

Article

# Gaussian Mixture Model-Based Ensemble Kalman Filtering for State and Parameter Estimation for a PMMA Process

Ruoxia Li, Vinay Prasad \* and Biao Huang

Department of Chemical and Materials Engineering, University of Alberta, 12th Floor—Donadeo Innovation Centre for Engineering (ICE), 9211—116 Street, Edmonton, AB T6G 1H9, Canada; ruoxia@ualberta.ca (R.L.); bhuang@ualberta.ca (B.H.)

\* Correspondence: vprasad@ualberta.ca; Tel.: +1780-2481-595; Fax: +1780-4922-881

Academic Editor: Masoud Soroush

Received: 24 November 2015; Accepted: 21 March 2016; Published: 30 March 2016

**Abstract:** Polymer processes often contain state variables whose distributions are multimodal; in addition, the models for these processes are often complex and nonlinear with uncertain parameters. This presents a challenge for Kalman-based state estimators such as the ensemble Kalman filter. We develop an estimator based on a Gaussian mixture model (GMM) coupled with the ensemble Kalman filter (EnKF) specifically for estimation with multimodal state distributions. The expectation maximization algorithm is used for clustering in the Gaussian mixture model. The performance of the GMM-based EnKF is compared to that of the EnKF and the particle filter (PF) through simulations of a polymethyl methacrylate process, and it is seen that it clearly outperforms the other estimators both in state and parameter estimation. While the PF is also able to handle nonlinearity and multimodality, its lack of robustness to model-plant mismatch affects its performance significantly.

**Keywords:** Gaussian mixture model; ensemble Kalman filter; particle filter; expectation maximization; polymethyl methacrylate; state and parameter estimation

## 1. Introduction

Polymerization reactors offer unique challenges for process modeling, monitoring, and control. The production of polymers of different grades means that the process conditions are changed relatively often. Product quality specifications (usually expressed in terms of constraints on the properties of the molecular weight distribution) and dynamic operation lead to the need for on-line monitoring and control, which further require accurate process models and real-time estimation of states and parameters of the system. Over the years, the most popular estimator used in nonlinear chemical processes—both in general and specifically for polymerization reactors, too—is the extended Kalman filter (EKF) (e.g., [1–8]). However, this estimator involves linearization of the original model at each step, and can be inaccurate for highly nonlinear systems. Our focus in this work is on particle-based estimators, which are derivative free estimators using different sampling methods to generate an ensemble of particles to represent the distributions of the dynamic states of the system.

The most commonly used estimators based on the use of an ensemble of particles are the ensemble Kalman filter (EnKF) [9], the unscented Kalman filter (UKF) [10,11] and the particle filter (PF) [12]. While the EnKF and the UKF provide only the mean and variance of the posterior distribution of the states (since they use a Gaussian assumption for the distributions), the PF, which works on Bayesian principles, can provide estimates for the full distribution of the states even in situations where the distribution is not Gaussian (which occurs in nonlinear systems) by using a set of particles associated with different weights. In practice, the application of the PF to chemical processes is very recent. Chen *et al.* [13] compared the performance of the auxiliary particle filter with an EKF for a batch

polymethyl methacrylate process to show that it outperformed the EKF in terms of the root mean squared error for state and parameter estimation. Shenoy *et al.* [14] compared the UKF, EKF, and PF in a case study on a polyethylene reactor simulation to demonstrate that the PF provided more accurate estimation results, but was less robust to plant-model mismatch. Shao *et al.* [15] compared the performance of the PF, EKF, UKF, and moving horizon estimation for constrained state estimation and showed that the constrained PF provides more accurate estimation results compared to other methods.

An important issue with the PF relates to its performance for high dimensional systems. The ensemble Kalman filter (EnKF), on the other hand, has the advantage of being scalable to high-dimensional systems without a prohibitive increase in the size of the ensemble required; however, as stated earlier, the algorithm is based on the assumption that both the prior and posterior distribution of the states can be approximated by the Gaussian distribution, and it may be unreliable when this assumption is not valid.

Polymerization processes can be of high dimension when they are described using population balance models [16,17] and a multimodal distribution of properties such as the particle size and molecular weight, may be desirable [18–20]. This, especially in the presence of model-plant mismatch, creates challenges for both the EnKF and the PF. Also, the nonlinearity of the systems may lead to multimodality in the state distributions.

Recently, the Gaussian mixture model (GMM) has been combined with the ensemble Kalman filter to create a new category of estimators: Gaussian mixture filters. Bengtsson *et al.* [21] proposed the GMM to approximate the prior distribution of the states, but the means and variances of the GMM were approximated directly from the ensemble. In [22], Smith proposed the expectation maximization (EM) algorithm to learn the parameters of the prior distribution modeled by the GMM. In the update step, the idea of Kalman-based filtering was extended to the multimodal scenario; however, the posterior distribution is constrained to be a Gaussian distribution. Dovera and Della Rossa [23] used a different update technique and retained the posterior distribution as a GMM.

In this work, we propose an estimator that belongs to the category of Gaussian mixture filters and provides a full state distribution at each time step that is approximated by the GMM. We extend the idea of the EnKF to priors with multimodal features that are described by the GMM. We present results on the application of this estimator to a polymethyl methacrylate (PMMA) process and compare its performance to that of the EnKF and the PF.

## 2. State Estimation Techniques for Nonlinear Systems

Consider a dynamic nonlinear system represented by:

$$\begin{cases} x_n = f(x_{n-1}, u_{n-1}, \theta) + v_n \\ y_n = Hx_n + e_n \end{cases} \quad (1)$$

where  $x_n$  are the hidden states.  $u_n$  and  $y_n$  are the inputs and outputs of the system.  $\theta$  represents the parameters in the model.  $v_n$  and  $e_n$  are process noise and measurement noise respectively.

In this section, we will introduce the particle filter and the ensemble Kalman filter for these systems, and then describe the GMM-based ensemble Kalman filter that we propose to employ. The performance of the three estimators will be compared for the PMMA system in later sections.

### 2.1. Particle Filter (PF)

The PF employs a sequential Monte Carlo method that uses a set of sampling techniques to generate samples from a sequence of probability distribution functions.

The particle filter approximates the posterior probability  $p(x_n|y_n)$  with a set of  $N_s$  particles  $\{x_n^{(i)}\}$ . Each particle is assigned a weight  $w_n^{(i)}$  and the sum of all weights is unity. Since the probability distribution of the states conditioned on the measurements of the outputs,  $p(x_n|y_{1:n})$ , is usually

unknown, these particles are drawn from the importance distribution  $q(x_n | y_{1:n})$ . The posterior distribution is given by:

$$p(x_n | y_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)} \delta(x_n - x_n^{(i)}) \quad (2)$$

where the recursive update of the weights  $w_n^{(i)}$  is given by:

$$w_n^{(i)} = w_{n-1}^{(i)} \frac{p(x_n^{(i)} | x_{n-1}^{(i)})}{q(x_n^{(i)} | x_{n-1}^{(i)}, y_n)} p(y_n | x_n^{(i)}) \quad (3)$$

In the sequential importance resampling (SIR) version of the PF, we choose  $q(x_n^{(i)} | x_{n-1}^{(i)}) = p(x_n^{(i)} | x_{n-1}^{(i)})$ , so that  $w_n^{(i)} = w_{n-1}^{(i)} p(y_n | x_n^{(i)})$ , i.e., we draw particles directly from the prior distribution at time instant  $n$ .

The  $N_s$  particles at time step  $(n-1)$  are forwarded through the state transition equation  $x_{n|n-1}^{(i)} = f(x_{n-1|n-1}^{(i)}, u_{n-1}, v_{n-1}^{(i)})$  to get a new series of particles  $\{x_{n|n-1}^{(i)}\}_{i=1}^{N_s}$  to approximate the prior density  $p(x_n | y_{1:n-1})$  at time instant  $n$ . The weight  $w_n^{(i)}$  associated with each particle is calculated using Equation (3). Then, a resampling step is performed on the prior particles  $\{x_{n|n-1}^{(i)}\}_{i=1}^{N_s}$  based on their weights  $w_n^{(i)}$  to generate the posterior particles  $\{x_{n|n}^{(i)}\}_{i=1}^{N_s}$  such that the weights of all the posterior particles are set to be equal. The full state distribution and its properties can be calculated from the posterior particles.

## 2.2. Ensemble Kalman Filter (EnKF)

The EnKF was first proposed as a data assimilation technique for highly nonlinear ocean models by Evensen [9] and is a Monte Carlo sampling based variant of the Kalman filter. Like the PF, it also uses an ensemble of particles from which the statistical information of the distribution of the states can be calculated, but it uses the Kalman update. In order to have an explicit analytical expression for the Kalman gain, both the prior and posterior distributions are approximated by the Gaussian distribution. The framework of this algorithm is as follows:

At time step  $k$ ,  $N_e$  particles are drawn from the prior distribution to form the prior ensemble  $\{x_{n-1|n-1}^i\}_{i=1, \dots, N_e}$ . In the prediction step, each member of the ensemble  $x_{n-1|n-1}^i$  is forwarded through the state transition equation  $x_{n|n-1}^i = f(x_{n-1|n-1}^i, u_{n-1}, v_{n-1}^i)$  to get its predicted value, thus forming a predicted ensemble  $\{x_{n|n-1}^i\}_{i=1, \dots, N_e}$ . Corresponding to each member of the ensemble, a predicted observation value is obtained; this can be achieved by perturbing the measurement of the output with random measurement error. Let  $\{y_{n|n-1}^i\}_{i=1, \dots, N_e}$  denote the predicted observation data.

In the update step, two error matrices are calculated. The error matrix of the predicted state ensemble is defined as:

$$e_{n|n-1}^i = x_{n|n-1}^i - \mu_{n|n-1}^x \quad (4)$$

where  $\mu_{n|n-1}^x = \frac{1}{N_e} \sum_{i=1}^{N_e} x_{n|n-1}^i$ .

The error matrix of the predicted measurement ensemble is defined as:

$$\varepsilon_{n|n-1}^i = y_{n|n-1}^i - \mu_{n|n-1}^y \quad (5)$$

where  $\mu_{n|n-1}^y = \frac{1}{N_e} \sum_{i=1}^{N_e} y_{n|n-1}^i$ .

The cross-covariance between the state prediction ensemble and measurement ensemble is given in Equation (6), and the covariance matrix of the measurement ensemble is given in Equation (7).

$$P_{n|n-1}^{e,\varepsilon} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (e_{n|n-1}^i)(\varepsilon_{n|n-1}^i)^T \quad (6)$$

$$P_{n|n-1}^{\varepsilon,\varepsilon} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\varepsilon_{n|n-1}^i)(\varepsilon_{n|n-1}^i)^T \quad (7)$$

with the two covariance matrices, the Kalman gain is calculated as:

$$K = P_{n|n-1}^{e,\varepsilon} (P_{n|n-1}^{\varepsilon,\varepsilon} + R)^{-1} \quad (8)$$

where  $R$  is the covariance of the measurement noise.

Each member of the ensemble is updated as:

$$x_{n|n}^i = x_{n|n-1}^i + K(y_n^{obs} - \hat{y}_{n|n-1}^i) \quad (9)$$

where  $y_n^{obs}$  is the true measurement value at time step  $n$ .

### 2.3. Gaussian Mixture Model Based Ensemble Kalman Filter (EnKF-GMM)

#### 2.3.1. Expectation Maximization (EM) for Clustering of the Gaussian Mixture Model

The probability distribution function of a random vector  $x$  following a finite Gaussian mixture distribution is given by:

$$p_X(x) = \sum_{j=1}^M \pi_j \times N(x; \mu_j, P_j) \quad (10)$$

subject to constraints that  $\pi_j \geq 0$  and  $\sum_{j=1}^M \pi_j = 1$ , where  $\pi_j, \mu_j, P_j$  are the prior probability, mean and

covariance of mode  $j$  and  $N(x; \mu_j, P_j) = \frac{1}{(2\pi)^{n/2} |P_j|^{1/2}} e^{-\frac{1}{2}(x-\mu_j)^T P_j^{-1}(x-\mu_j)}$ .

Given a set of data  $\{x_i\}_{i=1,\dots,N}$  randomly generated by a GMM, the expectation maximization (EM) algorithm is used to estimate the parameters of the GMM,  $\theta = \{\pi_1, \dots, \pi_M, \mu_1, \dots, \mu_M, P_1, \dots, P_M\}$  [24]. EM is a variant of maximum likelihood estimation when there exist hidden variables or missing data. In this case, the mode identity of each data point is considered as the missing or hidden variable. Let  $\{(c_i)_j\}$  be a binary indicator vector representing the identity of the component that generates  $x_i$ . Its value is given by:

$$(c_i)_j = \begin{cases} 1, & \text{if data point is generated by component } j \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

In the EM algorithm, an E-step is performed first to compute the  $Q$  function, the expectation of the log likelihood of the complete data set, by computing the probability of each data  $x_i$  belonging to each component  $j$  given the current parameters  $\theta^k$  estimated from the previous iteration. Specifically,  $Q(\theta|\theta^k) = E[L(p(z|\theta))|\{x\}, \theta^k]$ , where  $\{x\}$  is the observed data set;  $\{z\}$  is the complete data set consisting of both observed and missing data,  $\{z\} = \{c_1, x_1, \dots, c_N, x_N\}$ ,  $c_i$  is the membership of each data point, and  $\theta^k$  is the estimate of the last iteration. This becomes

$$Q(\theta|\theta^k) = \sum_{i=1}^N \sum_{j=1}^M p[(c_i)_j|\{x\}, \theta^k] \left( \log \pi_j N(x_i; \mu_j, P_j) \right) \quad (12)$$

$$w_{ij} = p[(c_i)_j|\{x\}, \theta^k] = \frac{\pi_j^k N(x_i; \mu_j^k, P_j^k)}{\sum_{m=1}^M \pi_m^k N(x_i; \mu_m^k, P_m^k)} \quad (13)$$

Next, the M-step is performed to maximize the Q function and calculate the corresponding  $\theta^{k+1}$ .

$$\pi_j^{k+1} = \frac{N^k}{N} \quad (14)$$

$$\mu_j^{k+1} = \frac{1}{N^k} \sum_{i=1}^N w_{ij} x_i \quad (15)$$

$$P_j^{k+1} = \frac{1}{N^k} \sum_{i=1}^N w_{ij} (x_i - \mu_j^{k+1})(x_i - \mu_j^{k+1})^T \quad (16)$$

where  $N^k = \sum_{i=1}^N w_{ij}$ .

The E-step and the M-step are performed iteratively until the estimates converge. During this process, the problem of singularity may arise when one of the components collapses onto one data point. This usually happens due to over-fitting in the maximum likelihood estimation (MLE). To avoid this problem, one approach is to adopt a Bayesian regularization method [25] to replace the MLE with the maximum *a posteriori* (MAP) estimate. Based on this method, the update of the covariance is modified to become

$$P_j^{k+1} = \frac{\sum_{i=1}^N w_{ij} (x_i - \mu_j^{k+1})(x_i - \mu_j^{k+1})^T + \lambda I_d}{N^k + 1} \quad (17)$$

where  $I_d$  is an n-dimensional unit matrix and  $\lambda$  is a regularization constant determined by some validation data [26]. An alternate (*ad hoc*) method to deal with the problem of singularity is to detect when the singularity occurs and reset the means of all components randomly and the covariance to some larger value.

**The pseudo-code for the EM algorithm is provided below.**

---

Algorithm 1: Expectation Maximization algorithm. Inputs are data set  $\{x_i\}_{i=1,\dots,N}$ , component number M and initial values  $\{\theta^0\}$  of  $\{\pi_j\}_{j=1,\dots,M}$ ,  $\{\mu_j\}_{j=1,\dots,M}$ ,  $\{P_j\}_{j=1,\dots,M}$ ,  $\theta^k = \theta^0$ .

---

EM $\{x, M, \{\theta^k\}\}$

// E step

**while**  $\varepsilon \leq 1e-6$

**for**  $i = 1: N$

**for**  $j = 1:M$

$$p[(c_i)_j | x_i, \theta^k] = p(x_i | (c_i)_j, \theta^k) p((c_i)_j | \theta^k) / p(x_i)$$

**end for**

**end for**

// M step

**for**  $j = 1:M$

$$\pi_j^{k+1} = \sum_{i=1}^N p[(c_i)_j | x_i, \theta^k] / N$$

$$\mu_j^{k+1} = \sum_{i=1}^N p[(c_i)_j | x_i, \theta^k] x_i / \sum_{i=1}^N p[(c_i)_j | x_i, \theta^k]$$

$$P_j^{k+1} = \frac{\sum_{i=1}^N p[(c_i)_j | x_i, \theta^k] (x_i - \mu_j^{k+1})(x_i - \mu_j^{k+1})^T + \lambda I_d}{\sum_{i=1}^N p[(c_i)_j | x_i, \theta^k] + 1}$$

**end for**

$$\varepsilon = \mu^{k+1} - \mu^k$$

**end while**

**return**  $\theta^{k+1}$

---

### 2.3.2. EnKF-GMM Algorithm

In this section, a GMM-based EnKF (EnKF-GMM) filter is proposed to obtain estimates of the full state distribution. As with the particle filter, it also uses a set of particles to represent the posterior

probability distribution function (PDF) of the states. The difference is that the PDF is constrained to be a GMM at every time step.

At each time step, the EnKF-GMM has two steps—forecast and update. The forecast step is identical to the EnKF. An ensemble of size  $N$ ,  $\{x_i\}_{i=1,\dots,N}$ , is drawn from the prior distribution of the states and forwarded through the model to obtain a predicted ensemble for the next time step. Then, the EM algorithm is performed on the predicted ensemble to obtain the estimates of the GMM with  $M$  components. Next, the Kalman update is performed based on each component in the GMM to get an ensemble of size  $N \times M$ . Finally, these ensemble members are combined based on their weights and reduced to a size of  $N$ . The details of the algorithmic sequence are as follows:

#### Forecast:

1. The first portion of the forecast step is to determine the number of components  $M$  in the multimodal distribution.  $M$  can be determined using the Bayesian or other information criteria [27,28], or using prior knowledge. For example, in reservoir models, petrophysical properties (such as porosity or permeability) are typically related to geological units (facies), and variables inside the facies are characterized by underlying multimodal distributions which are known beforehand [9]. In our work, this information can be considered as prior knowledge if we know the distribution of the process noise.
2. With the knowledge of the process model and the number of components  $M$ , the prior ensemble  $\{x_i\}_{i=1,\dots,N}$  is propagated through the model to get the predicted values of the ensemble  $\{x_i^f\}_{i=1,\dots,N}$ . These are the realizations of the predicted state space  $x^f$ .

Assuming the predicted state  $x^f$  at the forecast step is a GMM,

$$p(x^f) = \sum_{j=1}^M \tau_j^f p_j(x^f) = \sum_{j=1}^M \tau_j N(x^f; \mu_j^f, P_j^f) \quad (18)$$

The EM algorithm is applied on  $\{x_i^f\}_{i=1,\dots,N}$  to give us the parameters of the prior distribution ( $\tau_j^f$ ,  $\mu_j^f$  and  $P_j^f$ ) of each component  $j$ .

#### Update:

3. For each component  $j$  of the distribution, the Kalman gain matrix for each Gaussian component is computed by utilizing the membership probability matrix  $W$ .

$$P[j]^f H^T = \sum_{i=1}^N w_{i,j} (x_i^f - \mu_j) (Hx_i^f - H\mu_j)^T / n_j \quad (19)$$

$$HP[j]^f H^T = \sum_{i=1}^N w_{i,j} (Hx_i^f - H\mu_j) (Hx_i^f - H\mu_j)^T / n_j \quad (20)$$

$$K[j] = P[j]^f H^T (HP[j]^f H^T + R)^{-1} \quad (21)$$

where  $w_{i,j} = \frac{\pi_j N(x_i; \mu_j, P_j)}{\sum_{m=1}^M \pi_m N(x_i; \mu_m, P_m)}$ ,  $n_j = \sum_{i=1}^N w_{i,j}$ , and  $H$  is the linearized measurement function.

4. In the update step, assuming one Gaussian component  $j$  claims the ownership of all the ensemble members, the Kalman update can be performed for each component member under component  $j$ . This gives us an ensemble size of  $N \times M$ .

$$x_i^{a,j} = x_i^f + K[k] (d - Hx_i^f - e_i) \quad (22)$$

5. The  $N \times M$  ensemble members can be combined to form  $N$  members by using the probability matrix. This gives us the final posterior ensemble  $\{x_i^a\}_{i=1,\dots,N}$ .

$$x_i^a = \sum_{j=1}^M w_{i,j} x_i^{a,j} \quad (23)$$

The mean and covariance of the posterior can be computed as:

$$\mu_j^a = \sum_{i=1}^N w_{i,j} x_i^{a,j} / n_j \quad (24)$$

$$P[j]^a = \sum_{i=1}^N w_{i,j} (x_i^{a,j} - \mu_j^a) (x_i^{a,j} - \mu_j^a)^T / n_j \quad (25)$$

6. The posterior weight of each component of the distribution can be computed based on the observed data  $d$ , which contains the measurements  $y$ .

$$\tau_j^a = p(\mu_j, \Sigma_j, R | d) = \frac{p(d | \mu_j, \Sigma_j, R) n_j}{\sum_{j=1}^M p(d | \mu_j, \Sigma_j, R) n_j} \quad (26)$$

$$p(d | \mu_j, \Sigma_j, R) = \frac{\exp[-\frac{1}{2} (d - H\mu_j)^T (H\Sigma_j H^T + R)^{-1} (d - H\mu_j)]}{\sqrt{(2\pi)^m |H\Sigma_j H^T + R|}} \quad (27)$$

7. The point estimate is given by:

$$x^a = \sum_{j=1}^M \tau_j^a \mu_j^a \quad (28)$$

The pseudo-code for the EnKF-GMM algorithm is provided below.

---

Algorithm 2: EnKF-GMM algorithm. Inputs include the initial distribution of  $x$ , the total number of the particles  $N$ , the components  $M$ , and the time steps  $T$ . Inputs and observations at each time step are  $u_n$  and  $d_n$ .

---

$[\{x_i^a\}_{i=1}^N, \{\mu_j^a, P_j^a, \tau_j^a\}_{j=1}^M] = \text{EnKF-GMM}[\{x_i\}_{i=1}^N, d_i]$

---

**for**  $n = 1:T$

**for**  $I = 1 : N$

    Draw  $x_i^f \sim f(I, u_{n-1}, v_{n-1}^i)$

    Calculate  $y_i = Hx_i^f + v_n^i$

**end for**

  Apply the EM algorithm on  $\{x_i^f\}_{i=1,\dots,N}$  using algorithm 1:

$\{\tau_j^f, \mu_j^f, P_j^f\}_{j=1}^M = EM[\{x_i^f\}_{i=1,\dots,N}, M, \{\theta^k\}]$

**for**  $j = 1 : M$

    Calculate the Kalman gain of each component  $K[j]$  using Equation (21)

**for**  $i = 1 : N$

      Calculate the updated particles for each component  $\{x_i^{a,j}\}_{i=1}^N$  using Equation (22)

**end for**

    Combine  $\{x_i^{a,j}\}_{i=1}^N$  to obtain the posterior particles  $\{x_i^a\}_{i=1}^N$  using Equation (23)

    Calculate the parameters of the posterior distribution  $\mu_j^a, P_j^a, \tau_j^a$  using

Equations (24)–(26).

**end for**

  Calculate the point estimate  $x^a$  using Equation (28)

**end for**

---

While the PF and the EnKF-GMM both can, in principle, account for multimodality, the use of the Gaussian mixture model provides the EnKF-GMM with greater flexibility in capturing a wide variety of distributions under varying levels of model-plant mismatch, as will be shown in the results.

### 3. Results and Discussion

#### 3.1. Mathematical Model of the Methyl Methacrylate (MMA) Polymerization Process

Simulations of a free-radical methyl methacrylate (MMA) polymerization process are used to demonstrate the performance of the estimation method proposed in this paper. The process is assumed to take place in a continuous stirred tank reactor (CSTR), and uses AIBN as the initiator and toluene as the solvent. The mathematical model of this process is described below in Equations (29)–(35), and further details can be found in [29,30]. Parameter values are provided in Table 1. The six states to be estimated include the monomer concentration  $C_M$ , the initiator concentration  $C_I$ , the reactor temperature  $T$ , the moments of the polymer distribution,  $D_0$  and  $D_1$ , and the jacket temperature  $T_j$ . Only the temperatures are measured. The number average molecular weight (NAMW), which is the primary quality variable for the process, is defined as the ratio  $D_1/D_0$ .

$$\frac{dC_M}{dt} = - (k_p + k_{fm}) C_M P_0 + \frac{F(C_{min} - C_M)}{V} \quad (29)$$

$$\frac{dC_I}{dt} = -k_I C_I + \frac{(F_I C_{Iin} - F C_I)}{V} \quad (30)$$

$$\frac{dT}{dt} = \frac{(-\Delta H) k_p C_M P_0}{\rho C_p} - \frac{UA}{\rho C_p V} (T - T_j) + \frac{F(T_{in} - T)}{V} \quad (31)$$

$$\frac{dD_0}{dt} = (0.5k_{tc} + k_{td}) P_0^2 + k_{fm} C_M P_0 - \frac{FD_0}{V} \quad (32)$$

$$\frac{dD_1}{dt} = M_m (k_p + k_{fm}) C_M P_0 - \frac{FD_1}{V} \quad (33)$$

$$\frac{dT_j}{dt} = \frac{F_{cw} (T_{w0} - T_j)}{V_0} + \frac{UA}{\rho_w C_{pw} V_0} (T - T_j) \quad (34)$$

$$P_0 = \sqrt{\frac{2f^* + C_I k_I}{k_{td} + k_{tc}}} \quad (35)$$

**Table 1.** Operational parameters for the methyl methacrylate (MMA) polymerization reactor.

$F = 1.0 \text{ m}^3/\text{h}$	$M_m = 100.12 \text{ kg/kgmol}$
$F_I = 0.0032 \text{ m}^3/\text{h}$	$f^* = 0.58$
$F_{cw} = 0.1588 \text{ m}^3/\text{h}$	$R = 8.314 \text{ kJ/kgmol} \cdot \text{K}$
$C_{min} = 6.4678 \text{ kgmol/m}^3$	$-\Delta H = 57800 \text{ kJ/kgmol}$
$C_{Iin} = 8.0 \text{ kgmol/m}^3$	$E_p = 1.8283 \times 10^4 \text{ kJ/kgmol}$
$T_{in} = 350 \text{ K}$	$E_I = 1.2877 \times 10^5 \text{ kJ/kgmol}$
$T_{w0} = 293.2 \text{ K}$	$E_{fm} = 7.4478 \times 10^4 \text{ kJ/kgmol}$
$U = 720 \text{ kJ/h} \cdot \text{K} \cdot \text{m}^2$	$E_{tc} = 2.9442 \times 10^3 \text{ kJ/kgmol}$
$A = 2.0 \text{ m}^2$	$E_{td} = 2.9442 \times 10^3 \text{ kJ/kgmol}$
$V = 0.1 \text{ m}^3$	$A_p = 1.77 \times 10^9 \text{ m}^3/\text{kgmol} \cdot \text{h}$
$V_0 = 0.02 \text{ m}^3$	$A_I = 3.792 \times 10^{18} \text{ 1/h}$
$\rho = 866 \text{ kg/m}^3$	$A_{fm} = 1.0067 \times 10^{15} \text{ m}^3/\text{kgmol} \cdot \text{h}$
$\rho_w = 1000 \text{ kg/m}^3$	$A_{tc} = 3.8223 \times 10^{10} \text{ m}^3/\text{kgmol} \cdot \text{h}$
$C_p = 2.0 \text{ kJ}/(\text{kg} \cdot \text{K})$	$A_{td} = 3.1457 \times 10^{11} \text{ m}^3/\text{kgmol} \cdot \text{h}$
$C_{pw} = 4.2 \text{ kJ}/(\text{kg} \cdot \text{K})$	

In all the simulations whose results are described in the following sections, the number of particles used for each estimator,  $N$ , is 100. The number of components,  $M$ , is set to 2. The parameters of the bimodal noise in all simulations are  $\mu=[0.1,0.8]$ ,  $P=\text{diag}(0.1,0.1)$  for states  $C_m$ ,  $C_I$  and  $D_0$ ;  $\mu=[8,64]$ ,  $P=\text{diag}(8,8)$  for state  $D_1$ ; and  $\mu=[0.6,4.8]$ ,  $P=\text{diag}(0.6,0.6)$  for states  $T$  and  $T_j$ .

The simulations we perform are introduced here: Case Study 1 provides a comparison of the EnKF-GMM, the PF, and the EnKF for a case with bimodal distributions and insignificant model-plant mismatch. Case Study 2 provides a comparison of the three estimators where the model-plant mismatch is significant. Case Study 3 compares the estimators for state estimation with uncertain parameters, but with the uncertain parameter not being estimated. Case Study 4 considers the same case as Case Study 3, but with combined state and parameter estimation. In Case Study 5, we consider an alternate version of the PF and use the simulation conditions of Case Study 2.

### 3.2. Comparison of State Estimation with the EnKF-GMM, EnKF, and PF (Case Study 1)

In this section, we present the results of applying the EnKF-GMM, EnKF, and PF algorithms on the PMMA process. To illustrate the performance of the estimators in cases where the states have multimodal distributions, bimodal process noise is applied to all the six states. The measurement noise is assumed to be Gaussian. The prior distribution of the state is also assumed to follow a GM distribution which contains two modes.

For Case Study 1, the true initial values of the states are:

$$x_0 = \left[ 5 \frac{\text{kgmol}}{\text{m}^3}, 3 \frac{\text{kgmol}}{\text{m}^3}, 320\text{K}, 0.5 \frac{\text{kgmol}}{\text{m}^3}, 0.5 \frac{\text{kg}}{\text{m}^3}, 300 \text{ K} \right]$$

The dynamics of the simulation describe how the system relaxes to a steady state from this initial condition. For the estimators, the initial particles are drawn from the prior distribution. The tuning parameters for the prior distribution are its mean and covariance. In the first case, a prior distribution with a small amount of bimodal process noise is tested for the three algorithms. The means of the two Gaussian modes of the prior distribution are:

$$\mu_1 = \left[ 4 \frac{\text{kgmol}}{\text{m}^3}, 2 \frac{\text{kgmol}}{\text{m}^3}, 310 \text{ K}, 0.49 \frac{\text{kgmol}}{\text{m}^3}, 0.49 \frac{\text{kg}}{\text{m}^3}, 295 \text{ K} \right];$$

$$\mu_2 = \left[ 6 \frac{\text{kgmol}}{\text{m}^3}, 4 \frac{\text{kgmol}}{\text{m}^3}, 330 \text{ K}, 0.51 \frac{\text{kgmol}}{\text{m}^3}, 0.51 \frac{\text{kg}}{\text{m}^3}, 305 \text{ K} \right]$$

The covariances of the modes of the prior distribution are:

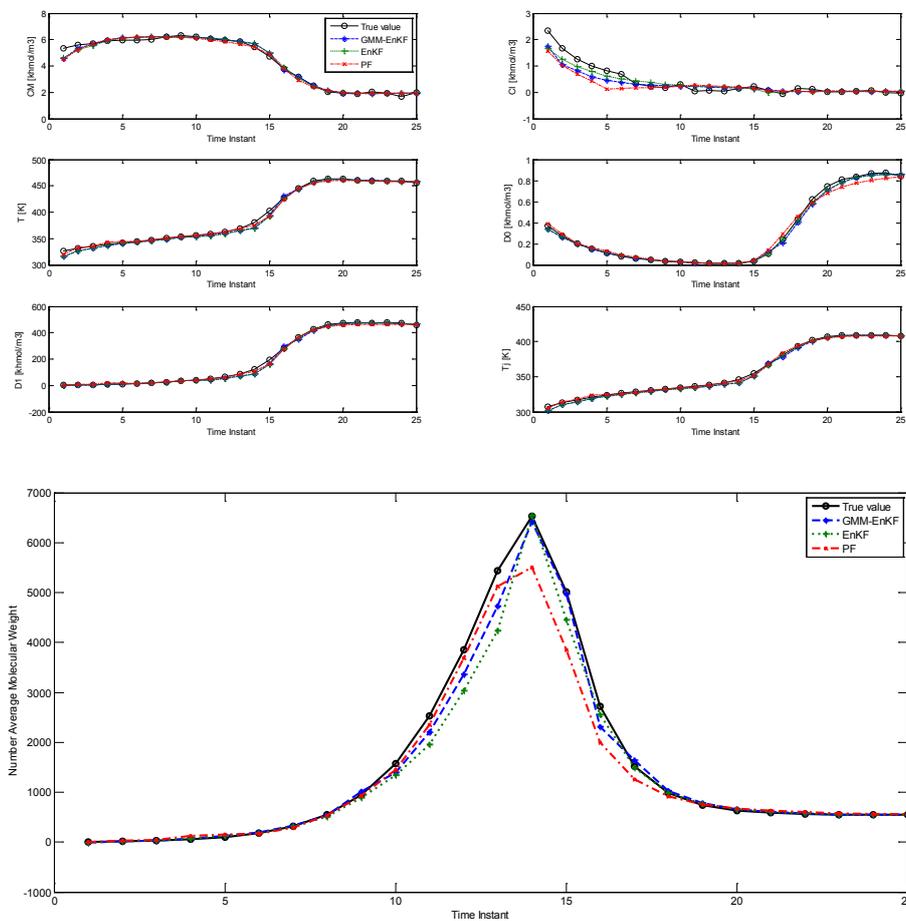
$$P_1 = \text{diag} (4, 4, 28, 8e - 1, 8e - 4, 6);$$

$$P_2 = \text{diag} (4, 4, 28, 8e - 1, 8e - 4, 6)$$

The tuning parameters of the initial distribution indicate a state distribution with insignificant bimodality. The purpose of this simulation is to demonstrate the estimation performance of the three algorithms in the scenario where the state distribution shows insignificant multimodality.

The comparison of estimation results using the EnKF-GMM, EnKF, and PF is shown in Figure 1, with time steps on the  $x$ -axis (each time step is 0.3 h = 18 min). Table 2 shows the root mean squared error (RMSE) over the 25 time steps of the simulation for the six states and the NAMW for the three algorithms. In this case, the estimation results from Figure 1 and Table 2 show that the three algorithms have similar performance in estimation of the six states. However, the EnKF-GMM has the best performance in the estimation of the NAMW. In addition, the converged variance of the estimates of the states, obtained from the estimated covariance matrix with the EnKF-GMM, are  $[10^{-4}, 10^{-4}, 1.2 \times 10^{-4}, 10^{-5}, 2 \times 10^{-4}, 4 \times 10^{-4}]$ , respectively, confirming the significance of the estimates. The PF performs better than the EnKF only for some states. Increasing the number of particles for each of

the algorithms to 200 (results not shown) improves the performance of the PF slightly, but the same conclusions hold.



**Figure 1.** Comparison of the estimation performance of the ensemble Kalman filter (EnKF)-Gaussian mixture model (GMM), EnKF, and particle filter (PF) for the polymethyl methacrylate (PMMA) process with multimodal process noise (Case Study 1).

**Table 2.** RMSE of the Gaussian mixture model based ensemble Kalman filter (EnKF-GMM), ensemble Kalman filter (EnKF), and particle filter (PF) for the polymethyl methacrylate (PMMA) process with multimodal process noise (Case Study 1).

Variable	EnKF-GMM	EnKF	PF
$C_M, \text{kg} \cdot \text{mol}/\text{m}^3$	0.20	0.20	0.33
$C_1, \text{kg} \cdot \text{mol}/\text{m}^3$	0.24	0.20	0.33
$T, \text{K}$	4.3	4.4	3.1
$D_0, \text{kg} \cdot \text{mol}/\text{m}^3$	0.019	0.014	0.032
$D_1, \text{kg}/\text{m}^3$	11.85	11.53	10.44
$T_j, \text{K}$	2.3	2.2	1.4
NAMW	209	338	357

In Case Study 2, the multimodal features of the prior distribution are made more significant compared with the first case. The parameters of the prior distribution given below indicate that both modes lie far away from the true value, which also means that the initial condition mismatch is much larger. The true initial values of the states remain the same as the first case, and the process noise and measurement noise applied to the plant remain unchanged as well. The modified prior distribution is specified by:

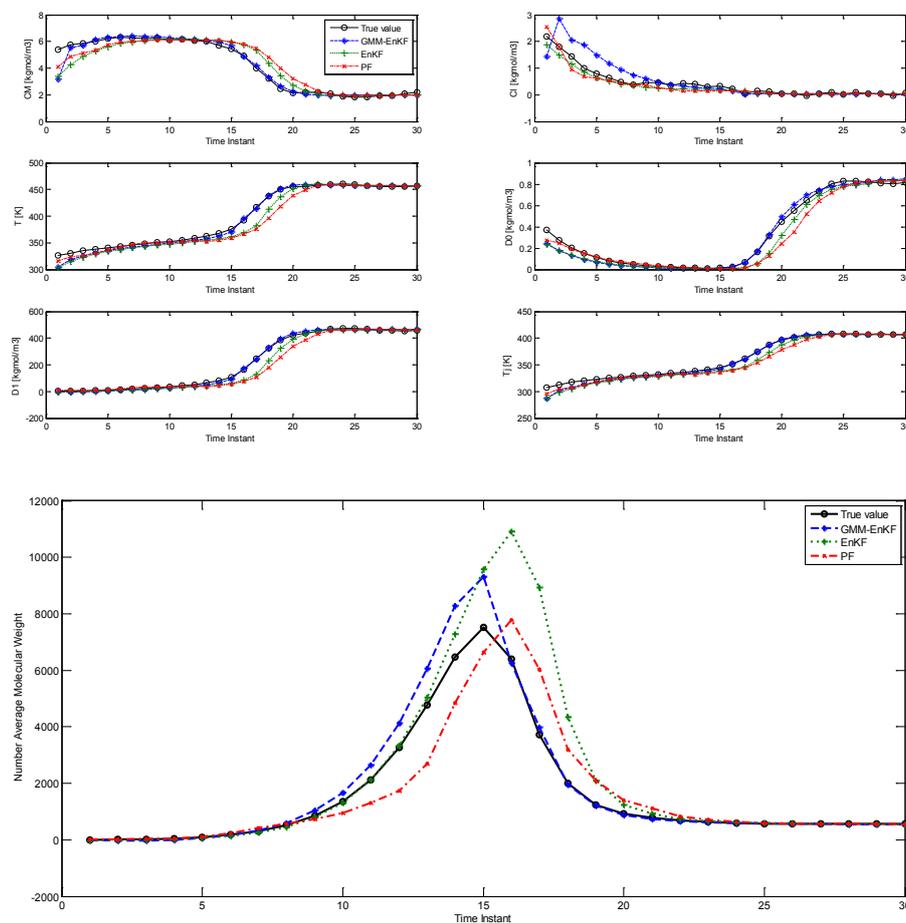
$$\mu_1 = \left[ 1 \frac{\text{kgmol}}{\text{m}^3}, 1 \frac{\text{kgmol}}{\text{m}^3}, 290 \text{ K}, 0.49 \frac{\text{kgmol}}{\text{m}^3}, 0.49 \frac{\text{kg}}{\text{m}^3}, 270 \text{ K} \right];$$

$$\mu_2 = \left[ 10 \frac{\text{kgmol}}{\text{m}^3}, 8 \frac{\text{kgmol}}{\text{m}^3}, 350 \text{ K}, 0.51 \frac{\text{kgmol}}{\text{m}^3}, 0.51 \frac{\text{kg}}{\text{m}^3}, 330 \text{ K} \right];$$

$$P_1 = \text{diag} (0.8, 0.8, 5.6, 8e - 2, 8e - 3, 5.6);$$

$$P_2 = \text{diag} (0.8, 0.8, 5.6, 8e - 2, 8e - 3, 5.6)$$

In this case, the parameters of the prior distribution indicate that both of the modes lie near the tail of the likelihood function. The initial particles not only show significant multimodality, but also some degree of model-plant mismatch. The comparison of estimation using the EnKF-GMM, EnKF, and PF is shown in Figure 2 and the RMSE is shown in Table 3, and it is clear that the EnKF-GMM outperforms the other two estimators. As expected, the performance of the EnKF has worsened in this case because its Gaussian assumption on the prior and posterior distributions is violated in a significant manner. The PF does not show good performance either, and it is outperformed by the EnKF in the estimation of the NAMW. This is because the PF lacks robustness to plant-model mismatch [14], which is present in this case. Increasing the number of particles for all the estimators does not change these conclusions.



**Figure 2.** Comparison of the estimation performance of the EnKF-GMM, EnKF, and PF for the PMMA process with more significant multimodal process noise (Case Study 2).

Figure 3 shows the evolution of the multimodal posterior distribution of the one of the states (the monomer concentration) at time steps 1, 3, 4, and 9. Table 4 lists the corresponding estimation errors of the three algorithms at those time steps with respect to the true value of  $C_M$ . Figure 4

shows the evolution of the posterior distribution of another state (the jacket temperature) at time steps 2, 6, 9, and 10, and Table 5 shows the corresponding estimation errors of the three algorithms. These distributions are bimodal, and this clearly shows that the EnKF-GMM outperforms the other estimators in the presence of multimodal distributions.

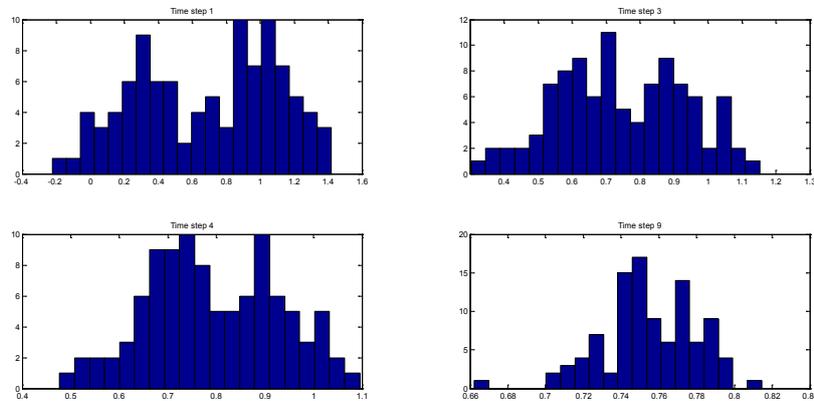


Figure 3. Evolution of the multimodal posterior distributions of  $C_M$  at time steps 1, 2, 4, and 9 (Case Study 2).

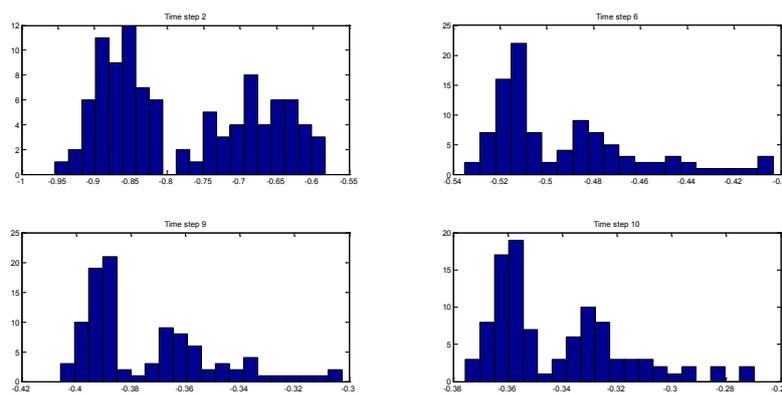


Figure 4. Evolution of the multimodal posterior distributions of  $T_j$  at time steps 2, 6, 9, and 10 (Case Study 2).

Table 3. RMSE of the EnKF-GMM, EnKF, and PF for the PMMA process with more significant multimodal process noise (Case Study 2).

Variable	EnKF-GMM	EnKF	PF
$C_M, \text{kg}\cdot\text{mol}/\text{m}^3$	0.44	0.68	0.69
$C_I, \text{kg}\cdot\text{mol}/\text{m}^3$	0.37	0.14	0.17
$T, \text{K}$	5.8	11.8	14.4
$D_0, \text{kg}\cdot\text{mol}/\text{m}^3$	0.042	0.062	0.078
$D_1, \text{kg}/\text{m}^3$	9.73	36.13	51.38
$T_j, \text{K}$	5.1	8.2	9.2
NAMW	559	1400	831

Table 4. Comparison of the estimation errors of the EnKF-GMM, EnKF, and PF for  $C_M$  at time steps 1, 3, 4, and 9 (in  $\text{kg}\cdot\text{mol}/\text{m}^3$ ) (Case Study 2).

Estimator	Time Step 1	Time Step 3	Time Step 4	Time Step 9
EnKF-GMM	0.23	0.14	0.40	0.04
EnKF	2.06	1.06	0.80	0.10
PF	3.60	2.20	1.65	0.22

**Table 5.** Comparison of the estimation errors of the EnKF-GMM, EnKF, and PF for  $T_j$  at time steps 2, 6, 9, and 10 (in  $K$ ) (Case Study 2).

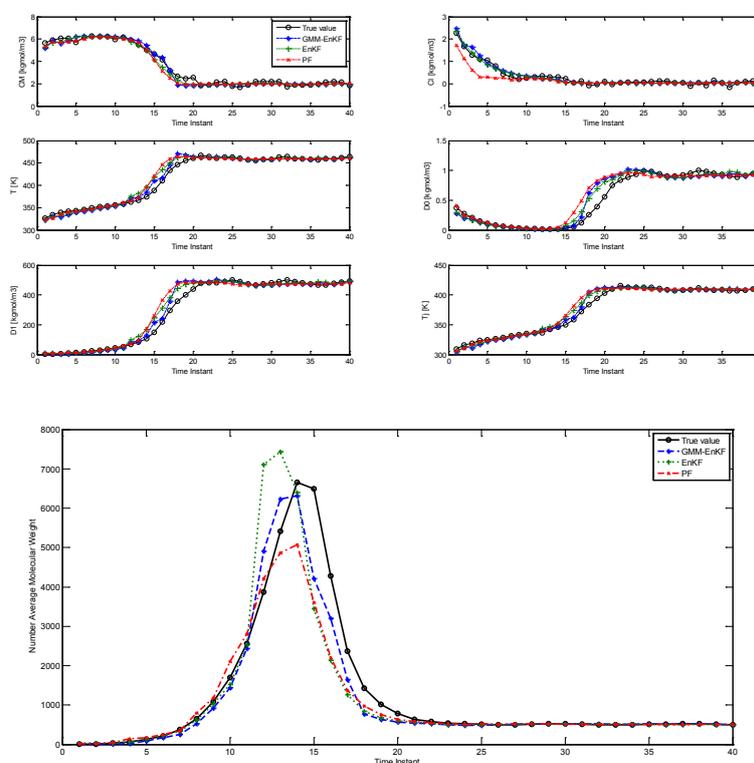
Estimator	Time Step 1	Time Step 3	Time Step 4	Time Step 9
EnKF-GMM	6.4	2.8	1.5	1.3
EnKF	6.6	3.0	1.9	1.7
PF	13.5	4.5	2.9	2.9

3.3. Comparison of State and Parameter Estimation with the EnKF-GMM, EnKF and PF (Case Studies 3 and 4)

We consider the effects of parametric uncertainty in this section. The uncertain parameter chosen for these studies is  $E_p$ , which is the activation energy associated with the reaction rate parameter  $k_p$ . We choose  $E_p$  as the uncertain parameter because (based on dimensionless sensitivity analysis) the NAMW is highly sensitive to the values of this parameter. We consider state estimation and joint state and parameter estimation in this section.

3.3.1. State Estimation with Uncertain Parameter (Case Study 3)

In this sub-section, while  $E_p$  is an uncertain parameter and noise is added to its value at each time step in the simulation, the parameter is not estimated. The nominal value of  $E_p$  is set to be  $E_p = 1.8283 \times \frac{10^4 \text{ kJ}}{\text{kgmol}}$ , and bimodal Gaussian noise with means of the modes  $\mu_1 = -100$ ,  $\mu_2 = 100$  and covariances  $P_1 = 50$ ,  $P_2 = 50$  is added to it. In addition, process and measurement noise with the same distributions as in the second case in Section 3.2 are included. Figure 5 shows the comparison of the estimation results using the three algorithms over 40 time steps, and Table 6 shows the corresponding RMSE. In this case, the EnKF-GMM shows a small improvement in state estimation performance over the other estimators, especially in the estimation of the NAMW.



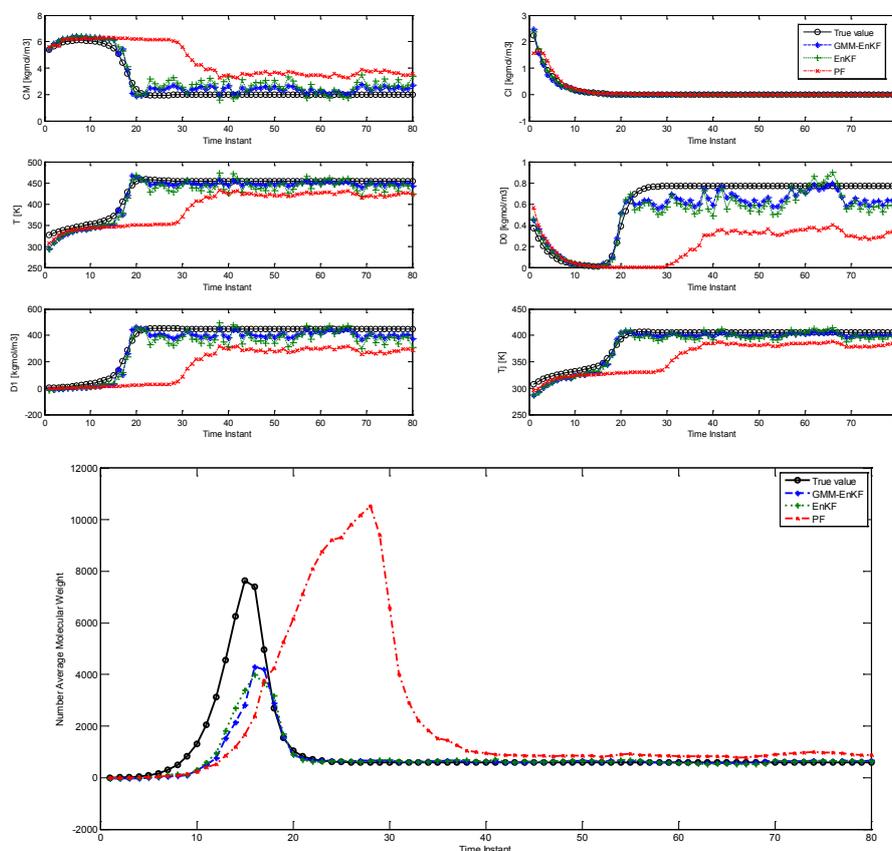
**Figure 5.** Comparison of state estimation with the EnKF-GMM, EnKF, and PF for the PMMA process with uncertain parameter  $E_p$  (Case Study 3).

**Table 6.** RMSE of the EnKF-GMM, EnKF, and PF for state estimation in the case with uncertain parameter  $E_p$  (Case Study 3).

Variable	EnKF-GMM	EnKF	PF
$C_M, \text{kg} \cdot \text{mol}/\text{m}^3$	0.29	0.26	0.32
$C_I, \text{kg} \cdot \text{mol}/\text{m}^3$	0.12	0.10	0.27
T, K	7.2	8.9	10.3
$D_0, \text{kg} \cdot \text{mol}/\text{m}^3$	0.111	0.092	0.144
$D_1, \text{kg}/\text{m}^3$	32.27	35.11	45.34
$T_j, \text{K}$	5.5	5.7	7.5
NAMW	487	869	653

### 3.3.2. State and Parameter Estimation with Uncertain Parameter (Case Study 4)

Next, we compare the performance of the estimators for joint state and parameter estimation. Once again,  $E_p$  is the uncertain parameter and its nominal value is kept the same as in Case Study 3. The parameter  $E_p$  is treated as an augmented state for estimation. The prior distribution for  $E_p$  has the following characteristics: means of  $\mu_1 = 1.9 \times 10^4$ ,  $\mu_2 = 2.5 \times 10^4$  for its two modes, and covariances of  $P_1 = 500$ ,  $P_2 = 500$ . Bimodal noise is added to each particle of the parameter, with means  $\mu_1 = -100$ ,  $\mu_2 = 100$  and covariances  $P_1 = 50$ ,  $P_2 = 50$ . Except for the exclusion of process noise, the properties of the simulation are kept the same as in Case Study 3. Figure 6 shows the performance of the estimators in state estimation, and Figure 7 shows their performance in estimating the parameter  $E_p$ . While the performance of the EnKF in state estimation is comparable to that of the EnKF-GMM, the EnKF-GMM is clearly superior in parameter estimation. The PF has the worst performance among the estimators.



**Figure 6.** Comparison of state estimation with the EnKF-GMM, EnKF, and PF for the PMMA process with uncertain parameters (Case Study 4).

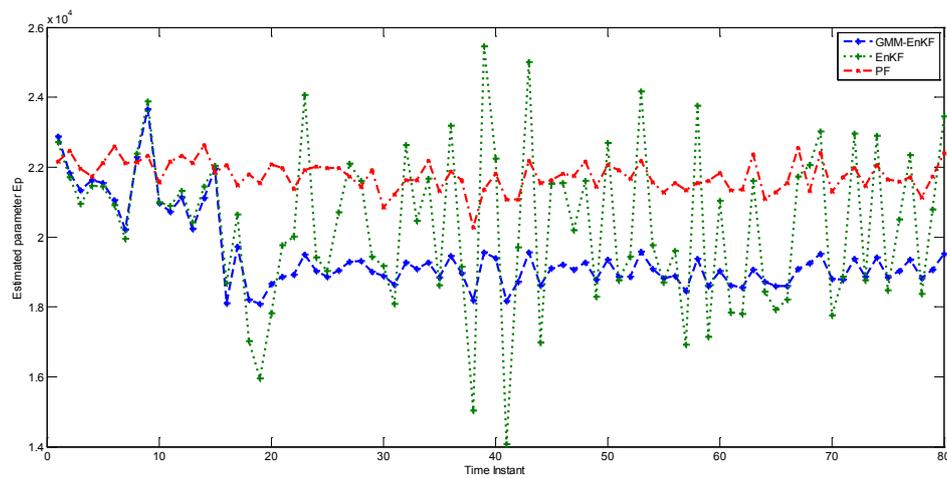


Figure 7. Parameter estimation using the EnKF-GMM, EnKF, and PF (Case Study 4).

### 3.4. Alternate Point Estimates for the PF (Case Study 5)

In the PF, even though the full distribution is obtained, a point estimate for the states is usually obtained by choosing the expectation (mean) of the posterior particles. This is the method we have employed for the PF in the simulations described in the previous sections. However, if the distribution is multimodal, the mean may not necessarily represent the best point estimate, and the mode of the distribution (which is equivalent to the maximum *a posteriori* estimate) can provide a better estimate [14,31]. We investigate whether this approach can improve the performance of the PF, since we are considering cases where the distributions are multimodal. We apply *k*-means clustering on the posterior distribution of the particles to identify the modes and the maximum *a posteriori* estimate with the particle filter, and compare the estimation performance of this PF, called the PF-mode, with the other estimators. The parameters of the simulations are similar to the second case study. Figure 8 shows the performance of the estimators, and the RMSE is described in Table 7. The PF-mode clearly outperforms the PF and the EnKF; however, the EnKF-GMM has superior performance.

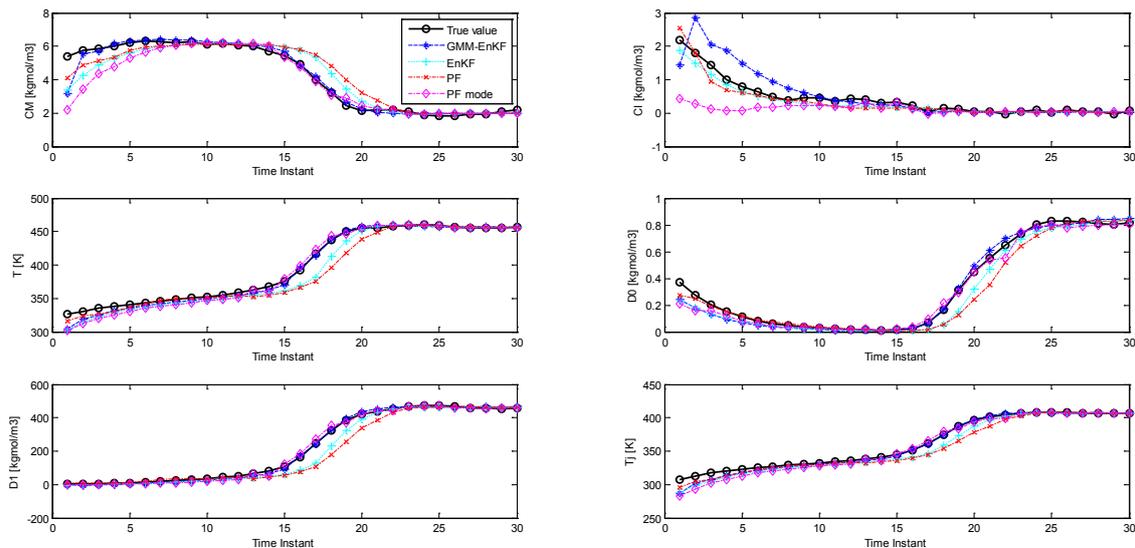
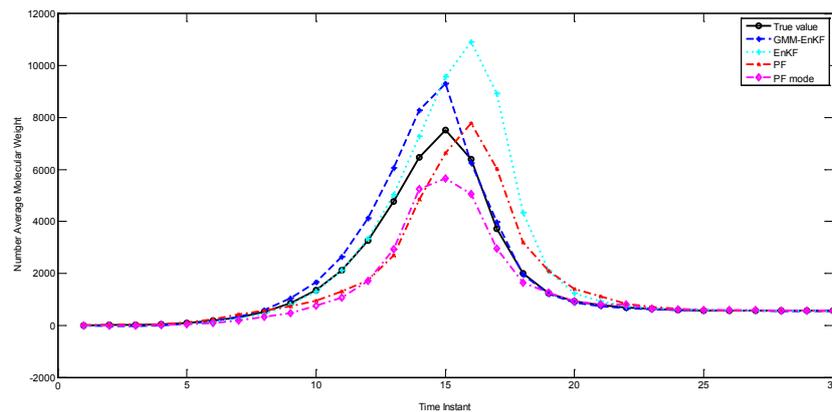


Figure 8. Cont.



**Figure 8.** Comparison of state estimation with the EnKF-GMM, EnKF, PF, and PF-mode (Case Study 5).

**Table 7.** RMSE of the EnKF-GMM, EnKF, PF, and PF-mode for state estimation (Case Study 5).

Variable	EnKF-GMM	EnKF	PF	PF-mode
$C_M$ , kg·mol/m <sup>3</sup>	0.44	0.68	0.68	0.85
$C_I$ , kg·mol/m <sup>3</sup>	0.37	0.14	0.17	0.55
$T$ , K	5.8	11.8	14.4	8.31
$D_0$ , kg·mol/m <sup>3</sup>	0.042	0.062	0.078	0.047
$D_1$ , kg/m <sup>3</sup>	9.73	36.13	51.38	13.05
$T_j$ , K	5.1	8.2	9.2	7.9
NAMW	559	1400	831	706

The idea of the PF-mode is very similar to that of the EnKF-GMM. Both of them use clustering to extract modes from the posterior distribution and generate a point estimate based on the information in the modes. However, the EnKF-GMM outperforms the PF-mode because it is more robust to poor initial estimates and model-plant mismatch. Also, if the number of modes in the state distributions varies with time, perhaps even becoming unimodal at some times, using the mode as a point estimate is not necessarily superior to the mean. The EnKF-GMM combines the modes of the distribution in proportion based on the calculated weights to get a point estimate, and can adjust its estimation results in these cases by adjusting the weights of the modes.

#### 4. Conclusions

We have proposed an estimator based on a Gaussian mixture model coupled with an ensemble Kalman filter (EnKF-GMM) that is capable of handling multimodal state distributions, and demonstrated its performance in simulations on a polymethyl methacrylate process. The EnKF-GMM clearly outperforms the particle filter (PF) and the EnKF in both state and parameter estimation with multimodal distributions. The EnKF is limited by the assumption of Gaussian distributions, and the particle filter's performance is affected by its lack of robustness with respect to model-plant mismatch. A different choice for obtaining a point estimate with the particle filter, leading to a maximum *a posteriori* estimate, improves the performance of the PF, but the EnKF-GMM is still superior, indicating that it is the estimator of choice for systems with multimodal state distributions such as polymer processes.

**Acknowledgments:** The authors acknowledge financial support from the China Scholarship Council and the Natural Sciences and Engineering Research Council of Canada.

**Author Contributions:** All three authors conceived the work and participated in defining its scope. Ruoxia Li developed the algorithms and conducted the simulations described in the manuscript with inputs from Vinay Prasad and Biao Huang. Ruoxia Li wrote the initial drafts of the manuscript, and all authors contributed to the editing of the final manuscript and to the revisions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wilson, D.; Agarwal, M.; Rippin, D. Experiences implementing the extended Kalman filter on an industrial batch reactor. *Comput. Chem. Eng.* **1998**, *22*, 1653–1672. [[CrossRef](#)]
2. Prasad, V.; Schley, M.; Russo, L.P.; Bequette, W.B. Product property and production rate control of styrene polymerization. *J. Process Control* **2002**, *12*, 353–372. [[CrossRef](#)]
3. Jo, J.; Bankoff, S. Digital monitoring and estimation of polymerization reactor. *AIChE J.* **1976**, *22*, 361–368. [[CrossRef](#)]
4. Kozub, D.; MacGregor, J. State estimation for semi-batch polymerization reactors. *Chem. Eng. Sci.* **1992**, *47*, 1047–1062. [[CrossRef](#)]
5. McAuley, K.; MacGregor, J. On-line inference of polymer properties in an industrial polymer properties in an industrial polyethylene reactor. *AIChE J.* **1991**, *37*, 825–835. [[CrossRef](#)]
6. McAuley, K.; MacGregor, J. Nonlinear product property control in industrial gas-phase polyethylene reactors. *AIChE J.* **1993**, *39*, 855–866. [[CrossRef](#)]
7. Srinivas, G.; Arkun, Y.; Schork, F. Estimation and control of an alpha-olefin polymerization reactor. *J. Process Control* **1994**, *5*, 303–313. [[CrossRef](#)]
8. Gopalakrishnan, A.; Kaisare, N.S.; Narasimhan, S. Incorporating delayed and infrequent measurements in extended Kalman filter based nonlinear state estimation. *J. Process Control* **2011**, *21*, 119–129. [[CrossRef](#)]
9. Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **1994**, *99*, 10143–10162. [[CrossRef](#)]
10. Julier, S.; Uhlmann, J.; Durrant-Whyte, H. A new approach for filtering nonlinear systems. In Proceedings of the American Control Conference, Seattle, WA, USA, 21–23 June 1995.
11. Julier, S.; Uhlmann, J.; Durrant-Whyte, H. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Autom. Control* **2000**, *45*, 477–482. [[CrossRef](#)]
12. Arulampalam, S.; Maskell, S.; Gordon, N.; Clapp, T. A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Trans. Autom. Control* **2002**, *30*, 174–189.
13. Chen, T.; Morris, J.; Martin, E. Particle filters for state and parameter estimation in batch processes. *J. Process Control* **2005**, *15*, 665–673. [[CrossRef](#)]
14. Shenoy, A.V.; Prakash, J.; Prasad, V.; Shah, S.L.; McAuley, K.B. Practical issues in state estimation using particle filters: Case studies with polymer reactors. *J. Process Control* **2013**, *23*, 120–131. [[CrossRef](#)]
15. Shao, X.; Huang, B.; Lee, J.M. Constrained Bayesian state estimation- A comparative study and a new particle filter based approach. *J. Process Control* **2010**, *20*, 143–157. [[CrossRef](#)]
16. Crowley, T.J.; Meadows, E.S.; Kostoulas, E.; Doyle, F.D. Control of particle size distribution described by a population balance model of semibatch emulsion polymerization. *J. Process Control* **2000**, *10*, 419–432. [[CrossRef](#)]
17. Kiparissides, C. Challenges in particulate polymerization reactor modeling and optimization: A population balance perspective. *J. Process Control* **2006**, *16*, 205–224. [[CrossRef](#)]
18. Sajjadi, S.; Brooks, B.W. Unseeded semibatch emulsion polymerization of butyl acrylate: Bimodal particle size distribution. *J. Polym. Sci. A Polym. Chem.* **2000**, *38*, 528–545. [[CrossRef](#)]
19. Doyle, F.J.; Soroush, M.; Cordeiro, C. Control of product quality in polymerization processes. In Proceedings of the Sixth International Conference on Chemical Process Control, Tucson, AZ, USA, 7–12 January 2001; AIChE Press: New York, NY, USA, 2002; pp. 290–306.
20. Flores-Cerrillo, J.; MacGregor, J.F. Control of particle size distributions in emulsion semibatch polymerization using mid-course correction policies. *Ind. Eng. Chem. Res.* **2002**, *41*, 1805–1814. [[CrossRef](#)]
21. Bengtsson, T.; Snyder, C.; Nychka, D. Toward a nonlinear ensemble filter for high-dimensional systems. *J. Geophys. Res.* **2003**, *108*, 35–45. [[CrossRef](#)]
22. Smith, K.W. Cluster ensemble Kalman filter. *Tellus* **2007**, *59A*, 749–757. [[CrossRef](#)]
23. Dovera, L.; Della Rossa, E. Multimodal ensemble Kalman filtering using Gaussian mixture models. *Comput. Geosci.* **2011**, *15*, 307–323. [[CrossRef](#)]
24. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **1977**, *39*, 1–38.

25. Ormoneit, D.; Tresp, V. Improved Gaussian mixture density estimates using Bayesian penalty terms and network averaging. In *Advances in Neural Information Processing Systems 8*; Touretzky, D.S., Tesauro, G., Leen, T.K., Eds.; MIT Press: Cambridge, MA, USA, 1996; pp. 542–548.
26. Ueda, N.; Nakano, R.; Ghahramani, Z.; Hinton, G.E. SMEM algorithm for mixture models. *Neural Comput.* **2000**, *12*, 2109–2128. [[CrossRef](#)] [[PubMed](#)]
27. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [[CrossRef](#)]
28. Hu, X.; Xu, L. Investigation on several model selection criteria for determining the number of cluster. *Neural Inf. Process.-Lett. Rev.* **2004**, *4*, 1–10.
29. Silva-Beard, A.; Flores-Tlacuahuac Silva, A. Effect of process design/operation on the steady-state operability of a methyl methacrylate polymerization reactor. *Ind. Eng. Chem. Res.* **1999**, *38*, 4790–4804. [[CrossRef](#)]
30. Shenoy, A.V.; Prasad, V.; Shah, S.L. Comparison of unconstrained nonlinear state estimation techniques on a MMA polymer reactor. In Proceedings of the 9th International Symposium on Dynamics and Control of Process Systems (DYCOPS 2010), Leuven, Belgium, 5–9 July 2010.
31. Bavdekar, V.A.; Shah, S.L. Computing point estimates from a non-Gaussian posterior distribution using a probabilistic k-means clustering approach. *J. Process Control* **2014**, *24*, 487–497. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).