Taylor & Francis
Taylor & Francis Group

International Journal of Control

ISSN: 0020-7179 (Print) 1366-5820 (Online) Journal homepage: http://www.tandfonline.com/loi/tcon20

Analysis of single Gaussian approximation of Gaussian mixtures in Bayesian filtering applied to mixed multiple-model estimation

U. Orguner & M. Demırekler

To cite this article: U. Orguner & M. Demırekler (2007) Analysis of single Gaussian approximation of Gaussian mixtures in Bayesian filtering applied to mixed multiple-model estimation, International Journal of Control, 80:6, 952-967, DOI: 10.1080/00207170701261952

To link to this article: http://dx.doi.org/10.1080/00207170701261952

International Journal of Control
Vol. 80, No. 6, June 2007, 952–967

Analysis of single Gaussian approximation of Gaussian mixtures in Bayesian filtering applied to mixed multiple-model estimation

U. ORGUNER* and M. DEMİREKLER

Department of Electrical and Electronics Engineering,
Middle East Technical University, TR-06531, Ankara, Turkey

(*Received 15 June 2006; in final form 5 February 2007*)

This paper examines the effect of the moment-matched single Gaussian approximation, which is made in various multiple-model filtering applications to approximate a Gaussian mixture, on the Bayesian filter performance. The estimation error caused by the approximation is analysed for both the prediction and the measurement updates of a Bayesian filter. An approximate formula is found for the covariance of the error caused by the approximation for a general Gaussian mixture with arbitrary components. The calculated error covariance is used for obtaining a mixed multiple-model estimation algorithm which has a performance near that of GPB2 with less computations.

1. Introduction

Kalman filter is the universally adopted state estimation technique used in many applications (Gelb 1974, Anderson and Moore 1979). On the other hand, its underlying assumptions are generally too strict for most of the cases. Specifically, the assumption of perfectly known system dynamics and measurement model parameters is usually problematic due to the inherent uncertainty involved in the modelling process. Multiple-model techniques are proposed for solving this problem in the cases where the model uncertainties can be covered by a finite number of models. However, the computational cost of obtaining the optimal minimum mean-square error (MMSE) estimate of the state in multiple-model configurations increases exponentially in time (Bar-Shalom and Li 1993). Therefore, approximations are necessarily made to obtain suboptimal but computationally cheaper estimates. The most well-known examples of these suboptimal approaches are the generalized pseudo Bayesian (GPB) (Ackerson and Fu 1970; Tugnait, 1982) and the interacting multiple model (IMM) (Blom and Bar-Shalom 1988) algorithms. In those algorithms, the multiple component Gaussian mixtures are approximated by a single Gaussian, matching the mean and covariance of the densities. In this paper, the effect of this approximation applied to the input of a optimal Bayesian filter is investigated. The first and second moments of the error between the estimate resulting from the approximation (which is the estimate of a Kalman filter) and MMSE-optimal estimate which is obtained from Bayesian density recursions are examined. An analytical expression for the covariance of the resulting error due to the approximation in the filter estimate which can be calculated before filtering, is found. This measure is then used as a tool for generating a mixed IMM-GPB2 algorithm. The resulting mixed algorithm can combine the two methods while achieving the performance of the better-performance algorithm (GPB2) with a computational load near to that of the computationally cheaper algorithm (IMM).

The outline of the paper is as follows. In § 2, the problem definition is made and the optimal solution in terms of Bayesian density updates is given. The estimate after the approximation is obtained in § 3. Section 4 is composed of the investigation of the first two moments of the error between optimal and the approximate estimates. Up to § 5, where we proceed with a brief

*Corresponding author. Email: umutor@yahoo.com

International Journal of Control
ISSN 0020–7179 print/ISSN 1366–5820 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00207170701261952

review of the IMM and GPB2 algorithms, quite a stand-alone presentation is adopted in order to make the analysis accessible to the readers with limited multiple model estimation background. Section 5 is followed by the introduction of a mixed IMM-GPB2 algorithm in § 6. In § 7, the simulation results are presented for the mixed IMM-GPB2 algorithm. The paper is finalized with conclusions in § 8.

2. Problem definition and optimal solution

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let x_k, y_k, w_k and v_k be random variables defined on this space. We assume that the probability density functions for all the defined random variables exist and that all the random variables are integrable. Suppose we are given the following discrete-time state-space representation

$$x_{k+1} = Ax_k + Bw_k, \tag{1}$$

$$y_k = Cx_k + Dv_k, \tag{2}$$

where w_k and v_k are uncorrelated, white Gaussian noises with zero mean and covariance Q and R respectively. The prior density $p(x_0)$ for the initial state x_0 is the Gaussian mixture which is

$$p(x_0) = \sum_{i=1}^{N} p_i \mathcal{N}(x_0; \bar{x}_i, P_i), \tag{3}$$

where the notation $\mathcal{N}(x; \bar{x}, P)$ denotes the multivariate normal density with dummy variable x, mean $\bar{x}$ and covariance P. The component probabilities $\{p_i\}_{i=1}^N$ sum up to unity, i.e., $\Sigma_{i=1}^N p_i = 1$.

2.1 *Problem definition*

We are interested the optimal MMSE estimate $\hat{x}_{k|k}$ and its covariance $P_{k|k}$ given as

$$\hat{x}_{k|k} = E\{x_k | Y^k\}, \tag{4}$$

$$P_{k|k} = E\left\{\left(x_k - \hat{x}_{k|k}\right)\left(x_k - \hat{x}_{k|k}\right)^T \middle| Y^k\right\}, \tag{5}$$

where

$$Y^k \triangleq \{y_0, y_1, \ldots, y_k\} \tag{6}$$

or more rigorously Y^k denotes the σ-algebra generated by the random variables $y_0, y_1, \ldots, y_k$. In the remaining parts of the paper, the prediction and measurement updates for the Bayesian (or Kalman) filters will be examined separately. In each case, we assume that we are at an intermediate stage of a recursive estimation process and our input density (information state to be updated) is a Gaussian mixture. Therefore, in prediction (measurement) update, it is assumed that the input density $p(x_{k-1}| Y^{k-1})$ $(p(x_k| Y^{k-1}))$ is a Gaussian mixture. In a non-recursive framework, the input densities correspond to prior density of the estimatee at the corresponding level of estimation.

2.2 *Optimal solution*

The optimal solution to the problem stated above is well-known and given, for example, in Sorenson and Alspach (1971) or Anderson and Moore (1979). In the following, we simply state these facts.

2.2.1 Prediction update. The general Bayesian density update equation for MMSE prediction update is given as follows (Kumar and Varaiya 2000):

$$p\left(x_k | Y^{k-1}\right) = \int p(x_k|x_{k-1}) p\left(x_{k-1} | Y^{k-1}\right) \mathrm{d}x_{k-1}. \tag{7}$$

Let $p(x_{k-1} \mid Y^{k-1})$ be the N component Gaussian mixture

$$p\left(x_{k-1} | Y^{k-1}\right) = \sum_{i=1}^{N} p_i \mathcal{N}(x_{k-1}; \bar{x}_i, P_i). \tag{8}$$

Using the Kalman filter prediction update formulas, the prediction density $p(x_k| Y^{k-1})$ can be obtained as

$$p\left(x_k | Y^{k-1}\right) = \sum_{i=1}^{N} p_i \mathcal{N}\left(x_k; \bar{x}_i^-, P_i^-\right), \tag{9}$$

where

$$\bar{x}_i^- = A\bar{x}_i, \tag{10}$$

$$P_i^- = AP_iA^T + BQB^T \tag{11}$$

for $i = 1, \ldots, N$. The optimal MMSE prediction $\hat{x}_{k|k-1}^{op}$ and its covariance $P_{k|k-1}^{op}$ is therefore given by

$$\hat{x}_{k|k-1}^{op} = \sum_{i=1}^{N} p_i \bar{x}_i^-, \tag{12}$$

$$P_{k|k-1}^{op} = \sum_{i=1}^{N} p_i \left[P_i^- + \left(\bar{x}_i^- - \hat{x}_{k|k-1}^{op}\right)\left(\bar{x}_i^- - \hat{x}_{k|k-1}^{op}\right)^T\right]. \tag{13}$$

2.2.2 Measurement update. The general Bayesian density update equation for MMSE measurement update is given as follows (Kumar and Varaiya 2000):

$$p\left(x_k | Y^k\right) = \frac{p(y_k|x_k) p\left(x_k | Y^{k-1}\right)}{\int p(y_k|x_k) p(x_k | Y^{k-1}) \mathrm{d}x_k}. \tag{14}$$

Let $p(x_k|Y^{k-1})$ be the N component Gaussian mixture

$$p\left(x_k|Y^{k-1}\right) = \sum_{i=1}^{N} p_i \mathcal{N}(x_k; \bar{x}_i, P_i). \tag{15}$$

Using the Kalman filter measurement update formulas and defining the probabilities

$$p_i^+ \triangleq \frac{p_i \mathcal{N}(y_k; \bar{y}_i, S_i)}{\sum_{j=1}^{N} p_j \mathcal{N}\left(y_k; \bar{y}_j, S_j\right)}, \tag{16}$$

the posterior density $p(x_k|Y^k)$ is given as

$$p\left(x_k|Y^k\right) = \sum_{i=1}^{N} p_i^+ \mathcal{N}\left(x_k; \bar{x}_i^+, P_i^+\right), \tag{17}$$

where

$$\bar{x}_i^+ = \bar{x}_i + K_i(y_k - C\bar{x}_i) \triangleq \bar{x}_i + K_i(y_k - \bar{y}_i), \tag{18}$$

$$P_i^+ = P_i - K_i S_i K_i^T = P_i - P_i C^T S_i^{-1} C P_i \tag{19}$$

with

$$\bar{y}_i = C\bar{x}_i, \tag{20}$$

$$K_i = P_i C^T S_i^{-1} = P_i^+ C^T R^{-1}, \tag{21}$$

$$S_i = C P_i C^T + D R D^T. \tag{22}$$

Calculating the mean and the covariance of this density gives us the following optimal filtered estimate $\hat{x}_{k|k}^{op}$ and covariance $P_{k|k}^{op}$ as

$$\hat{x}_{k|k}^{op} = \sum_{i=1}^{N} p_i^+ \bar{x}_i^+, \tag{23}$$

$$P_{k|k}^{op} = \sum_{i=1}^{N} p_i^+ \left[P_i^+ + \left(\bar{x}_i^+ - \hat{x}_{k|k}^{op}\right)\left(\bar{x}_i^+ - \hat{x}_{k|k}^{op}\right)^T \right]. \tag{24}$$

3. Approximate solution

In this section, we are going to approximate the Gaussian mixtures which are input to the prediction and measurement updates of the optimal Bayesian recursions in the previous section by a single Gaussian whose mean and covariance are matched to those of the Gaussian mixture. Therefore, we will assume in each update, the input Gaussian mixture

$$p(x) = \sum_{i=1}^{N} p_i \mathcal{N}(x, \bar{x}_i, P_i) \tag{25}$$

is approximated by

$$p_{app}(x) = \mathcal{N}\left(x, \bar{x}_{app}, P_{app}\right), \tag{26}$$

where

$$\bar{x}_{app} \triangleq \sum_{i=1}^{N} p_i \bar{x}_i, \tag{27}$$

$$P_{app} \triangleq \sum_{i=1}^{N} p_i \left[P_i + \left(\bar{x}_i - \bar{x}_{app}\right)\left(\bar{x}_i - \bar{x}_{app}\right)^T \right]. \tag{28}$$

The Bayesian recursions after this approximation simply turn into the standard Kalman filter updates.

3.1 *Prediction update*

Now, assuming that the input Gaussian mixture

$$p\left(x_{k-1}|Y^{k-1}\right) = \sum_{i=1}^{N} p_i \mathcal{N}(x_{k-1}, \bar{x}_i, P_i) \tag{29}$$

is approximated by a single Gaussian

$$p\left(x_{k-1}|Y^{k-1}\right) \approx \mathcal{N}\left(x_{k-1}; \bar{x}_{app}, P_{app}\right), \tag{30}$$

where $\bar{x}_{app}$ and P_{app} are given in (27) and (28) respectively, the suboptimal prediction resulting from this approximation is

$$\hat{x}_{k|k-1}^{sub} = A\bar{x}_{app}, \tag{31}$$

$$P_{k|k-1}^{sub} = A P_{app} A^T + BQB^T. \tag{32}$$

Substituting $\bar{x}_{app}$ given in (27) into (31),

$$\hat{x}_{k|k-1}^{sub} = A\sum_{i=1}^{N} p_i \bar{x}_i = \sum_{i=1}^{N} p_i A\bar{x}_i = \sum_{i=1}^{N} p_i \bar{x}_i^- = \hat{x}_{k|k-1}^{op}. \tag{33}$$

Substituting P_{app} in (28) into (32), we obtain

$$P_{k|k-1}^{sub} = A\sum_{i=1}^{N} p_i \left[P_i + \left(\bar{x}_i - \bar{x}_{app}\right)\left(\bar{x}_i - \bar{x}_{app}\right)^T \right] A^T + BQB^T, \tag{34}$$

$$= \sum_{i=1}^{N} p_i \Big[A P_i A^T + BQB^T$$

$$+ A(\bar{x}_i - \bar{x}_{app})(\bar{x}_i - \bar{x}_{app})^T A^T \Big], \tag{35}$$

$$= \sum_{i=1}^{N} p_i \left[P_i^- + \left(\bar{x}_i^- - A\bar{x}_{app}\right)\left(\bar{x}_i^- - A\bar{x}_{app}\right)^T \right], \tag{36}$$

$$= \sum_{i=1}^{N} p_i \left[P_i^- + \left(\bar{x}_i^- - \hat{x}_{k|k-1}^{sub}\right)\left(\bar{x}_i^- - \hat{x}_{k|k-1}^{sub}\right)^T \right], \tag{37}$$

$$= \sum_{i=1}^{N} p_i \left[P_i^- + \left(\bar{x}_i^- - \hat{x}_{k|k-1}^{op} \right) \left(\bar{x}_i^- - \hat{x}_{k|k-1}^{op} \right)^T \right] = P_{k|k-1}^{op}. \tag{38}$$

These two results prove that the Kalman filter prediction update yields the same first two moments as the Bayesian density recursion under the moment matched single Gaussian approximation. In other words, the single Gaussian approximation causes no change in the MMSE estimate and covariance. This invariance property of the Kalman filter prediction update is already known and is the main motive underlying the IMM-type hypothesis merging (instead of GPB2-type merging) (Bar-Shalom and Li 1993).

3.2 *Measurement update*

Assuming that the input Gaussian mixture

$$p(x_k | Y^{k-1}) = \sum_{i=1}^{N} p_i \mathcal{N}(x_k, \bar{x}_i, P_i) \tag{39}$$

is approximated by a single Gaussian

$$p(x_k | Y^{k-1}) \approx \mathcal{N}(x_k; \bar{x}_{opp}, P_{app}) \tag{40}$$

where $\bar{x}_{app}$ and P_{app} are given in (27) and (28) respectively, the Kalman filter measurement update gives the suboptimal estimate and covariance as

$$\hat{x}_{k|k}^{sub} = \bar{x}_{app} + K_{sub}(y_k - C\bar{x}_{app}), \tag{41}$$

$$P_{k|k}^{sub} = P_{app} - K_{sub} S_{sub} K_{sub}^T = P_{app} - P_{app} C^T S_{sub}^{-1} C P_{app}, \tag{42}$$

where

$$K_{sub} = P_{app} C^T S_{sub}^{-1} = P_{k|k}^{sub} C^T R^{-1}, \tag{43}$$

$$S_{sub} = C P_{app} C^T + DRD^T = \sum_{i=1}^{N} p_i \left[S_i + (\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^T \right], \tag{44}$$

$$\bar{y} = \sum_{i=1}^{N} p_i \bar{y}_i. \tag{45}$$

Unfortunately, these equations show that the optimal estimate and covariance given in (23) and (24) respectively are different than the sub-optimal ones above. Therefore, under the moment-matched single Gaussian approximation, using the standard Kalman filter measurement update (unlike the prediction update) causes our final estimate to deviate from the optimal MMSE one.

This type of approximation appears in IMM and GPB filtering used especially in maneuvering target tracking and possibly in other multiple-model estimation applications. Specifically, this approximation is made during the mixing process of the IMM filter at the input of each component Kalman filter and it is the main approximation that makes the IMM filter different from the GPB2 filter. Up to now, the difference between the two filters was analysed only by means of Monte-Carlo runs (Blom and Bar-Shalom 1988) and to the authors' knowledge, our work is the first one which tries to quantify the deviation analytically. In the next section, the first two moments of the error caused by the moment-matched single Gaussian approximation will be examined.

4. Error analysis

In the previous section, it was shown that the Kalman filter measurement update results in erroneous estimate and covariance under moment-matched single Gaussian approximation unlike the prediction update whose resulting estimate (i.e., prediction) is invariant under the same approximation. Defining the error Δ caused by the measurement update (in an IMM framework, this error may correspond to the error caused in the filtered estimates of individual (component) Kalman filters by the moment-matched single Gaussian approximation made in the mixing process at the input of the filters) as

$$\Delta \triangleq \hat{x}_{k|k}^{sub} - \hat{x}_{k|k}^{op}, \tag{46}$$

in this section, we are interested in the conditional expected value

$$\bar{\Delta} \triangleq E_{y_k}\{\Delta | Y^{k-1}\} \tag{47}$$

and the conditional covariance

$$\sum\nolimits_{\Delta} \triangleq E_{y_k}\left\{ (\Delta - \bar{\Delta})(\Delta - \bar{\Delta})^T \middle| Y^{k-1} \right\} \tag{48}$$

of this error. Note that the expectations given above are to be taken with respect to y_k which has a density

$$\begin{aligned} p(y_k | Y^{k-1}) &= \int p(y_k | x_k) p(x_k | Y^{k-1}) \mathrm{d}x_k \\ &= \sum_{j=1}^{N} p_j \mathcal{N}(y_k; \bar{y}_j, S_j) \end{aligned} \tag{49}$$

Therefore, the density with respect to which the expectations are to be taken is also a Gaussian mixture.

4.1 *Calculation of the mean of* "

In order to find $\bar{\Delta}$, we need the expected values of the optimal and sub-optimal estimates

$$E_{yk}\left\{\hat{x}^{sub}_{k|k}|\,Y^{k-1}\right\}$$
$$= E_{y_k}\left\{\bar{x}_{app} + K_{sub}\big(y_k - C\bar{x}_{app}\big)\big|\,Y^{k-1}\right\}, \tag{50}$$

$$= \bar{x}_{app} + K_{sub}\big(E_{y_k}\left\{y_k|\,Y^{k-1}\right\} - C\bar{x}_{app}\big), \tag{51}$$

$$= \bar{x}_{app} + K_{sub}\left(\sum_{i=1}^{N} p_i\bar{y}_i - C\bar{x}_{app}\right). \tag{52}$$

Replacing $C\bar{x}_{app}$ by $\Sigma_{i=1}^{N} p_i C\bar{x}_{app}$,

$$E_{y_k}\left\{\hat{x}^{sub}_{k|k}|\,Y^{k-1}\right\} = \bar{x}_{app}. \tag{53}$$

Note that this is just a rephrasing of the fact that

$$E_{y_k}\left\{\hat{x}_{k|k}|\,Y^{k-1}\right\} = E_{y_k}\left\{E\left\{x_k|\,Y^k\right\}|\,Y^{k-1}\right\}$$
$$= E\left\{x_k|\,Y^{k-1}\right\} = \hat{x}_{k|k-1}. \tag{54}$$

In a more technical language, this is the manifestation of the fact that, for an integrable random variable ξ, the sequence of random variables ξ_n defined as

$$\xi_n \triangleq E\{\xi|\mathcal{F}_n\} \tag{55}$$

is a martingale with respect to the filtration $\{\mathcal{F}_n\}_{n=1}^{\infty}$ (Brzeźniak and Zastawniak 1999). In our case, $\xi = x_k$ and $\mathcal{F}_n$ is the σ-field generated by Y^n. Another interpretation is that the Kalman filter and the optimal Bayesian density measurement updates are unbiased estimators. Using the same fact, we can easily conclude that

$$E_{yk}\left\{\hat{x}^{op}_{k|k}|\,Y^{k-1}\right\} = \bar{x}_{app}. \tag{56}$$

However, we are going to prove this fact the long way as well to obtain some expected values that will be utilized in the covariance calculation in the next subsection and the appendix

$$E_{y_k}\left\{\hat{x}^{op}_{k|k}|\,Y^{k-1}\right\} = E_{y_k}\left\{\sum_{i=1}^{N} p_i^+\bar{x}_i^+|\,Y^{k-1}\right\}, \tag{57}$$

$$= E_{y_k}\left\{\sum_{i=1}^{N} p_i^+[\bar{x}_i + K_i(y_k - C\bar{x}_i)]|\,Y^{k-1}\right\}, \tag{58}$$

$$= \sum_{i=1}^{2} E_{y_k}\left\{p_i^+|\,Y^{k-1}\right\}[I - K_iC]\bar{x}_i + K_iE_{y_k}\left\{p_i^+y_k|\,Y^{k-1}\right\}. \tag{59}$$

Therefore, for the evaluation of this expected value, we need other intermediate exected values $E_{y_k}\{p_i^+|\,Y^{k-1}\}$ and $E_{y_k}\{p_i^+y_k|\,Y^{k-1}\}$. Note that

$$p_i^+ \triangleq \frac{p_i\mathcal{N}(y_k;\,\bar{y}_i, S_i)}{\sum_{j=1}^{N} p_j\mathcal{N}\big(y_k;\,\bar{y}_j, S_j\big)} = \frac{p_i\mathcal{N}(y_k;\,\bar{y}_i, S_i)}{p(y_k|\,Y^{k-1})}. \tag{60}$$

Then,

$$E_{y_k}\left\{p_i^+|\,Y^{k-1}\right\} \triangleq \int p_i^+p\big(y_k|\,Y^{k-1}\big)\mathrm{d}y_k$$
$$= p_i\int\mathcal{N}(y_k;\,\bar{y}_i, S_i)\mathrm{d}y_k = p_i. \tag{61}$$

As a result,

$$E_{y_k}\left\{p_i^+|\,Y^{k-1}\right\} = p_i. \tag{62}$$

In the same way,

$$E_{y_k}\left\{p_i^+y_k|\,Y^{k-1}\right\} \triangleq \int p_i^+y_kp\big(y_k|\,Y^{k-1}\big)\mathrm{d}y_k$$
$$= p_i\int y_k\mathcal{N}(y_k;\,\bar{y}_i, S_i)\mathrm{d}y_k = p_i\bar{y}_i. \tag{63}$$

Consequently,

$$E_{y_k}\left\{p_i^+y_k|\,Y^{k-1}\right\} = p_i\bar{y}_i. \tag{64}$$

Using these expected values, we calculate the expected value of the optimal filtered estimate $\hat{x}^{op}_{k|k}$ as

$$E_{y_k}\left\{\hat{x}^{op}_{k|k}|\,Y^{k-1}\right\} = \sum_{i=1}^{N} p_i[I - K_iC]\bar{x}_i + p_iK_i\bar{y}_i$$
$$= \sum_{i=1}^{N} p_i\bar{x}_i = \bar{x}_{app}. \tag{65}$$

Having calculated the required expected values of the sub-optimal and optimal estimates, the mean of Δ is given as

$$\bar{\Delta} \triangleq E_{y_k}\left\{\Delta|\,Y^{k-1}\right\} = 0. \tag{66}$$

which means that the moment-matched single Gaussian approximation does not cause any bias in the estimate.

4.2 *Calculation of the covariance of* "

The covariance Σ_Δ of Δ then can be calculated by

$$\sum\nolimits_\Delta = E_{y_k}\left\{\Delta\Delta^T|\,Y^{k-1}\right\} \tag{67}$$

where

$$\Delta \triangleq \hat{x}^{sub}_{k|k} - \hat{x}^{op}_{k|k} = \bar{x}_{app} + K_{sub}\big(y_k - C\bar{x}_{app}\big) - \sum_{i=1}^{N} p_i^+ \bar{x}_i^+, \tag{68}$$

$$= \bar{x}_{app} + K_{sub}(y_k - \bar{y}) - \sum_{i=1}^{N} p_i^+ [\bar{x}_i + K_i(y_k - C\bar{x}_i)], \tag{69}$$

$$= \bar{x}_{app} + K_{sub}(y_k - \bar{y}) - \sum_{i=1}^{N} p_i^+ [\bar{x}_i + K_i(y_k - \bar{y}_i)]. \tag{70}$$

Since the expression is too long, we are going to assign the terms to some auxiliary variables. Define

$$T_1 \triangleq \bar{x}_{app}, \tag{71}$$

$$T_2 \triangleq K_{sub}(y_k - \bar{y}), \tag{72}$$

$$T_3 \triangleq \sum_{i=1}^{N} p_i^+ \bar{x}_i, \tag{73}$$

$$T_4 \triangleq \sum_{i=1}^{N} p_i^+ K_i(y_k - \bar{y}_i). \tag{74}$$

Then,

$$\sum_{\Delta} = E_{y_k}\{(T_1 + T_2 - T_3 - T_4)(T_1 + T_2 - T_3 - T_4)^T | Y^{k-1}\}. \tag{75}$$

Thus, the covariance calculation requires the expected values of the form

$$E_{y_k}\left\{T_i T_j^T | Y^{k-1}\right\} \quad \text{for} \quad i, j = 1, \ldots, 4. \tag{76}$$

The evaluation of these expected values involves second order marginal and cross moments of the posterior probabilities p_i^+ and y_k with respect to the density $p(y_k | Y^{k-1})$. These expected values are extremely difficult (if not impossible) to evaluate analytically, and therefore, some approximations have to be made. The evaluation of the second marginal moment of the probability p_i^+ shown below illustrates the approximation used to calculate these integrals

$$E_{y_k}\left\{\left(p_i^+\right)^2 | Y^{k-1}\right\} \triangleq \int \left(p_i^+\right)^2 p\big(y_k | Y^{k-1}\big) \mathrm{d}y_k$$

$$= \int \frac{p_i^2 \mathcal{N}^2(y_k; \bar{y}_i, S_i)}{\sum_{j=1}^{N} p_j \mathcal{N}\big(y_k; \bar{y}_j, S_j\big)} \mathrm{d}y_k, \tag{77}$$

$$= p_i \int \frac{p_i \mathcal{N}(y_k; \bar{y}_i, S_i)}{\sum_{j=1}^{N} p_j \mathcal{N}\big(y_k; \bar{y}_j, S_j\big)} \mathcal{N}(y_k; \bar{y}_i, S_i) \mathrm{d}y_k. \tag{78}$$

At this point, we see that the numerator of the integrand is the square of a Gaussian density which decreases quite fast. Due to this, the integration is effectively around the mean value $\bar{y}_i$. Assuming that the means $\bar{y}_i$ of the Gaussian components are separated sufficiently, in the effective integration range, the Gaussian mixture in the denominator can be approximated as

$$\sum_{j=1}^{N} p_j \mathcal{N}\big(y_k; \bar{y}_j, S_j\big) \approx p_i \mathcal{N}(y_k; \bar{y}_i, S_i). \tag{79}$$

This corresponds too assuming that the measurement dependent posterior probability p_i^+ is approximately unity around the mean $\bar{y}_i$ within the 2σ covariance ellipse. Many other approximation schemes such as Taylor series expansions etc. are also possible but they turn out to yield negative variance values for the probabilities p_i^+ in some extreme cases. After the approximation,

$$E_{y_k}\left\{\left(p_i^+\right)^2 | Y^{k-1}\right\} \approx p_i \int \mathcal{N}(y_k; \bar{y}_i, S_i) \mathrm{d}y_k = p_i. \tag{80}$$

Using this identity, we can calculate the variance $\sigma^2_{p_i^+}$ of the probability p_i^+ as

$$\sigma^2_{p_i^+} = E_{y_k}\left\{\left(p_i^+\right)^2 | Y^{k-1}\right\} - E^2_{y_k}\{p_i^+ | Y^{k-1}\} \approx p_i - p_i^2. \tag{81}$$

This variance vs. the prior probability p_i is shown in figure 1. With the help of the same approximation, the expected values defined in (76) can be evaluated easily. After some involved algebraic manipulations which are presented in Appendix A, the covariance of the error resulting from moment-matched single Gaussian

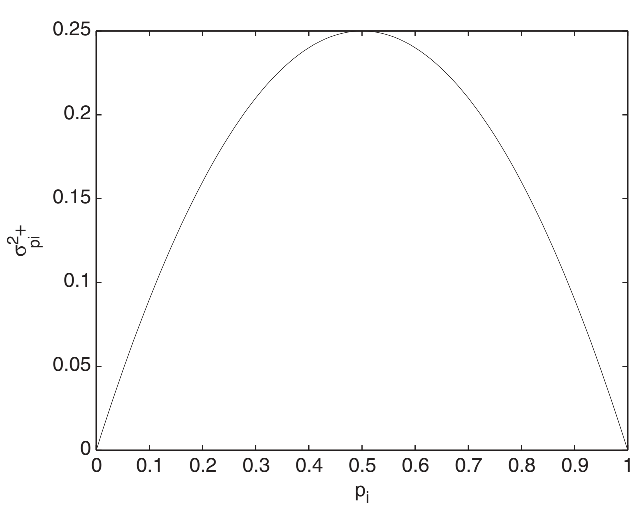

Figure 1. Variance of the probability p_i^+ vs. p_i.

approximation is found to be

$$\sum\nolimits_{\Delta} = \sum_{i=1}^{N} p_i P_i C^T S_i^{-1} C P_i - P_{app} C^T S_{sub}^{-1} C P_{app} + \sum_{i=1}^{N} p_i \bar{x}_i \bar{x}_i^T - \bar{x}_{app} \bar{x}_{app}^T \tag{82}$$

where

$$\bar{x}_{app} = \sum_{i=1}^{N} p_i \bar{x}_i, \tag{83}$$

$$\bar{y} = C \bar{x}_{app}, \tag{84}$$

$$S_{sub} = \sum_{i=1}^{N} p_i \big[S_i + (\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^T \big], \tag{85}$$

$$P_{app} = \sum_{i=1}^{N} p_i \Big[P_i + \big(\bar{x}_i - \bar{x}_{app}\big)\big(\bar{x}_i - \bar{x}_{app}\big)^T \Big]. \tag{86}$$

Although the covariance formula in equation (82) is calculated using some intense approximations (a more explicit illustration of the amount of approximations can be observed in the derivation given in Appendix A), it is still in accordance with intuition. Notice that when only a single component probability p_i is dominant in the mixture, i.e., $p_i \approx 1$, the error covariance goes to zero since $P_{app} \approx P_i$, $S_{sub} \approx S_i$ and $\bar{x}_{app} \approx \bar{x}_i$. This is intuitively clear since when one probability is dominant, the mixture can be approximated well by the single Gaussian corresponding to the dominant probability. A similar case appears when values of all the means and covariances of the mixture are very close to each other. In that case, the covariance formula again yields a "small" matrix. Moreover, when the means of the component densities in the mixture are well separated, the error covariance will be calculated with high accuracy. (It must be mentioned that, sufficient separation of means is to be interpreted taking the mixture covariances into account. For example, a mixture with two components with a fixed Euclidian distance between its means can either satisfy or violate the sufficient separation of means assumption according to the covariances of its components. If the covariances are "small" then it satisfies the assumption and viceversa.) Considering the two sets of Gaussian mixtures, for which the (82) would give correct error covariance values as described above, as sets of points on a high dimensional space, our error covariance calculation formula, being a function on this space, in a way, interpolates between these points to cover the intermediate cases. (The first set is composed of Gaussian mixtures which would give zero error covariance, e.g. The Gaussian mixtures where $p_i = 1$ for some i or the mixtures with the same component means and covariances. The second set is composed of Gaussian mixtures with infinite mean separation. These two sets represent the two extreme cases where our covariance formula would give correct results.) Obviously, this interpolation would give biased results due to the separated means assumption which can be hard to validate for most of the points on the described high dimensional space of Gaussian mixtures. However, it still gives us a quantification of the error and using thresholding mechanisms, it can be utilized in obtaining new multiple model estimation algorithms as shown in the subsequent parts of this paper.

4.3 *Discussion*

It must be emphasized here that the error caused by the moment-matched single Gaussian approximation in Bayesian filtering has already been known to exist in the literature. Especially, Bar-Shalom and Li (1993) are seen to be well-aware of its existence while deriving the IMM algorithm. In that derivation, they simply bypass the measurement update of each Kalman filter and then introduce the moment-matched single Gaussian approximation into the predicted state densities which are shown above to be transparent to the approximation. This deliberate selection of timing (of introducing the approximation) is an implicit consequence of the fact that the measurement update of the Kalman filter introduces errors into the estimate. Our analysis above is an explicit presentation of these observations and moreover it makes a quantification possible by means of the error covariance formula in (82).

An interesting observation about the IMM algorithm implied by our analysis is that its estimate covariances are optimistic. In IMM, although the means and covariances at the input of the Kalman filters are matched at each cycle, since the algorithm is not (made to be) aware of (errors caused by) the single Gaussian approximation, its calculated covariances must be "smaller" than they need to be. Using the same reasoning, one can come to the same conclusion about all mixing and/or merging based multiple model estimation algorithms (including GPB families). It is the authors' belief that this study will attract the attention of the community to introduce ways of compensating for these errors. The covariance formula in (82) can be used as a beginning for these future studies.

5. Review of the IMM and GPB2 algorithms

In this section, we are going to present a brief review of IMM and GPB2 algorithms. In this way, the approximations made in both of the algorithms and the differences will be illustrated. The IMM and GPB2 algorithms are used in the cases where the model uncertainty can be covered by a finite number of models. Each algorithm uses multiple Kalman filters matched to the models. We begin this section by defining the following so-called jump Markov linear system given below,

$$x_k = A_{r_k} x_{k-1} + B_{r_k} w_k, \tag{87}$$

$$y_k = C_{r_k} x_k + D_{r_k} v_k. \tag{88}$$

The parameter matrices $\{A_i,\ B_i,\ C_i,\ D_i\}$ are assumed to be known for $i \in \{1, \dots, N\}$ and $r_k \in \{1, \dots, N\}$ is a finite-state Markov chain with initial distribution $\pi_0 = [\pi_0^1, \pi_0^2 \dots, \pi_0^N]$ and probability transition matrix $\Pi = [\pi_{ij}]_{i,j=1}^N$. The noise processes w_k and v_k are white, uncorrelated and normally distributed with zero mean and covariances Q and R respectively. The following descriptions of the IMM and GPB2 algorithms have been adapted from Bar-Shalom and Li (1993).

5.1 *Steps of the GPB2 algorithm*

At each time step, an N-model GPB2 algorithm keeps N filtered estimates $\{\hat{x}^j_{k|k}\}_{j=1}^N$, covariances $\{P^j_{k|k}\}_{j=1}^N$ and mode probabilities $\{\mu^j_k\}_{j=1}^N$. At each cycle, given the previously kept filtered estimates $\{\hat{x}^j_{k-1|k-1}\}_{j=1}^N$, covariances $\{P^j_{k-1|k-1}\}_{j=1}^N$ and the mode probabilities $\{\mu^j_{k-1}\}_{j=1}^N$, the algorithm calculates the updated values of these quantities using the following steps.

- **Mode-matched Kalman filtering:** The algorithm takes each previous filtered estimate $\hat{x}^i_{k-1|k-1}$ and covariance $P^i_{k-1|k-1}$ and executes N Kalman filters each matched to a different model. All N Kalman filters use the filtered estimate $\hat{x}^i_{k-1|k-1}$ and covariance $P^i_{k-1|k-1}$ as their initial state and covariance. Since there are N previous filtered estimates, a total of N^2 Kalman filters are executed. At the end of the filtering, new filtered estimates $\hat{x}^{ij}_{k|k}$ and covariances $P^{ij}_{k|k}$ are obtained. The filtering equations are given as follows:

$$\hat{x}^{ij}_{k|k-1} = A_j \hat{x}^i_{k-1|k-1}, \tag{89}$$

$$P^{ij}_{k|k-1} = A_j P^i_{k-1|k-1} A_j^T + B_j Q B_j^T, \tag{90}$$

$$S^{ij}_k = C_j P^{ij}_{k|k-1} C_j^T + D_j R D_j^T, \tag{91}$$

$$K^{ij}_k = P^{ij}_{k|k-1} C_j^T \left(S^{ij}_k\right)^{-1}, \tag{92}$$

$$\hat{x}^{ij}_{k|k} = \hat{x}^{ij}_{k|k-1} + K^{ij}_k \left(y_k - C_j \hat{x}^{ij}_{k|k-1}\right), \tag{93}$$

$$P^{ij}_{k|k} = P^{ij}_{k|k-1} - K^{ij}_k S^{ij}_k \left(K^{ij}_k\right)^T. \tag{94}$$

- **Calculation of the merging probabilities:** At this step, the merging probabilities $\{\mu^{ij}_{k|k-1}\}_{i=1}^N$ are calculated for each j. These probabilities are used to merge the estimates $\{\hat{x}^{ij}_{k|k}\}_{i=1}^N$ and covariances $\{P^{ij}_{k|k}\}_{i=1}^N$ for each j. The probabilities are calculated as follows:

$$\mu^{ij}_{k|k-1} = \frac{1}{c_j} \Delta^{ij}_k \pi_{ij} \mu^i_{k-1}, \tag{95}$$

where

$$c_j = \sum_{i=1}^N \Delta^{ij}_k \pi_{ij} \mu^i_{k-1}, \tag{96}$$

$$\Delta^{ij}_k = \mathcal{N}\left(y_k;\ C_j \hat{x}^{ij}_{k|k-1}, S^{ij}_k\right). \tag{97}$$

- **Merging:** N filtered estimates $\{\hat{x}^{ij}_{k|k}\}_{i=1}^N$ and covariances $\{P^{ij}_{k|k}\}_{i=1}^N$ are merged for each j and the new filtered estimates $\{\hat{x}^j_{k|k}\}_{j=1}^N$ and covariances $\{P^j_{k|k}\}_{i=1}^N$ are obtained. The merging is done as follows:

$$\hat{x}^j_{k|k} = \sum_{i=1}^N \mu^{ij}_{k-1|k} \hat{x}^{ij}_{k|k}, \tag{98}$$

$$P^j_{k|k} = \sum_{i=1}^N \mu^{ij}_{k-1|k} \left[P^{ij}_{k|k} + \left(\hat{x}^{ij}_{k|k} - \hat{x}^j_{k|k}\right)\left(\hat{x}^{ij}_{k|k} - \hat{x}^j_{k|k}\right)^T\right]. \tag{99}$$

- **Mode probability update:** The previous mode probabilities $\{\mu^i_{k-1}\}_{i=1}^N$ are updated to obtain the new mode probabilities $\{\mu^j_k\}_{j=1}^N$ as follows:

$$\mu^j_k = \frac{1}{c} \sum_{i=1}^N \Delta^{ij}_k \pi_{ij} \mu^i_{k-1}, \tag{100}$$

where

$$c = \sum_{j=1}^N \sum_{i=1}^N \Delta^{ij}_k \pi_{ij} \mu^i_{k-1}. \tag{101}$$

- **Output estimate and covariance calculation by merging:** The output estimate and covariance are calculated by merging the filtered estimates $\{\hat{x}^j_{k|k}\}_{j=1}^N$ and covariances $\{P^j_{k|k}\}_{j=1}^N$. The merging is done using the

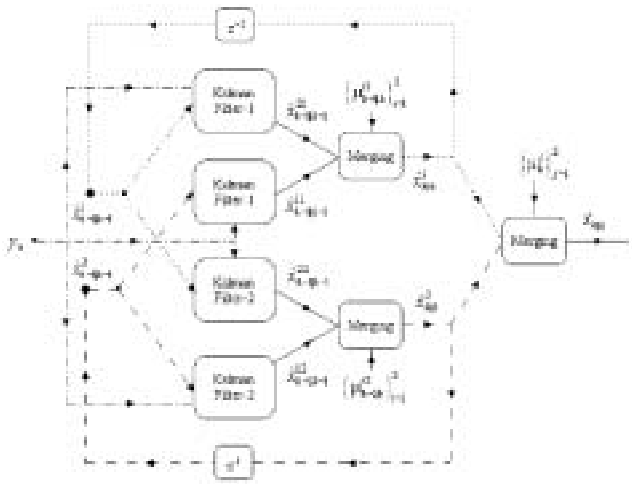

Figure 2. Block diagram of a two-model GPB2 algorithm.

updated mode probabilities $\{\mu_k^j\}_{j=1}^N$ as follows:

$$\hat{x}_{k|k} = \sum_{j=1}^{N} \mu_k^j \hat{x}_{k|k}^j, \tag{102}$$

$$P_{k|k} = \sum_{j=1}^{N} \mu_k^j \left[P_{k|k}^j + \left(\hat{x}_{k|k}^j - \hat{x}_{k|k}\right)\left(\hat{x}_{k|k}^j - \hat{x}_{k|k}\right)^T \right]. \tag{103}$$

The steps of the GPB2 algorithm related with the filtered state estimates are summarized for a two-model GPB2 algorithm in figure 2.

5.2 *Steps of the IMM algorithm*

At each time step, an N-model IMM algorithm keeps N filtered estimates $\{\hat{x}_{k|k}^j\}_{j=1}^N$, covariances $\{P_{k|k}^j\}_{j=1}^N$ and mode probabilities $\{\mu_k^j\}_{j=1}^N$. At each cycle, given the previously kept filtered estimates $\{\hat{x}_{k-1|k-1}^j\}_{j=1}^N$, covariances $\{P_{k-1|k-1}^j\}_{j=1}^N$ and the mode probabilities $\{\mu_{k-1}^i\}_{j=1}^N$, the algorithm calculates the updated values of these quantities using the following steps:

- **Calculation of mixing probabilities:** In contrast to GPB2, IMM algorithm merges the previous filtered estimates $\{\hat{x}_{k-1|k-1}^j\}_{j=1}^N$ and covariances $\{P_{k-1|k-1}^j\}_{j=1}^N$ to obtain N-different initial estimates $\{\hat{x}_{k-1|k-1}^{0j}\}_{j=1}^N$ and the covariances $\{P_{k-1|k-1}^{0j}\}_{j=1}^N$ the mode-matched Kalman filters. This process is called as mixing and the merging probabilities used for this purpose are called as the mixing probabilities. The mixing probabilities are calculated as follows:

$$\mu_{k-1|k-1}^{ij} = \frac{1}{\bar{c}_j} \pi_{ij} \mu_{k-1}^i, \tag{104}$$

where

$$\bar{c}_j = \sum_{i=1}^{N} \pi_{ij} \mu_{k-1}^i. \tag{105}$$

- **Mixing:** The previous filtered estimates $\{\hat{x}_{k-1|k-1}^j\}_{j=1}^N$ and covariances $\{P_{k-1|k-1}^j\}_{j=1}^N$ are merged to obtain N-different initial estimates $\{\hat{x}_{k-1|k-1}^{0j}\}_{j=1}^N$ and the covariances $\{P_{k-1|k-1}^{0j}\}_{j=1}^N$ as follows:

$$\hat{x}_{k-1|k-1}^{0j} = \sum_{i=1}^{N} \mu_{k-1|k-1}^{ij} \hat{x}_{k-1|k-1}^i, \tag{106}$$

$$P_{k-1|k-1}^{0j} = \sum_{i=1}^{N} \mu_{k-1|k-1}^{ij} \left[P_{k-1|k-1}^i + \left(\hat{x}_{k-1|k-1}^i - \hat{x}_{k-1|k-1}^{0j}\right) \times \left(\hat{x}_{k-1|k-1}^i - \hat{x}_{k-1|k-1}^{0j}\right)^T \right]. \tag{107}$$

Note that the mixing process described above represents the moment-matched single Gaussian approximation.

- **Mode-matched Kalman filtering:** The algorithm takes each initial estimate $\hat{x}_{k-1|k-1}^{0j}$ and covariance $P_{k-1|k-1}^{0j}$ and executes N Kalman filters each matched to a different model. All N Kalman filters use its corresponding initial estimate $\hat{x}_{k-1|k-1}^{0j}$ and covariance $P_{k-1|k-1}^{0j}$ as their initial state and covariance. Since there are N initial conditions, N Kalman filters are executed. At the end of this filtering the new filtered estimates $\hat{x}_{k|k}^j$ and covariances $\hat{x}_{k|k}^j$ are obtained. The filtering equations are given as follows:

$$\hat{x}_{k|k-1}^j = A_j \hat{x}_{k-1|k-1}^{0j}, \tag{108}$$

$$P_{k|k-1}^j = A_j P_{k-1|k-1}^{0j} A_j^T + B_j Q B_j^T, \tag{109}$$

$$S_k^j = C_j P_{k|k-1}^j C_j^T + D_j R D_j^T, \tag{110}$$

$$K_k^j = P_{k|k-1}^j C_j^T \left(S_k^j\right)^{-1}, \tag{111}$$

$$\hat{x}_{k|k}^j = \hat{x}_{k|k-1}^j + K_k^j\left(y_k - C_j \hat{x}_{k|k-1}^j\right), \tag{112}$$

$$P_{k|k}^j = P_{k|k-1}^j - K_k^j S_k^j \left(K_k^j\right)^T. \tag{113}$$

- **Mode probability update:** The previous mode probabilities $\{\mu_{k-1}^i\}_{i=1}^N$ are updated to obtain the new mode probabilities $\{\mu_k^j\}_{i=1}^N$ as follows:

$$\mu_k^j = \frac{1}{c} \Delta_k^j \sum_{i=1}^{N} \pi_{ij} \mu_{k-1}^i \tag{114}$$

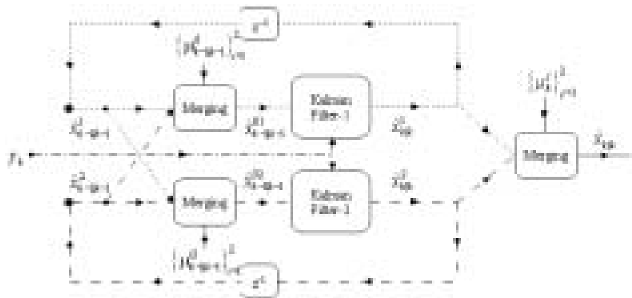

Figure 3. Block diagram of a two-model IMM algorithm.

where

$$\bar{c} = \sum_{j=1}^{N} \Delta_k^j \sum_{i=1}^{N} \pi_{ij}\mu_{k-1}^i, \tag{115}$$

$$\Delta_k^j = \mathcal{N}\left(y_k; C_j\hat{x}_{k|k-1}^j, S_k^j\right). \tag{116}$$

- **Output estimate and covariance calculation by merging:** The output estimate and covariance are calculated by merging the filtered estimates $\{\hat{x}_{k|k}^j\}_{j=1}^N$ and covariances $\{P_{k|k}^j\}_{j=1}^N$. The merging is done using the updated mode probabilities $\{\mu_k^j\}_{j=1}^N$ as follows:

$$\hat{x}_{k|k} = \sum_{j=1}^{N} \mu_k^j \hat{x}_{k|k}^j, \tag{117}$$

$$P_{k|k} = \sum_{j=1}^{N} \mu_k^j \left[P_{k|k}^j + \left(\hat{x}_{k|k}^j - \hat{x}_{k|k}\right)\left(\hat{x}_{k|k}^j - \hat{x}_{k|k}\right)^T \right]. \tag{118}$$

The steps of the IMM algorithm related with the filtered state estimates are summarized for a two-model IMM algorithm in figure 3.

6. Efficient mixed IMM-GPB2 algorithm

As observed in the discussion above, the moment-matched single Gaussian approximation is made at many steps of both IMM and GPB2 algorithms. In this section, we are going to concentrate specifically on the ones made in the mixing step of the IMM algorithm because they are the main approximations differentiating an IMM filter from a GPB2 filter. In GPB2 algorithm, for each model j, each input estimate $\hat{x}_{k-1|k-1}^i$ is passed through the Kalman filter matched to model j and the resulting estimates are merged to form the filtered estimate corresponding to that model (i.e., $\hat{x}_{k|k}^j$). In contrast to this, in the IMM filter, the input estimates $\{\hat{x}_{k-1|k-1}^i\}_{i=1}^N$ are mixed (merged) first to form a single initial estimate, and this initial estimate is then input to the Kalman filter matched to the model j to obtain filtered estimate corresponding to that model (i.e., $\hat{x}_{k|k}^j$). This initial mixing process in the IMM algorithm is the main reason for the increase in the errors of the IMM filter relative to the GPB2 filter. Comparing these processes with our error analysis, we see that our covariance formula calculated in § 4 can be used to calculate the covariance of the error between the filtered estimates that would be obtained using IMM-type merging (first mixing then filtering) and those that would be obtained by GPB2-type merging (first filtering then merging). Therefore, at the beginning of each algorithm cycle, we can decide, by observing the error covariance, whether to make IMM-type or GPB2-type merging for each model. For the models which would result in large errors with the IMM-type merging, GPB2-type merging can be selected to decrease the errors. This idea is the main motivation for our mixed IMM-GPB2 filter.

In the following, we are going to present the steps of our mixed IMM-GPB2 algorithm.

6.1 *Steps of the mixed IMM-GPB2 algorithm*

At each time step, an N-model mixed IMM-GPB2 algorithm keeps N filtered estimates $\{\hat{x}_{k|k}^j\}_{j=1}^N$, covariances $\{P_{k|k}^j\}_{j=1}^N$ and mode probabilities $\{\mu_k^j\}_{j=1}^N$. At the beginning of each cycle, the algorithm calculates a statistics γ_k^j for each model to decide whether to make IMM-type or GPB-type merging for that model. Given the previously kept filtered estimates $\{\hat{x}_{k-1|k-1}^j\}_{j=1}^N$, covariances $\{P_{k-1|k-1}^j\}_{j=1}^N$ and the mode probabilities $\{\mu_{k-1}^j\}_{j=1}^N$, the algorithm calculates the updated values of these quantities using the following steps.

- **Calculation of mixing probabilities:** The mixing probabilities are required for IMM-type merging (if any) and for the calculation of the statistics γ_k^j. The mixing probabilities are calculated as follows (same as the ones in IMM):

$$\mu_{k-1|k-1}^{ij} = \frac{1}{\bar{c}_j}\pi_{ij}\mu_{k-1}^i, \tag{119}$$

where

$$\bar{c}_j = \sum_{i=1}^{N} \pi_{ij}\mu_{k-1}^i. \tag{120}$$

- **Calculation of predicted estimates and merged predicted estimates:** This step is required only for the calculation of the statistics. The computational load in this step can be alleviated using the simplifications mentioned in § 6.2. The predicted estimates and covariances are calculated as follows:

$$\hat{x}^{ij}_{k|k-1} = A_j \hat{x}^{i}_{k-1|k-1}, \tag{121}$$

$$P^{ij}_{k|k-1} = A_j P^{i}_{k-1|k-1} A_j^T + B_j Q B_j^T, \tag{122}$$

$$S^{ij}_{k} = C_j P^{ij}_{k|k-1} C_j^T + D_j R D_j^T. \tag{123}$$

The merged predicted estimates and covariances for each model are calculated as follows:

$$\hat{x}^{j}_{app} = \sum_{i=1}^{N} \mu^{ij}_{k-1|k-1} \hat{x}^{ij}_{k|k-1}, \tag{124}$$

$$P^{j}_{app} = \sum_{i=1}^{N} \mu^{ij}_{k-1|k-1} \left[P^{ij}_{k|k-1} + \left(\hat{x}^{ij}_{k|k-1} - \hat{x}^{j}_{app} \right) \times \left(\hat{x}^{ij}_{k|k-1} - \hat{x}^{j}_{app} \right)^T \right], \tag{125}$$

$$S^{j}_{sub} = C_j P^{j}_{app} C_j^T + D_j R D_j^T. \tag{126}$$

- **Error covariance calculation:** At this step, the algorithm calculates for each model j, the covariance of the error Σ_j that would be induced in the filtered estimate $\hat{x}^{j}_{k|k}$ if one uses IMM-type merging instead of GPB2-type merging. This covariance is given using our analysis and calculated error covariance formula in § 4 as follows:

$$\sum\nolimits_j = \sum_{i=1}^{N} \mu^{ij}_{k-1|k-1} P^{ij}_{k|k-1} C_j^T \left(S^{ij}_k \right)^{-1} C_j P^{ij}_{k|k-1} - P^{j}_{app} C_j^T \left(S^{j}_{sub} \right)^{-1} C_j \left(P^{j}_{app} \right)^T + \sum_{i=1}^{N} \mu^{ij}_{k-1|k-1} \hat{x}^{ij}_{k|k-1} \left(\hat{x}^{ij}_{k|k-1} \right)^T - \hat{x}^{j}_{app} \left(\hat{x}^{j}_{app} \right)^T. \tag{127}$$

At this stage, the algorithm has to determine, for each model j, whether the matrix Σ_j is "big" enough to switch to GPB-type merging for that model. This can be done using a matrix norm to obtain a statistics γ^j_k and a simple thresholding. Also, the application specific heuristics can do well for this purpose. In fact, in § 7, we are going to use such a heuristics for the simulation. From this point on, we are going to assume that the required test has already been completed and a decision on whether to use IMM-type or GPB2-type merging has already been made for every model.

- **For every model j (If IMM-type merging is used):**

  (i) **Mixing.** The previous filtered estimates $\{\hat{x}^{j}_{k-1|k-1}\}_{j=1}^{N}$ and covariances $\{P^{j}_{k-1|k-1}\}_{j=1}^{N}$ are merged to obtain N-different initial estimates $\{\hat{x}^{0j}_{k-1|k-1}\}_{j=1}^{N}$ and the covariances $\{P^{0j}_{k-1|k-1}\}_{j=1}^{N}$ as follows:

$$\hat{x}^{0j}_{k-1|k-1} = \sum_{i=1}^{N} \mu^{ij}_{k-1|k-1} \hat{x}^{i}_{k-1|k-1}, \tag{128}$$

$$P^{0j}_{k-1|k-1} = \sum_{i=1}^{N} \mu^{ij}_{k-1|k-1} \left[P^{i}_{k-1|k-1} + \left(\hat{x}^{i}_{k-1|k-1} - \hat{x}^{0j}_{k-1|k-1} \right) \times \left(\hat{x}^{i}_{k-1|k-1} - \hat{x}^{0j}_{k-1|k-1} \right)^T \right]. \tag{129}$$

  (ii) **Mode-matched Kalman filtering.** The algorithm takes each initial estimate $\hat{x}^{0j}_{k-1|k-1}$ and covariance $P^{0j}_{k-1|k-1}$ and executes N Kalman filters each matched to a different model. All N Kalman filters uses its corresponding initial estimate $\hat{x}^{0j}_{k-1|k-1}$ and covariance $P^{0j}_{k-1|k-1}$ as their initial state and covariance. Since there are N initial conditions, N Kalman filters are executed. At the end of this filtering the new filtered estimates $\hat{x}^{j}_{k|k}$ and covariances $\hat{x}^{j}_{k|k}$ are obtained. The filtering equations are given as follows:

$$\hat{x}^{j}_{k|k-1} = A_j \hat{x}^{0j}_{k-1|k-1}, \tag{130}$$

$$P^{j}_{k|k-1} = A_j P^{0j}_{k-1|k-1} A_j^T + B_j Q B_j^T, \tag{131}$$

$$S^{j}_{k} = C_j P^{j}_{k|k-1} C_j^T + D_j R D_j^T, \tag{132}$$

$$K^{j}_{k} = P^{j}_{k|k-1} C_j^T \left(S^{j}_{k} \right)^{-1}, \tag{133}$$

$$\hat{x}^{j}_{k|k} = \hat{x}^{j}_{k|k-1} + K^{j}_{k} \left(y_k - C_j \hat{x}^{j}_{k|k-1} \right), \tag{134}$$

$$P^{j}_{k|k} = P^{j}_{k|k-1} - K^{j}_{k} S^{j}_{k} \left(K^{j}_{k} \right)^T. \tag{135}$$

  (iii) **Likelihood calculation.** The likelihood Δ^j_k of the current measurement y_k is calculated as follows:

$$\Delta^j_k = \mathcal{N}\left(y_k; C_j \hat{x}^{j}_{k|k-1}, S^{j}_{k} \right). \tag{136}$$

- **If GPB2-type merging is used:**

  (i) **Kalman filter measurement updates.** Since the Kalman filter prediction updates have been already done for statistics calculations, in this step, only measurement updates are required. These updates are done as follows:

$$K^{ij}_{k} = P^{ij}_{k|k-1} C_j^T \left(S^{ij}_{k} \right)^{-1}, \tag{137}$$

$$\hat{x}^{ij}_{k|k} = \hat{x}^{ij}_{k|k-1} + K^{ij}_{k} \left(y_k - C_j \hat{x}^{ij}_{k|k-1} \right), \tag{138}$$

$$P^{ij}_{k|k} = P^{ij}_{k|k-1} - K^{ij}_{k} S^{ij}_{k} \left(K^{ij}_{k} \right)^T. \tag{139}$$

(ii) **Calculation of the merging probabilities.** At this step, the merging probabilities $\{\mu^{ij}_{k|k-1}\}^N_{i=1}$ are calculated for each j. These probabilities are used to merge the estimates $\{\hat{x}^{ij}_{k|k}\}^N_{i=1}$ and covariances $\{P^{ij}_{k|k}\}^N_{i=1}$ for each j. The probabilities are calculated as follows:

$$\mu^{ij}_{k|k-1} = \frac{1}{c_j}\Delta^{ij}_k \pi_{ij}\mu^i_{k-1}, \tag{140}$$

where

$$c_j = \sum_{i=1}^{N} \Delta^{ij}_k \pi_{ij}\mu^i_{k-1}, \tag{141}$$

$$\Delta^{ij}_k = \mathcal{N}\left(y_k;\, C_j\hat{x}^{ij}_{k|k-1}, S^{ij}_k\right). \tag{142}$$

(iii) **Merging.** N filtered estimates $\{\hat{x}^{ij}_{k|k}\}^N_{i=1}$ and covariances $\{P^{ij}_{k|k}\}^N_{i=1}$ are merged for each j and the new filtered estimates $\{\hat{x}^{j}_{k|k}\}^N_{j=1}$ and covariances $\{P^{j}_{k|k}\}^N_{j=1}$ are obtained. The merging is done as follows:

$$\hat{x}^j_{k|k} = \sum_{i=1}^{N} \mu^{ij}_{k-1|k}\hat{x}^{ij}_{k|k}, \tag{143}$$

$$P^j_{k|k} = \sum_{i=1}^{N} \mu^{ij}_{k-1|k}\left[P^{ij}_{k|k} + \left(\hat{x}^{ij}_{k|k} - \hat{x}^j_{k|k}\right)\left(\hat{x}^{ij}_{k|k} - \hat{x}^j_{k|k}\right)^T\right]. \tag{144}$$

(iv) **Likelihood calculation.** The likelihood Δ^j_k of the current measurement y_k is calculated as follows:

$$\Delta^j_k = \sum_{i=1}^{N} \Delta^{ij}_k \mu^{ij}_{k-1|k-1}. \tag{145}$$

- **Mode probability update:** The previous mode probabilities $\{\mu^i_{k-1}\}^N_{i=1}$ are updated to obtain the new mode probabilities $\{\mu^j_k\}^N_{j=1}$ as follows:

$$\mu^j_k = \frac{1}{c}\Delta^j_k \sum_{i=1}^{N} \pi_{ij}\mu^i_{k-1}, \tag{146}$$

where

$$\bar{c} = \sum_{j=1}^{N} \Delta^j_k \sum_{i=1}^{N} \pi_{ij}\mu^i_{k-1}, \tag{147}$$

$$\Delta^j_k = \mathcal{N}\left(y_k;\, C_j\hat{x}^j_{k|k-1}, S^j_k\right). \tag{148}$$

- **Output estimate and covariance calculation by merging:** The output estimate and covariance are calculated by merging the filtered estimates $\{\hat{x}^j_{k|k}\}^N_{i=1}$ and covariances $\{P^j_{k|k}\}^N_{i=1}$. The merging is done using the updated mode probabilities $\{\mu^j_k\}^N_{j=1}$ as follows:

$$\hat{x}_{k|k} = \sum_{j=1}^{N} \mu^j_k \hat{x}^j_{k|k}, \tag{149}$$

$$P_{k|k} = \sum_{j=1}^{N} \mu^j_k\left[P^j_{k|k} + \left(\hat{x}^j_{k|k} - \hat{x}_{k|k}\right)\left(\hat{x}^j_{k|k} - \hat{x}_{k|k}\right)^T\right]. \tag{150}$$

6.2 *Possible simplifications*

Note that the calculations required for the statistics calculations are quite complicated. The following are some simplification suggestions that can be applied without reducing the performance of the algorithm substantially.

(i) The predicted quantities $\hat{x}^{ij}_{k|k-1}$, $P^{ij}_{k|k-1}$, S^{ij}_k can be replaced with their equivalents in the previous sampling period i.e., with $\hat{x}^{ij}_{k-1|k-2}$, $P^{ij}_{k-1|k-2}$, S^{ij}_{k-1} respectively. In this way, some or all of them might have been calculated already during the Kalman filtering. Also, in that case, the multiplications in the form $P^{ij}_{k-1|k}C^T_j(S^{ij}_{k-1})^{-1}$ might have been calculated during Kalman gain calculations. Obviously, if this simplification is made, then, while using the GPB2-type merging, one has to calculate the required predictions as well.

(ii) The calculation of the statistics can be done only at every $M \in \{2, 3, \ldots\}$ sampling periods and between the periods, the results of the last statistics calculation can be used.

(iii) Note that the statistics calculations need not require all the elements of the matrices Σ_j. For most of the case, for example, one might try to use only the diagonal elements of the matrices Σ_j for statistics calculation. In that case, the computations required for the calculation of the matrices Σ_j can be reduced significantly.

7. Simulation results

In this section, the performance of the mixed IMM-GPB2 algorithm will be observed and compared to those of the IMM and GPB2 algorithms. For this purpose, we consider a simplified example of a moving target whose acceleration evolves according to a finite-state Markov chain. This example is a slightly modified version of the one given in Jilkov and Li (2004). The target dynamics in one-dimension is given as

$$\underbrace{\begin{bmatrix} p_k \\ v_k \end{bmatrix}}_{x_k} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}\begin{bmatrix} p_{k-1} \\ v_{k-1} \end{bmatrix} + \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix}[a_k + w_k] \tag{151}$$

where p_k, v_k and a_k denote the target position, velocity and acceleration respectively. The initial state x_0 is normally distributed with mean $\bar{x}_0$ and covariance P_0 which are given as

$$\bar{x}_0 = \begin{bmatrix} 80000 \\ 400 \end{bmatrix} \quad P_0 = \begin{bmatrix} 10000 & 1000 \\ 1000 & 10000 \end{bmatrix}. \tag{152}$$

The acceleration process a_k is a finite-state Markov chain with states in the set $\{0, 20, -20\}$. The initial probability distribution for the states is given as $\pi_0 = [0.8, 0.1, 0.1]$. The transition probability matrix for the finite-state Markov chain is

$$\Pi = \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{bmatrix} \tag{153}$$

which corresponds to a highly maneuvering target. The white process noise $w_k \sim \mathcal{N}(w_k; 0, 2^2)$ represents small acceleration changes. It is assumed that only the positions are measured, i.e.,

$$y_k = p_k + v_k \tag{154}$$

where the terms $v_k \sim \mathcal{N}(w_k; 0, 100^2)$ stands for the normally distributed white measurement noise. The sampling period T is taken to be 10 secs.

Using the measurements coming from this system, we execute IMM, GPB2 and our mixed IMM-GPB2 algorithm. The mixed IMM-GPB2 algorithm is implemented using the velocity error standard deviation as the merging decision statistics. This means that we only use the elements corresponding to second row and second column of the 2×2 matrices Σ_j for the statistics calculation. Realizing also that the matrices $C_j = [1 \ 0]$ for $j = 1, 2, 3$, this reduces the computations required for the statistics calculations to an almost negligible level. We have taken the statistics threshold γ_{thresh} as 3 (m/sec) which means that if the statistics γ_k^j for model j is smaller than 3, the IMM-type merging will be used for that model. Otherwise, GPB-type merging is applied. RMS position errors resulting from 1000 Monte-Carlo runs are presented in figure 4. The corresponding velocity errors are shown in figure 5. As observed from the figures, the GPB2 shows the best performance in both of the cases as expected. The errors of the Mixed IMM-GPB2 algorithm are below those of the IMM filter and are very near to those of the GPB2 algorithm. The IMM and GPB2 algorithms uses 3 and 9 Kalman filters for each measurement respectively. The mixed IMM-GPB2 algorithm proves to use an average of 5.2 Kalman filters per each measurement. This shows that, to reach the performance of the GPB2 algorithm, one can use much less Kalman filters compared to the number of Kalman filters required for the GPB2 algorithm. The average

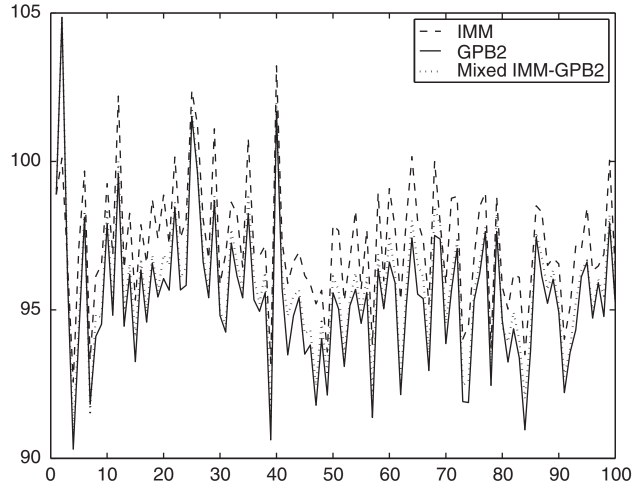

Figure 4. RMS position errors of the IMM, GPB2 and mixed IMM-GPB2 algorithms.

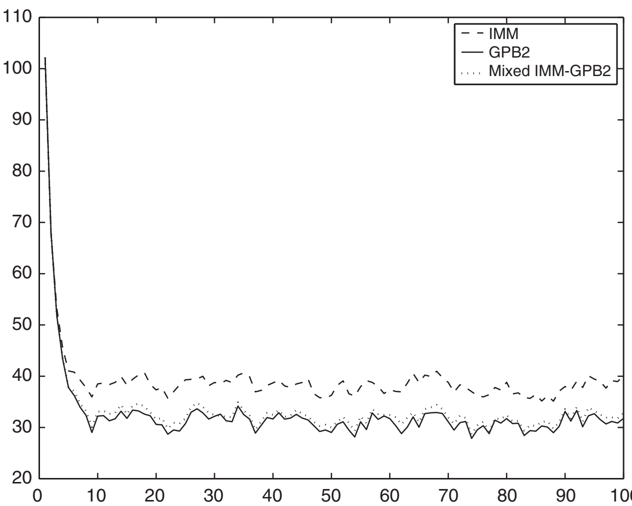

Figure 5. RMS velocity errors of the IMM, GPB2 and mixed IMM-GPB2 algorithms.

position RMS errors for the IMM filter and the GPB2 filter per measurement are 97.20 and 95.38 respectively. The corresponding RMS error value for the mixed IMM-GPB2 filter is 95.83. The average velocity RMS errors for the IMM and GPB2 algorithms per measurement are given as 39.33 and 32.69 respectively. The corresponding RMS error for the mixed IMM-GPB2 algorithm is 33.56 which is a significant reduction in the error relative to the IMM filter towards the performance of GPB2 algorithm. The results obtained using different values of the statistics threshold γ_{thresh} are shown in table 1. The error characteristics of the mixed IMM-GPB2 algorithm seem to reach a virtually indistinguishable level from that of GPB2 with only 7.6 Kalman filters on the average per measurement.

Table 1. The results obtained with different γ_{thresh} values.

γ_{thresh}	Average number of Kalman filters used per measurement	Average RMS position error per measurement	Average RMS velocity error per measurement
0.5	7.6	95.39	32.69
1	6.9	95.40	32.71
3	5.2	95.83	33.56
4	4.7	96.15	35.48
5	4.2	96.56	36.90
7	3	97.20	39.33
GPB2	9	95.38	32.69
IMM	3	97.20	39.33

8. Conclusions and future work

In this work, the error caused by the moment-matched single Gaussian approximation of the Gaussian mixtures applied at the input of the optimal Bayesian filter is examined analytically. The prediction update of the filter is shown to be invariant (in the MMSE sense) under the approximation. The measurement update, on the other hand, is shown to cause some difference (i.e., error) between the optimal and the approximated (sub-optimal) estimates. The resulting error proves to be of zero mean. An analytical formula to calculate its covariance approximately is found.

The calculated covariance is then used to obtain a mixed IMM-GPB2 algorithm which combines the IMM and GPB2 algorithms in which the single Gaussian approximations are abundant. The resulting algorithm turns out to reach the performance of the GPB2 filter with less number of Kalman filters (per measurement) than the GPB2 algorithm.

It is believed that the importance of this study is twofold implying different but possibly overlapping directions for future study.

- **Error covariance formula:** One of the main results of this paper is the error covariance formula given in equation (82). The importance of this result stems from the fact that this is the first attempt in the literature which analytically quantifies the error without resort to Monte Carlo simulations. It must, however, still be admitted that the formula will be biased for many practical cases due to the approximation used in its derivation. Although the simulation section of the paper shows how to obtain good performance using thresholding, the result, in the authors' opinion, can be improved in future studies. In this respect, the reference probability methods (Elliott *et al.* 1995) in which one defines and exploits a newly defined probability measure to calculate expectations very difficult to find otherwise are remarkable alternatives to the classical probabilistic methods used in this work. The form of the expectations, which needed to be approximated in this study, in the reference probability domain could allow better approximation schemes.

- **Mixing idea:** Although being popular in many other areas, to the authors' best knowledge, the idea of mixing better characteristics of different filters is first applied to multiple model filtering in this study. Although only the case of IMM and GPB2 is considered here, the same methodology can be straightforwardly used for combining different order GPB methods as well. In this way, every bit of performance can be extracted from low-cost algorithms and better and better multiple model filters which have a better place on the performance vs. computation curve can be obtained. In this direction, the use of error covariance formula as a decision criterion must be questioned from a practical perspective. Although being an explicit measure of error between different multiple model filters, the computational cost required in obtaining the error covariance could make it infeasible to use in real time applications in spite of the possible simplifications. Introduction of low-cost but still effective decision mechanisms seems to be necessary in this regard.

Appendix A: Derivation of the error covariance

In §4.2, the error covariance formula given in (82) is given without proof for the sake of simplicity. In this appendix, we give a brief derivation of that result.

After defining the quantities T_1, T_2, T_3 and T_4 it was noted in §4.2 that the covariance calculation requires the expected values of the form

$$E_{yk}\left[T_iT_j^T|Y^{k-1}\right] \quad \text{for} \quad i,j=1,\ldots,4 \tag{A1}$$

The approximation required to evaluate these expectations was suggested in §4.2 as either one of the following.

- Assuming that the means $\bar{y}_i$ of the Gaussian components of $p(y_k|Y^{k-1})$ are sufficiently separated.
- Assuming that the measurement dependent posterior probability p_i^+ is approximately unity around the mean $\bar{y}_i$ within the 2σ covariance ellipse.

Using this approximation, it was shown in §4.2 that the following equation is satisfied:

$$E_{y_k}\left[(p_i^+)^2|Y^{k-1}\right]=p_i. \tag{A2}$$

Similarly, using the same approximation, one can easily conclude that $p_i^+p_j^+\approx 0$ for all y_k values when $i\neq j$. Using these two facts, we can write the following result:

$$E_{y_k}\left[p_i^+p_j^k|Y^{k-1}\right]=\begin{cases}p_i, & i=j\\ 0, & i\neq j.\end{cases} \tag{A3}$$

Other basic expectations which can be evaluated similarly are listed below with their results.

- $E_{y_k}\left[p_i^+p_j^+y_k|Y^{k-1}\right]=\begin{cases}p_i\bar{y}_i, & i=j\\ 0, & i\neq j\end{cases},$
- $E_{y_k}\left[p_i^+p_j^+(y_k-\bar{y}_i)(y_k-\bar{y}_j)^T|Y^{k-1}\right]=\begin{cases}p_iS_i, & i=j\\ 0, & i\neq j.\end{cases}$

Note that the property $p_i^+p_j^+\approx 0$ when $i\neq j$ is so useful that it turns the double summations appearing in the expectations like $E_{y_k}[T_3T_3^T|Y^{k-1}]$, $E_{y_k}[T_3T_4^T|Y^{k-1}]$, $E_{y_k}[T_4T_4^T|Y^{k-1}]$, etc. into single summations.

A.1 *Calculation of the terms* $E_{y_k}[T_iT_j^T|Y^{k-1}]$

Using the expectations calculated in §4.2 and the basic expectation results given in the previous part of the appendix, in this section, the expectations $E_{y_k}[T_iT_j^T|Y^{k-1}]$ are evaluated one by one. Only the case $i\leq j$ is investigated since

$$E_{y_k}\left[T_iT_j^T|Y^{k-1}\right]=\left(E_{y_k}\left[T_jT_i^T|Y^{k-1}\right]\right)^T. \tag{A4}$$

- **Calculation of $E_{y_k}[T_1T_1^T|Y^{k-1}]$:**

$$E_{y_k}\left[T_1T_1^T|Y^{k-1}\right]=T_1T_1^T=\bar{x}_{app}\bar{x}_{app}^T. \tag{A5}$$

- **Calculation of $E_{y_k}[T_1T_2^T|Y^{k-1}]$:**

$$E_{y_k}\left[T_1T_2^T|Y^{k-1}\right]=T_1E_{y_k}\left[T_2^T|Y^{k-1}\right]=\bar{x}_{app}0=0. \tag{A6}$$

- **Calculation of $E_{y_k}[T_1T_3^T|Y^{k-1}]$:**

$$\begin{aligned}E_{y_k}\left[T_1T_3^T|Y^{k-1}\right]&=T_1E_{y_k}\left[T_3^T|Y^{k-1}\right]\\&=\bar{x}_{app}\sum_{i=1}^N E_{y_k}\left[p_i^+|Y^{k-1}\right]\bar{x}_i^T\\&=\bar{x}_{app}\sum_{i=1}^N p_i\bar{x}_i^T=\bar{x}_{app}\bar{x}_{app}^T.\end{aligned} \tag{A7}$$

- **Calculation of $E_{y_k}[T_1T_4^T|Y^{k-1}]$:**

$$\begin{aligned}E_{y_k}\left[T_1T_4^T|Y^{k-1}\right]&=T_1E_{y_k}\left[T_4^T|Y^{k-1}\right]\\&=\bar{x}_{app}\sum_{i=1}^N\Big(E_{y_k}\left[p_i^+y_k|Y^{k-1}\right]\\&\quad-E_{y_k}\left[p_i^+|Y^{k-1}\right]\bar{y}_i\Big)K_i^T\\&=\bar{x}_{app}\sum_{i=1}^N(p_i\bar{y}_i-p_i\bar{y}_i)K_i^T=0.\end{aligned} \tag{A8}$$

- **Calculation of $E_{y_k}[T_2T_2^T|Y^{k-1}]$:**

$$\begin{aligned}E_{y_k}\left[T_2T_2^T|Y^{k-1}\right]&=K_{sub}E_{y_k}\left[(y_k-\bar{y})(y_k-\bar{y})^T|Y^{k-1}\right]K_{sub}^T\\&=K_{sub}S_{sub}K_{sub}^T=P_{app}C^TS_{sub}^{-1}CP_{app}.\end{aligned} \tag{A9}$$

- **Calculation of $E_{y_k}[T_2T_3^T|Y^{k-1}]$:**

$$\begin{aligned}E_{y_k}\left[T_2T_3^T|Y^{k-1}\right]&=K_{sub}\sum_{i=1}^N\Bigg(E_{y_k}\left[p_i^+y_k|Y^{k-1}\right]\\&\quad-E_{y_k}\left[p_i^+|Y^{k-1}\right]\bar{y}\Bigg)\bar{x}_i^T\\&=K_{sub}\left[\sum_{i=1}^N p_i\bar{y}_i\bar{x}_i^T-\sum_{i=1}^N p_i\bar{y}\bar{x}_i^T\right]\\&=K_{sub}\sum_{i=1}^N p_i(\bar{y}_i-\bar{y})\bar{x}_i^T.\end{aligned} \tag{A10}$$

- **Calculation of $E_{y_k}[T_2 T_4^T | Y^{k-1}]$:**

$$
\begin{aligned}
E_{y_k}\left[T_2 T_4^T | Y^{k-1}\right] &= K_{sub} \sum_{i=1}^{N} E_{y_k}\left[p_i^+ (y_k - \bar{y})(y_k - \bar{y}_i)^T | Y^{k-1}\right] K_i^T \\
&= K_{sub} \sum_{i=1}^{N} E_{y_k}\left[p_i^+ (y_k - \bar{y}_i)(y_k - \bar{y}_i)^T | Y^{k-1}\right] K_i^T \\
&\quad - K_{sub} \sum_{i=1}^{N} E_{y_k}\left[p_i^+ (\bar{y}_i - \bar{y})(y_k - \bar{y}_i)^T | Y^{k-1}\right] K_i^T \\
&= K_{sub} \sum_{i=1}^{N} p_i S_i K_i^T - K_{sub} \sum_{i=1}^{N} (\bar{y}_i - \bar{y}) E_{y_k} \\
&\quad \times \left[p_i^+ (y_k - \bar{y}_i)^T | Y^{k-1}\right] K_i^T \\
&= K_{sub} \sum_{i=1}^{N} p_i S_i K_i^T = K_{sub} C \sum_{i=1}^{N} p_i \Sigma_i \\
&= K_{sub} C \left[P_{app} - \sum_{i=1}^{N} p_i \left(\bar{x}_i - \bar{x}_{app}\right)\left(\bar{x}_i - \bar{x}_{app}\right)^T \right] \\
&= K_{sub} C \left[P_{app} - \sum_{i=1}^{N} p_i \left(\bar{x}_i - \bar{x}_{app}\right) \bar{x}_i^T \right] \\
&= P_{app} C^T S_{sub}^{-1} C P_{app} - K_{sub} \sum_{i=1}^{N} p_i (\bar{y}_i - \bar{y}) \bar{x}_i^T.
\end{aligned}
$$

- **Calculation of $E_{y_k}[T_3 T_3^T | Y^{k-1}]$:**

$$
E_{y_k}\left[T_3 T_3^T | Y^{k-1}\right] = \sum_{i=1}^{N} E_{y_k}\left[\left(p_i^+\right)^2 | Y^{k-1}\right] \bar{x}_i \bar{x}_i^T = \sum_{i=1}^{N} p_i \bar{x}_i \bar{x}_i^T. \tag{A11}
$$

- **Calculation of $E_{y_k}[T_3 T_4^T | Y^{k-1}]$:**

$$
\begin{aligned}
E_{y_k}\left[T_3 T_4^T | Y^{k-1}\right] &= \sum_{i=1}^{N} \bar{x}_i E_{y_k}\left[\left(p_i^+\right)^2 (y_k - \bar{y}_i) | Y^{k-1}\right] K_i^T \\
&= \sum_{i=1}^{N} \bar{x}_i (p_i \bar{y}_i - p_i \bar{y}_i) K_i^T = 0.
\end{aligned} \tag{A12}
$$

- **Calculation of $E_{y_k}[T_4 T_4^T | Y^{k-1}]$:**

$$
\begin{aligned}
E_{y_k}\left[T_4 T_4^T | Y^{k-1}\right] &= \sum_{i=1}^{N} K_i E_{y_k}\left[\left(p_i^+\right)^2 (y_k - \bar{y}_i)(y_k - \bar{y}_i)^T | Y^{k-1}\right] K_i^T \\
&= \sum_{i=1}^{N} p_i K_i S_i K_i^T = \sum_{i=1}^{N} p_i P_i C^T S_i^{-1} C P_i.
\end{aligned} \tag{A13}
$$

A.2 *Calculation of the covariance* ' "

In this section, in order to evaluate the expectation

$$
\sum_{\Delta} = E_{y_k}\left[(T_1 + T_2 - T_3 - T_4)(T_1 + T_2 - T_3 - T_4)^T | Y^{k-1}\right], \tag{A14}
$$

we combine the results of the previous section as

$$
\begin{aligned}
\sum\nolimits_{\Delta} &= \bar{x}_{app} \bar{x}_{app}^T - \bar{x}_{app} \bar{x}_{app}^T + P_{app} C^T S_{sub}^{-1} C P_{app} \\
&\quad - K_{sub} \sum_{i=1}^{N} p_i (\bar{y}_i - \bar{y}) \bar{x}_i^T - P_{app} C^T S_{sub}^{-1} C P_{app} \\
&\quad + K_{sub} \sum_{i=1}^{N} p_i (\bar{y}_i - \bar{y}) \bar{x}_i^T - \bar{x}_{app} \bar{x}_{app}^T \\
&\quad - \sum_{i=1}^{N} p_i \bar{x}_i (\bar{y}_i - \bar{y})^T K_{sub}^T + \sum_{i=1}^{N} p_i \bar{x}_i x_i^T - P_{app} C^T S_{sub}^{-1} C P_{app} \\
&\quad + \sum_{i=1}^{N} p_i \bar{x}_i (\bar{y}_i - \bar{y})^T K_{sub}^T + \sum_{i=1}^{N} p_i P_i C^T S_i^{-1} C P_i \\
&= \sum_{i=1}^{N} p_i P_i C^T S_i^{-1} C P_i - P_{app} C^T S_{sub}^{-1} C P_{app} \\
&\quad + \sum_{i=1}^{N} p_i \bar{x}_i \bar{x}_i^T - \bar{x}_{app} \bar{x}_{app}^T
\end{aligned} \tag{A15}
$$

which is the same as (82). Note that while writing the right-hand side of (A15), we used the fact in (A4) extensively.

References

G.A. Ackerson and K.S. Fu, "On state estimation in switching environments", *IEEE Trans. Automat. Contr.*, 15, pp. 10–17, 1970.

B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Englewood Cliffs, NJ: Prentice Hall, 1979.

Y. Bar-Shalom and X.R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Norwell, MA: Artech House, 1993.

H.A.P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients", *IEEE Trans. Automat. Contr.*, 33, pp. 780–783, 1988.

Z. Brzeźniak and T. Zastawniak, *Basic Stochastic Processes: A Course Through Exercises*, London: Springer-Verlag, 1999.

R.J. Elliott, L. Aggoun and J.B. Moore, *Hidden Markov Models: Estimation and Control*, New York: Springer, 1995.

A. Gelb, *Applied Optimal Estimation*, Cambridge, MA: MIT Press, 1974.

V.P. Jilkov and X.R. Li, "Online Bayesian estimation of transition probabilities for Markovian jump systems", *IEEE Trans. Signal Processing*, 52, pp. 1620–1630, 2004.

P.R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*, Engle-wood Cliffs, NJ: Prentice Hall, 1986.

H.W. Sorenson and D.L. Alspach, "Recursive Bayesian estimation using Gaussian sums", *Automatica*, 7, pp. 465–479, 1971.

J.K. Tugnait, "Adaptive estimation and identification for discrete systems with Markov jump parameters", *IEEE Trans. Automat. Contr.*, 27, pp. 1054–1065, 1982.