γ_k^j

We note from (23) that in order to evaluate Γ_k^j, we just need to compute the KLD $D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k}))$. For this purpose, putting (12) and (16) into (21) yields

$$D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k=j,\mathbf{Z}_{1:k}))=-H(q_j(\mathbf{x}_k))-\psi_k^j.\tag{25}$$

Here, $H(q_j(\mathbf{x}_k))$ is the entropy of the multivariate Gaussian distribution $q_j(\mathbf{x}_k)$, which is equal to [14]

$$H(q_j(\mathbf{x}_k))=-\int q_j(\mathbf{x}_k)\log q_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k=\frac{1}{2}\log(|2\pi e\boldsymbol{\Sigma}_{k|k}^j|),\tag{26}$$

and ψ_k^j is defined as

$$\psi_k^j=\int q_j(\mathbf{x}_k)\log\sum_{i=1}^M w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})\mathrm{d}\mathbf{x}_k.\tag{27}$$

Unfortunately, ψ_k^j does not have a closed-form expression. Because it is easy to draw samples from $q_j(\mathbf{x}_k)$, the *naive* Monte Carlo method might be used for computing the integral in (27). However, as shown in [13], [14], even with 10^5 samples, the obtained result could still have large errors frequently, and the use of 10^6 samples is recommended, which is computationally quite expensive.

We shall follow [12] and develop a low-complexity control variate-based Monte Carlo method for evaluating (27). The idea is to first find a control variate $f_j(\mathbf{x}_k)$, which is an integrable function, such that

$$\psi_k^j=\int f_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k+\tilde{\psi}_k^j,\tag{28}$$

where $\tilde{\psi}_k^j$ is defined as

$$\tilde{\psi}_k^j=\int\left(q_j(\mathbf{x}_k)\log\sum_{i=1}^M w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})-f_j(\mathbf{x}_k)\right)\mathrm{d}\mathbf{x}_k.\tag{29}$$

With the control variate $f_j(\mathbf{x}_k)$, the problem of evaluating the integral in (27) reduces to calculating $\tilde{\psi}_k^j$. If $f_j(\mathbf{x}_k)$ is a good approximation to the original integrand $q_j(\mathbf{x}_k)\log\sum_{i=1}^M w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})$, $\tilde{\psi}_k^j$ can be computed using the naive Monte Carlo method with low variance. This makes it possible for the control variate-based approach to achieve relatively high accuracy in estimating ψ_k^j with a reasonably small number of samples.

We shall adopt the following control variate to find ψ_k^j

$$f_j(\mathbf{x}_k)=q_j(\mathbf{x}_k)\sum_{i=1}^M a_k^{ij}\log\frac{w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})}{a_k^{ij}},\tag{30}$$

where a_k^{ij} are non-negative parameters to be determined and they satisfy $\sum_{i=1}^M a_k^{ij}=1$. It can be shown by invoking the Jensen's inequality that

$$q_j(\mathbf{x}_k)\log\sum_{i=1}^M w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})\ge f_j(\mathbf{x}_k).\tag{31}$$

As a result, we have from (29) that with the control variate $f_j(\mathbf{x}_k)$ in (30), $\tilde{\psi}_k^j$ would be non-negative.

We shall find the parameters a_k^{ij} in $f_j(\mathbf{x}_k)$ by maximizing $\int f_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k$ with respect to a_k^{ij}. This could produce a control variate that is close to the integrand in (27), as desired. Mathematically, the associated optimization problem is

$$\max_{a_k^{ij}} \int f_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k, \quad \text{subject to } \sum_{i=1}^{M} a_k^{ij} = 1. \tag{32}$$

The method of Lagrange multipliers is applied to solve (32). In particular, the Lagrangian is

$$\mathcal{L}_j = \int f_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k + \gamma_j(\sum_{i=1}^{M} a_k^{ij} - 1), \tag{33}$$

where γ_j is the Lagrange multiplier. Setting the partial derivative of $\mathcal{L}_j$ with respect to a_k^{ij} to zero yields

$$a_k^{ij} = e^{\gamma_j - 1}\exp\left(\int q_j(\mathbf{x}_k)\log w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})\mathrm{d}\mathbf{x}_k\right). \tag{34}$$

Putting the above result into the equality constraint in (32) leads to

$$e^{\gamma_j - 1} = \frac{1}{\sum_{i=1}^{M} w_k^{ij} e^{\int q_j(\mathbf{x}_k)\log\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})\mathrm{d}\mathbf{x}_k}}. \tag{35}$$

As a result, the solution to (32) is

$$a_k^{ij} = \frac{w_k^{ij} e^{\int q_j(\mathbf{x}_k)\log\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})\mathrm{d}\mathbf{x}_k}}{\sum_{i'=1}^{M} w_k^{i'j} e^{\int q_j(\mathbf{x}_k)\log\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{i'j},\boldsymbol{\Sigma}_{k|k}^{i'j})\mathrm{d}\mathbf{x}_k}}. \tag{36}$$

After multiplying both the numerator and denominator in (36) with $e^{-\int q_j(\mathbf{x}_k)\log q_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k}$, we have

$$a_k^{ij} = \frac{w_k^{ij} e^{-D(q_j(\mathbf{x}_k)||\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij}))}}{\sum_{i'=1}^{M} w_k^{i'j} e^{-D(q_j(\mathbf{x}_k)||\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{i'j},\boldsymbol{\Sigma}_{k|k}^{i'j}))}}, \tag{37}$$

where $D(q_j(\mathbf{x}_k)||\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij}))$ is the KLD between two multivariate Gaussian PDFs, $q_j(\mathbf{x}_k)$ and $\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})$. It has the following closed-form expression [25]

$$\begin{aligned} D(q_j(\mathbf{x}_k)||\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})) &= \frac{1}{2}\log|\boldsymbol{\Sigma}_{k|k}^{ij}| - \frac{1}{2}\log|e\boldsymbol{\Sigma}_{k|k}^{j}| \\ &+ \frac{1}{2}\mathrm{Tr}\left((\boldsymbol{\Sigma}_{k|k}^{ij})^{-1}(\boldsymbol{\Sigma}_{k|k}^{j} + (\boldsymbol{\mu}_{k|k}^{j} - \boldsymbol{\mu}_{k|k}^{ij})(\boldsymbol{\mu}_{k|k}^{j} - \boldsymbol{\mu}_{k|k}^{ij})^T)\right), \end{aligned} \tag{38}$$

where $\mathrm{Tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$.

Substituting (37) back into (30) yields the optimized control variate function $f_j(\mathbf{x}_k)$. Integrating it gives

$$\int f_j(\mathbf{x}_k)\mathrm{d}\mathbf{x}_k = -H(q_j(\mathbf{x}_k)) + \beta_k^j, \tag{39}$$

where

$$\beta_k^j = \log\sum_{i=1}^{M} w_k^{ij} e^{-D(q_j(\mathbf{x}_k)||\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij}))}, \tag{40}$$

and $H(q_j(\mathbf{x}_k))$ is the entropy of $q_j(\mathbf{x}_k)$ (see (26)). The result in (39) is identical to the variational approximation result in [13], although no details on the theoretical developments are provided there.

Putting (39) into (28) and substituting the result back to (25), we obtain that the KLD between $q_j(\mathbf{x}_k)$ and the true model-matched state posterior can be found using

$$D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})) = -\beta_k^j - \tilde{\psi}_k^j. \tag{41}$$

The term $\tilde{\psi}_k^j$ now becomes, after utilizing (29) and (30),

$$\begin{aligned} \tilde{\psi}_k^j = \int q_j(\mathbf{x}_k)\Bigg(&\log\sum_{i=1}^{M} w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij}) \\ &- \sum_{i=1}^{M} a_k^{ij}\log\frac{w_k^{ij}\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})}{a_k^{ij}}\Bigg)\mathrm{d}\mathbf{x}_k. \end{aligned} \tag{42}$$

We evaluate the above integral using the naive Monte Carlo method with samples from the approximated model-matched state posterior $q_j(\mathbf{x}_k)$. Let us denote the result by $\hat{\tilde{\psi}}_k^j$. Replacing $\tilde{\psi}_k^j$ with $\hat{\tilde{\psi}}_k^j$ and substituting (41) into the definition of the correction factor in (23) yields an estimate of Γ_k^j given by

$$\hat{\Gamma}_k^j = \exp(\beta_k^j + \hat{\tilde{\psi}}_k^j). \tag{43}$$

This completes the development of the control variate-based method for evaluating the correction factor Γ_k^j in the enhanced GPB2 algorithm.

C. Upper & Lower Bounds for $D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}))$

In this subsection, we shall establish the upper and lower bounds for $D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}))$ to monitor the estimation accuracy of the Monte Carlo KLD result.

An upper bound for $D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}))$ is $-\beta_k^j$. This can be verified by noting from the discussion under (31) that with the used control variate in (30), $\tilde{\psi}_k^j$ would be non-negative. As a result, we have, from (41),

$$D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})) \leq -\beta_k^j. \tag{44}$$

The lower bound for $D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}))$ can be found by applying the Jensen's inequality in (27) such that

$$\begin{aligned} \psi_j^k &\leq \log\sum_{i=1}^{M} w_k^{ij}\int q_j(\mathbf{x}_k)\mathcal{N}(\mathbf{x}_k;\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{ij})\mathrm{d}\mathbf{x}_k \\ &= \log\sum_{i=1}^{M} w_k^{ij}\mathcal{N}(\boldsymbol{\mu}_{k|k}^{j};\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{j} + \boldsymbol{\Sigma}_{k|k}^{ij}), \end{aligned} \tag{45}$$

where the second equality is established using the update stage of a KF. Putting the above result into (25) and considering that $D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k}))$ is non-negative [23] yield

$$\alpha_k^j \leq D(q_j(\mathbf{x}_k)||p(\mathbf{x}_k|r_k = j, \mathbf{Z}_{1:k})), \tag{46}$$

where the lower bound α_j is equal to

$$\max\{0, -H(q_j(\mathbf{x}_k)) - \log\sum_{i=1}^{M} w_k^{ij}\mathcal{N}(\boldsymbol{\mu}_{k|k}^{j};\boldsymbol{\mu}_{k|k}^{ij},\boldsymbol{\Sigma}_{k|k}^{j} + \boldsymbol{\Sigma}_{k|k}^{ij})\}. \tag{47}$$

If the Monte Carlo KLD result $-(\beta_k^j + \hat{\tilde{\psi}}_k^j)$ falls out of the interval $[\alpha_k^j, -\beta_k^j]$, we replace it with the mean of the two bounds, which is $(\alpha_k^j - \beta_k^j)/2$. In this case, the correction factor in the enhanced GPB2 algorithm is computed via

$$\hat{\Gamma}_k^j = \exp\Big((\alpha_k^j - \beta_k^j)/2\Big). \tag{48}$$

IV. Simulations

In this section, we examine the performance of the newly proposed enhanced GPB2 algorithm in maneuvering target tracking. The benchmark methods are the standard IMM [1] and GPB2 [2]. In future works, we shall study its performance in angle and range-based [26] as well as time difference of arrival (TDOA) and frequency difference of arrival (FDOA)-based [27] target tracking problems.

A. Tracking Scenario

Consider tracking a point target on the 2-D plane. The target motion starts at $[234.9\text{km}, 85.5\text{km}]^T$ with a velocity $[-176.8\text{m/s}, -176.8\text{m/s}]^T$. Its motion state at time k is denoted by $\mathbf{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k]^T$, where $[x_k, y_k]^T$ contains the target position coordinates along the x-axis and y-axis, and $[\dot{x}_k, \dot{y}_k]^T$ is the velocity vector. $\mathbf{x}_k$ evolves according to the constant turn (CT) model [5]

$$\mathbf{x}_k = \mathbf{F}_k(\omega)\mathbf{x}_{k-1} + \mathbf{G}_k\mathbf{w}_k, \tag{49}$$

where ω is the turn rate, the state transition matrix $\mathbf{F}_k(w)$ is

$$\mathbf{F}_k(\omega) = \begin{bmatrix} 1 & \sin(\omega T)/\omega & 0 & -(1-\cos(\omega T))/\omega \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) \\ 0 & (1-\cos(\omega T))/\omega & 1 & \sin(\omega T)/\omega \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) \end{bmatrix}$$

and $T = 3$s is the sampling period. $\mathbf{G}_k = \mathbf{I}_2 \otimes [T^2/2, T]^T$ is the gain matrix for the process noise $\mathbf{w}_k$, which is assumed to be white Gaussian with zero mean and covariance $\sigma_w^2\mathbf{I}_2$. Here, $\sigma_w = 0.1$m/s and $\otimes$ denotes the kronecker product. By setting the turn rate w to zero, the CT model reduces to the constant velocity (CV) model with the following state transition matrix

$$\mathbf{F}_k(0) = \mathbf{I}_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}. \tag{50}$$

In the whole tracking process of 600s, the target has zero turn rate, except that during the two intervals $(150\text{s}, 360\text{s})$ and $(399\text{s}, 600\text{s})$, its turn rate is changed to $\omega = -0.0147$rad/s and $\omega = 0.0208$rad/s, respectively.

A stationary sensor located at the origin is used to perform target tracking. At time k, the sensor measurement can be either linearly related to the target motion state $\mathbf{x}_k$ as

$$\mathbf{z}_k = [x_k, y_k]^T + \mathbf{n}_k, \tag{51}$$

or nonlinearly dependent on $\mathbf{x}_k$ as

$$\mathbf{y}_k = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan(x_k/y_k) \end{bmatrix} + \mathbf{m}_k. \tag{52}$$

$\mathbf{n}_k$ is the noise in the *linear* measurements and it is white Gaussian with zero mean and covariance $\mathbf{R}_{k,n} = \sigma_p^2\mathbf{I}_2$ ($\sigma_p = 5$km). On the other hand, $\mathbf{m}_k$ is the noise in the *nonlinear* measurements. It is assumed to be zero-mean Gaussian noise with covariance $\mathbf{R}_{k,m} = \text{diag}(\sigma_r^2, \sigma_\theta^2)$, where $\sigma_r = 25$km and $\sigma_\theta = 0.5^o$.

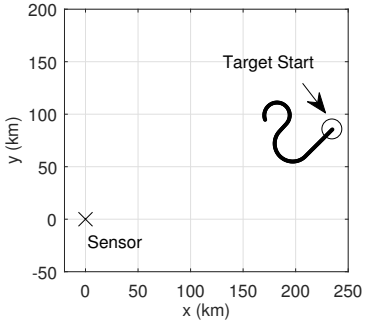

Fig. 1. The 2-D maneuvering target tracking scenario in consideration.

The three algorithms considered, IMM, GPB2 and the enhanced GPB2, all have $M = 11$ state-space models. These models have the *same* measurement equation. For the linear measurement case, the measurement equation is given in (51), while it is given in (52) for the nonlinear measurement case. The state-space models differ in their process equations. Specifically, they follow the CT model in (49) but with turn rates being 0, $\pm$0.25rad/s, $\pm$0.05rad/s, $\pm$0.025rad/s, $\pm$0.0167rad/s and $\pm$0.0125rad/s. The models have the same model probability at the beginning of the tracking process. The model transition probability matrix $\mathbf{P}$ have its diagonal elements equal to 0.95 and its off-diagonal elements equal to 0.005. In each simulation, 100 ensemble runs are conducted. The estimation root mean square error (RMSE) for the target position and the velocity RMSE *normalized* with respect to the true velocity are used as performance metrics.

B. Results and Discussions

1) Linear measurements: In the first simulation, we consider the case of linear measurements (see (51)). Every model is initialized using the measurement at time 0. Their initial model-matched state posteriors have the same Gaussian distribution with mean $[\mathbf{z}_0(1), 0, \mathbf{z}_0(2), 0]^T$ and covariance $\text{diag}(\sigma_p^2, V_{max}^2, \sigma_p^2, V_{max}^2)$, where $V_{max} = 300$m/s. The enhanced GPB2 uses $N = 1000$ samples to compute the KLD in (21) and the associated correction factor Γ_k^j via the control variate-based approach developed in Section III.B.

The results are summarized in Figs. 2 and 3, where the target position and velocity RMSEs are plotted as a function of time. It can be seen that the enhanced GPB2 algorithm offers significant performance improvement over the standard IMM and GPB2, especially during the time when the target performs turn motion. This might be explained as follows. The assumed turn rates of the models are different from those of the target. As a result, there normally exist several models with non-negligible model probabilities. Merging their model-matched state posteriors into a single Gaussian density could lead to

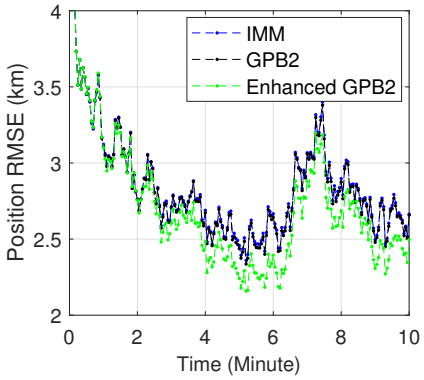

Fig. 2. Comparison of target position RMSE in the linear measurement case.

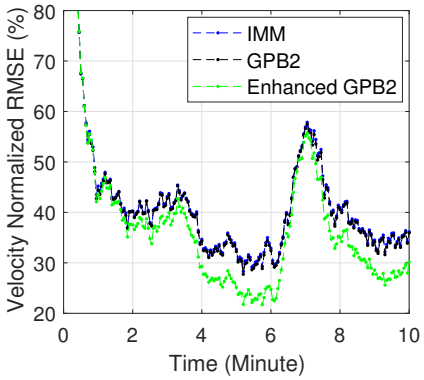

Fig. 3. Comparison of target velocity normalized RMSE in the linear measurement case.

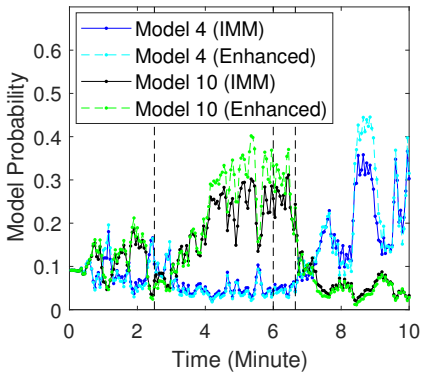

Fig. 4. Comparison of model identity probabilities of the IMM and enhanced GPB2 algorithms from a certain ensemble run in the linear measurement case.

large approximation error that needs to be accounted for in model identity probability updating, as in enhanced GPB2.

To further investigate the reasons underlying the observed performance improvement, Fig. 4 depicts as a function of time the model identity probabilities of the IMM and enhanced GPB2 algorithms from a certain ensemble run. For clarity, only the temporal evolution of the model identity probabilities of model 10 (with a turn rate $w = -0.0167$rad/s) and model 4 (with a turn rate $w = 0.025$rad/s) are shown. Their assumed turn rates are respectively the closest among all the 11 models to the true target turn rates during the two intervals (150s, 360s) and (399s, 600s). We can see from Fig. 4 that compared with the IMM technique, the use of the correction factor in the model probability updating also enables the enhanced GPB2 algorithm to emphasize more the models matching well with the target motion. This could contribute to the improvement in the state estimation performance observed in Figs. 2 and 3.

2) Nonlinear measurements: This simulation experiment examines the performance of the enhanced GPB2 algorithm when being applied to a nonlinear filtering problem. The setup is the same as the previous experiment, except that the measurements are now nonlinearly related to the target motion state (see (52)). The models are again initialized using the measurement at time 0 such that the initial model-matched state posteriors have the same Gaussian PDF with mean $[\mathbf{y}_0(1)\sin(\mathbf{y}_0(2)), 0, \mathbf{y}_0(1)\cos(\mathbf{y}_0(2)), 0]^T$. The associated covariance can be found via applying the first-order perturbation analysis, which is omitted here due to save space. When realizing the IMM, GPB2 and enhanced GPB2 algorithms, the nonlinear model-matched filtering (see (4)) and the evaluation of the model-conditioned measurement likelihood in (13) are carried out using a cubature KF (CKF) [20].

The obtained results are shown in Figs. 5 and 6. The observations are very similar to those from Figs. 2 and 3. Specifically, in the presence of nonlinear measurements, the enhanced GPB2 continues to provide superior performance over IMM and standard GPB2 when the target conducts turn motion. The model identity probabilities of models 10 and 4 from the considered IMM and enhanced GPB2 algorithms in a certain ensemble run are plotted in Fig. 7 as a function of time. Again, it can be seen that the models that better match the true target motion are assigned with higher model probabilities under the enhanced GPB2 algorithm over the IMM technique. This might be part of the reason for the performance improvement demonstrated in Figs. 5 and 6.

In this simulation, it takes the enhanced GPB2 algorithm 0.066s on average to process a newly obtained measurement. This is almost two times higher than the average amount of running time required by GPB2, which is 0.035s. The increase in the computational time mainly comes from applying the control variate-based Monte Carlo method to find the correction factor Γ_k^j (see Section III.B).

3) Effect of the number of Monte Carlo samples: In the previous two experiments, the enhanced GPB2 algorithm utilizes $N = 1000$ samples when computing the correction factor Γ_k^j. In this experiment, the value of N is varied to investigate the impact of different numbers of Monte Carlo samples used to compute Γ_k^j on the estimation accuracy of the enhanced GPB2. For this purpose, we repeat the previous simulation on tracking the maneuvering target using nonlinear measurements but with N set to be equal to 1000, 2000 and 5000. The results are shown in Figs. 8 and 9. We see that increasing the number of Monte Carlo samples does not lead to significant improvement, which indicates that with the developed control

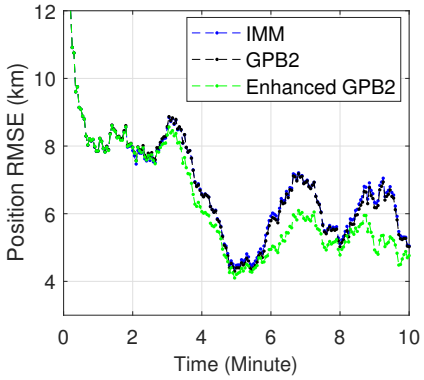

Fig. 5. Comparison of target position RMSE in the nonlinear measurement case.

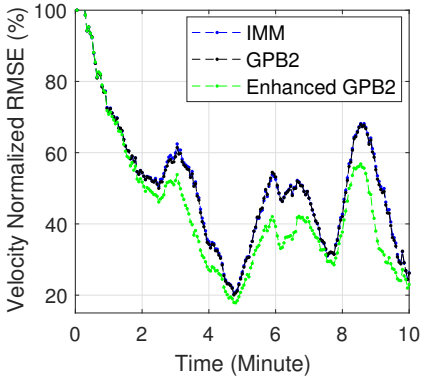

Fig. 6. Comparison of target velocity normalized RMSE in the nonlinear measurement case.

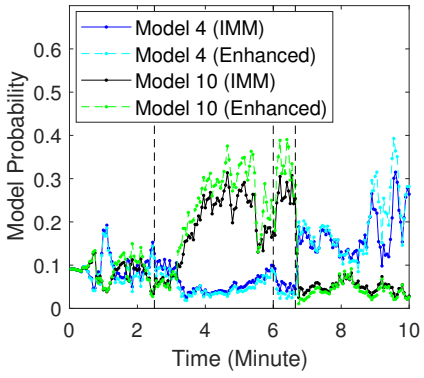

Fig. 7. Comparison of model identity probabilities of the IMM and enhanced GPB2 algorithms from a certain ensemble run in the nonlinear measurement case.

variate-based approach, the use of $N = 1000$ samples already yields a satisfactory performance.

V. Conclusions

This paper investigated the effects of the approximation used in the standard GPB2 for achieving computational tractability, which at every time instant, replaces each model-matched state posterior, a Gaussian mixture, with a single Gaussian density. It was found by re-visiting the derivation of GPB2 that this approximation is not taken into consideration when updating the model identity probabilities. Resorting to variational inference, we proposed an enhanced GPB2 algorithm that introduces a correction factor into the model identity probability updating to account for the above approximation and improve performance. The correction factor depends on the KLD between the true model-matched state posterior and the corresponding single Gaussian density approximation obtained via moment matching. As the KLD cannot be evaluated analytically, a control variate-based Monte Carlo approach was developed. Besides, the upper and lower bounds for the KLD were established and the estimated KLD is corrected if it falls out of bounds. The enhanced GPB2 algorithm was shown via simulations to be able to outperform, in a linear and a nonlinear maneuvering target tracking tasks, the standard IMM and GPB2 techniques.

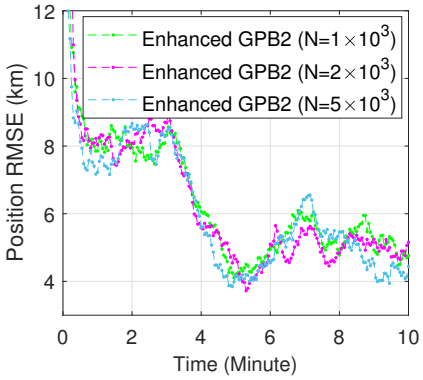

Fig. 8. Target position RMSE of the enhanced GPB2 algorithm with different number of samples in the control variate-based Monte Carlo.

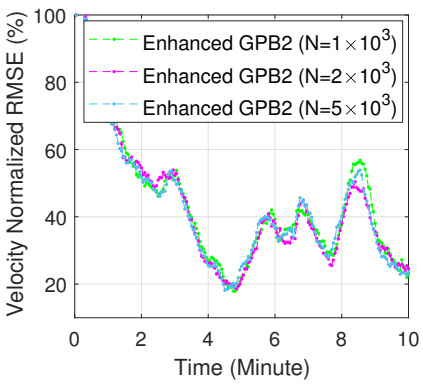

Fig. 9. Target velocity normalized RMSE of the enhanced GPB2 algorithm with different number of samples in the control variate-based Monte Carlo.

Acknowledgment

The work of Le Yang is supported by the Natural Science Foundation of China under Grant No. 61304264. Le Yang also wants to thank Yi Liu for her support, patience and love.

REFERENCES

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: Theory, algorithms and software*. New York: Wiley, 2001.

[2] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.

[3] A. H. Sayed, "A framework for state-space estimation with uncertain models," *IEEE Trans. Autom. Control*, vol. 46, pp. 998–1013, July 2001.

[4] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers, "A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices," *IEEE Trans. Autom. Control*, vol. 63, pp. 594–601, Feb. 2018.

[5] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, pp. 1333–1364, Oct. 2003.

[6] M. Zorzi, "Robust Kalman filtering under model perturbations," *IEEE Trans. Autom. Control*, vol. 62, pp. 2902–2907, June 2017.

[7] J. Zhao, "Dynamic state estimation with model uncertainties using H_∞ extended Kalman filter," *IEEE Trans. Power Syst.*, vol. 33, pp. 1099–1100, Jan. 2018.

[8] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part V. Multiple-model methods," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, pp. 1255–1321, Oct. 2005.

[9] R. R. Pitre, V. P. Jilkov, and X. R. Li, "A comparative study of multiple model algorithms for maneuvering target tracking," in *Proc. SPIE 5809, Signal Process., Sensor Fusion, and Target Recognition*, Orlando, Florida, USA, 2005.

[10] U. Orguner and M. Demirekler, "Analysis of single Gaussian approximation of Gaussian mixtures in Bayesian filtering applied to mixed multiple-model estimation," *International Journal of Control*, vol. 80, pp. 952–967, June 2007.

[11] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.

[12] J. Y. Chen, J. R. Hershey, P. A. Olsen, and E. Yashchin, "Accelerated Monte Carlo for Kullback-Leibler divergence between Gaussian mixture models," in *Proc. Intl. Conf. Acoustics, Speech and Signa Process.*, Las Vegas, Nevada, USA, April 2008.

[13] J. R. Hershey and P. A. Olsen, "Approximating the Kullback-Leibler divergence between Gaussian mixture models," in *Proc. Intl. Conf. Acoustics, Speech and Signa Process.*, Honolulu, Hawai, USA, April 2007.

[14] J. L. Durrieu, J. P. Thiran, and F. Kelly, "Lower and upper bounds for approximation of the Kullback-Leibler divergence between Gaussian mixture models," in *Proc. Intl. Conf. Acoustics, Speech and Signa Process.*, Las Vegas, Nevada, USA, April 2012.

[15] J. Johnston and V. Krishnamurthy, "An improvement to the interacting multiple model (IMM) algorithm," *IEEE Trans. Signal Process.*, vol. 49, pp. 2909–2923, Dec. 2001.

[16] Y. Ma, S. Zhao, and B. Huang, "Multiple-model state estimation based on variational Bayesian inference," *to appear in IEEE Trans. Autom. Control*, 2019.

[17] W. Li and Y. Jia, "Location of mobile station with maneuvers using an IMM-based cubature Kalman filter," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 4338–4348, Nov. 2012.

[18] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, pp. 477–482, March 2000.

[19] I. Arasaratnam, S. Haykin, and R. J. Elliott, "Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature," *Proc. IEEE*, vol. 95, no. 5, pp. 953–977, May. 2007.

[20] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.

[21] S. Bhaumik and Swati, "Cubature quadrature Kalman filter," *IET Signal Process.*, vol. 7, pp. 533–541, Sept. 2013.

[22] J. Duník, O. Straka, M. Šimandl, and E. Blasch, "Random-point-based filters: Analysis and comparison in target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, pp. 1403–1421, April 2015.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[24] L. Yang, K. Wang, and L. S. Mihaylova, "Online sparse multi-output Gaussian process regression and learning," *IEEE Trans. Signal and Info. Process. over Networks*, vol. 5, pp. 258–272, June 2019.

[25] M. J. Beal, "Variational algorithms for approximate Bayesian inference," Ph.D. dissertation, University of London, 2003.

[26] X. Chen, Z. Liu, and X. Wei, "Unambiguous parameter estimation of multiple near-field sources via rotating uniform circular array," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 872–875, 2017.

[27] X. Li, L. Yang, L. S. Mihaylova, F. Guo, and M. Zhang, "Enhanced GMM-based filtering with measurement update ordering and innovation-based pruning," in *Proc. Intl. Conf. Information Fusion (FUSION)*, July 2018, pp. 2572–2579.