An Extended Moving Horizon Estimation embedded with an Abridged Gaussian Sum Extended Kalman Filer to handle non-Gaussian noises

Mahshad Valipour*, Luis A. Ricardez-Sandoval**

**Chemical Engineering Department, University of Waterloo, Waterloo, Canada (Tel: 519-781-2161; email: mahshad.valipour@uwaterloo.ca).*
*** Chemical Engineering Department, University of Waterloo, Waterloo, Canada (Tel: 519-888-4567x38667; email: laricard@uwaterloo.ca).*

Abstract: Most industrial processes involve changes in the operating conditions that may lead to non-Gaussian process uncertainties and measurement noises. Recently, an Abridged Gaussian Sum Extended Kalman Filter (AGS-EKF) and Extended Moving Horizon Estimation (EMHE) frameworks were proposed to capture non-Gaussian random uncertainties and noises often present in the chemical systems. Gaussian mixture models (GMM) are used in both estimation schemes to efficiently approximate the non-Gaussian distributions. Previous studies on EMHE considered a sufficiently long estimation horizon to minimize the arrival cost effect. The present work aims to further improve the performance of EMHE by introducing a suitable arrival cost estimator to shorten the estimation horizon. As the focus of this study is on systems involving non-Gaussian noises, AGS-EKF as a non-Gaussian state estimator is selected to estimate the arrival cost. The performance of the proposed estimation framework was tested using the open-loop unstable Williams-Otto reactor considering non-Gaussian uncertainties and noises. The results revealed that the proposed estimation framework improves the estimation in the presence of non-Gaussian noises when compared to the standard framework (MHE combined with EKF) thus making it a suitable estimation method for systems involving non-Gaussian noises.

Keywords: Non-Gaussian noise, Abridged Gaussian sum extended Kalman filter, Extended moving horizon estimation, Gaussian mixture model, Arrival cost

1. INTRODUCTION

Online measurement technologies are not often available for key state variables that are critical for online monitoring and control. Kalman Filter (KF) and its extensions, e.g., Extended Kalman Filter (EKF), as well as Moving Horizon Estimation (MHE) are well-known model-driven state estimation schemes for applications featuring zero-mean Gaussian process uncertainties and measurement noises. However, chemical processes often involve changes in the operating conditions that may lead to a general class of non-Gaussian process uncertainties and measurement noises (e.g., uniform, multimodal). It is thus essential that state estimation schemes can properly capture the non-Gaussianity to successfully monitor and control chemical plants. In our previous works, an efficient EKF-based framework, referred to as *Abridged Gaussian Sum Extended Kalman Filter* (AGS-EKF) was introduced to capture the non-Gaussianity in the process variables (Valipour and Ricardez-Sandoval, 2021a)(Valipour and Ricardez-Sandoval, 2021b). Similarly, an *Extended Moving Horizon Estimation* (EMHE) that relaxes the Gaussian assumption in the standard MHE was proposed (Valipour and Ricardez-Sandoval, 2021c). That approach improves the estimation accuracy (in the context of a moving horizon estimation framework) for applications involving a general class of non-Gaussian process uncertainties and measurement noises. The key idea in both AGS-EKF and EMHE schemes is to approximate the non-Gaussian distributions of the process uncertainties and measurement noises using Gaussian mixture models (GMM). In the conventional Gaussian Sum Filter (GSF), multiple EKFs are performed simultaneously (one EKF using one Gaussian component), and a summation over the weighted estimation provided by each EKF represents the point estimates. Performing multiple EKFs based on individual Gaussian components may increase the computational costs and lead to biased estimations. To avoid these issues, AGS-EKF and EMHE avoid performing multiple EKFs and MHEs for each of the Gaussian components in the GMM. Instead, the main characteristics (i.e., the mean value and the covariance matrix) of the overall GMM of the corresponding process uncertainty variable (or measurement noise variable) are used in AGS-EKF and EMHE schemes to represent the non-Gaussian distributions in the state estimation framework. Both AGS-EKF and EMHE schemes provide the point estimates by performing EKF and MHE only once, thus avoiding the additional computational costs and biased estimations often observed in GSFs. Moreover, for applications where distributions of the process uncertainty and measurement noise are known *a priori* and remain unchanged throughout the plant operation, the GMMs of the random uncertainties and noises can be approximated offline. In such scenarios, EMHE and AGS-EKF can improve the accuracy in the estimation at no additional computational costs.

Similar to the standard MHE, a shorter horizon is favourable for the purpose of online estimation and closed-loop operation

though it requires a more accurate approximation of the arrival cost. However, in (Valipour and Ricardez-Sandoval, 2021c), a sufficiently long estimation horizon was considered to minimize the effect of arrival cost. Thus, the performance of EMHE while considering the effect of arrival cost is a topic of interest that have not been investigated to date. Motivated by this, the current work aims to investigate the performance of EMHE coupled with AGS-EKF as the arrival cost estimator. As the focus of this work is on applications involving non-Gaussian process uncertainties and measurement noises, AGS-EKF is a suitable candidate as the arrival cost estimator due to its efficiency and accuracy for such applications. The Williams-Otto reactor as a highly nonlinear and open-loop unstable system is considered as the case study to investigate the performance of the proposed framework.

This study is organized as follows: The general formulation of EMHE (the lead state estimation scheme) and AGS-EKF (the arrival cost estimator) is presented next. Section 3 presents the computational experiments performed on the Williams-Otto reactor to show the benefits of proposed estimation scheme. Concluding remarks and future work are presented at the end.

2. NON-GAUSSIAN STATE ESTIMATION SCHEME

Consider a general nonlinear dynamic system with states $\boldsymbol{x}$, inputs $\boldsymbol{u}$, measurements $\boldsymbol{y}$, and non-Gaussian process uncertainties ($\boldsymbol{w}$) and measurement noises ($\boldsymbol{v}$):

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k \tag{1}$$
$$\boldsymbol{y}_k = h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{v}_k$$
$$\boldsymbol{w}, \boldsymbol{x} \in \mathbb{R}^{n_x}, \boldsymbol{v}, \boldsymbol{y} \in \mathbb{R}^{n_y}, \boldsymbol{u} \in \mathbb{R}^{n_u},$$
$$f: \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_x}, h: \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_y}$$

where k denotes the time interval. Note that $\boldsymbol{w}$ and $\boldsymbol{v}$ are mutually uncorrelated. Moreover, f and h describe the process model and the measurement model, respectively. Given the nonlinear model presented in (1), the GMMs of the process uncertainties and measurement noises are described next. The non-Gaussian state estimation framework involving EMHE as the lead state estimation scheme engaged with the AGS-EKF as the arrival cost estimator are presented in section 2.2.

2.1 Gaussian Mixture Model

GMMs is a mixture of several Gaussian distributions that can be used to approximate non-Gaussian densities of the process uncertainties and measurement noises. That is, the process uncertainties $\boldsymbol{w}$ can be approximated by a mixture of ngp Gaussian components, i.e., $\mathcal{N}(\boldsymbol{\mu}^i, \boldsymbol{Q}^i)$, as follows:

$$p(\boldsymbol{w}) = \sum_{i=1}^{ngp} (\boldsymbol{\alpha}^i \, \mathcal{N}[\boldsymbol{w};\, \boldsymbol{\mu}^i, \boldsymbol{Q}^i]);$$
$$\boldsymbol{\alpha}^i, \boldsymbol{\mu}^i \in \mathbb{R}^{n_x}; \boldsymbol{Q}^i \in \mathbb{R}^{n_x \times n_x} \tag{2}$$

$$\sum_{i=1}^{ngp} \boldsymbol{\alpha}^i = \mathbf{1}; \quad \boldsymbol{\alpha}^i \geq \mathbf{0}; \tag{3}$$

where $\boldsymbol{\alpha}^i$ represents the weight assigned to the i^{th} Gaussian components in the mixture. These non-negative weights are normalized; thus, for each of the state variables included in n_x the summation of the weights adds to unity, as shown in (3). Likewise, GMMs involving ngm Gaussian components, i.e., $\mathcal{N}(\boldsymbol{\tau}^i, \boldsymbol{R}^i)$, are considered to approximate the non-Gaussian distributions of the measurement noises, i.e.,

$$p(\boldsymbol{v}) = \sum_{i=1}^{ngm} (\boldsymbol{\beta}^i \, \mathcal{N}[\boldsymbol{v};\, \boldsymbol{\tau}^i, \boldsymbol{R}^i]);$$
$$\boldsymbol{\beta}^i, \boldsymbol{\tau}^i \in \mathbb{R}^{n_y}; \boldsymbol{R}^i \in \mathbb{R}^{n_y \times n_y} \tag{4}$$

$$\sum_{i=1}^{ngm} \boldsymbol{\beta}^i = \mathbf{1}; \quad \boldsymbol{\beta}^i \geq \mathbf{0}; \tag{5}$$

where $\boldsymbol{\beta}^i$ is the corresponding weights assigned to the i^{th} Gaussian components in the mixture. The mean-value ($\boldsymbol{\tau}^i$ and $\boldsymbol{\mu}^i$), covariance matrix ($\boldsymbol{R}^i$ and $\boldsymbol{Q}^i$), and the weights ($\boldsymbol{\beta}^i$ and $\boldsymbol{\alpha}^i$) corresponding to each Gaussian component in the mixture represent the parameters that need to be estimated in the GMM. The Expectation-Maximization (EM) algorithm is an efficient method used to estimate the GMM parameters. This method starts with an adequate user-defined initial guess for the Gaussian components' parameters. The first step in the EM algorithm is to draw a sufficiently large number of samples from the original non-Gaussian distribution. The Expectation-step (E-step) determines the possibility that a sample follow the i^{th} Gaussian component in the mixture (membership weight). The membership weight for each sample is determined based on the probability density function (likelihood) of each sample given the i^{th} Gaussian component. The next step, Maximization-step (M-step), the membership weights provided by E-step are used to update the corresponding weight, mean value, and covariance of each Gaussian component in the GMM. Both the E-step and M-step are performed in a recursive fashion until the likelihood does not change significantly from one iteration to another. More details about this method are found elsewhere (Dinov, 2008)(Bilmes, 1998).

The number of Gaussian components in the mixture (i.e., ngm and ngp), is a user defined parameter, which is a trade-off between the accuracy of GMM and CPU time (i.e., the accuracy improves as the number of Gaussian components increases, at the expense of larger computational costs). Preliminary tests are thus required to determine a suitable number of ngm and ngp. These tests begin with the smallest reasonable guess for the number of Gaussian components. For instance, if the distribution is multi-modal, an appropriate choice is to set the Gaussian components to the number of modes in the non-Gaussian density. Given this initial guess for ngm and ngp, the EM algorithm is performed to obtain the GMMs. The test stops when the resulting GMMs can adequately capture the non-Gaussian distributions of the random noises and process uncertainties. Otherwise, the tests continue with a one-to-one increase in the number Gaussian components until the desired accuracy for the GMM is satisfied. As these tests are not intensive and are performed offline, they do not affect the computational costs of the online state estimation schemes. A detailed analysis on the CPU costs was presented in our previous work (Valipour and Ricardez-Sandoval, 2021a).

Once the individual components in the GMMs presented in (2)-(5) have been identified and estimated, the mean value ($\boldsymbol{\mu}^{\boldsymbol{GM}}$ and $\boldsymbol{\tau}^{\boldsymbol{GM}}$) and covariance matrix ($\boldsymbol{Q}^{\boldsymbol{GM}}$ and $\boldsymbol{R}^{\boldsymbol{GM}}$) of the overall GMMs of the uncertainties and noises can be determined based on the mean value, covariance, and the weights corresponding to each i^{th} Gaussian components in the mixture, i.e.,

$$\boldsymbol{\mu}^{\boldsymbol{GM}} = \sum_{i=1}^{ngp} \boldsymbol{\alpha}^i \odot \boldsymbol{\mu}^i \tag{6}$$

$$\boldsymbol{Q}^{\boldsymbol{GM}} = \sum_{i=1}^{ngp} \boldsymbol{\alpha}^i \odot \boldsymbol{Q}^i + \sum_{i=1}^{ngp} \boldsymbol{\alpha}^i \odot (\boldsymbol{\mu}^i - \boldsymbol{\mu}^{\boldsymbol{GM}})(\boldsymbol{\mu}^i - \boldsymbol{\mu}^{\boldsymbol{GM}})^{\boldsymbol{T}}$$

$$\boldsymbol{\tau}^{\boldsymbol{GM}} = \sum_{i=1}^{ngm} \boldsymbol{\beta}^i \odot \boldsymbol{\tau}^i$$

$$\boldsymbol{R}^{\boldsymbol{GM}} = \sum_{i=1}^{ngm} \boldsymbol{\beta}^i \odot \boldsymbol{R}^i + \sum_{i=1}^{ngm} \boldsymbol{\beta}^i \odot (\boldsymbol{\tau}^i - \boldsymbol{\tau}^{\boldsymbol{GM}})(\boldsymbol{\tau}^i - \boldsymbol{\tau}^{\boldsymbol{GM}})^{\boldsymbol{T}}$$

$$\boldsymbol{\mu}^{\boldsymbol{GM}} \in \mathbb{R}^{n_x}; \boldsymbol{\tau}^{\boldsymbol{GM}} \in \mathbb{R}^{n_y}; \boldsymbol{Q}^{\boldsymbol{GM}} \in \mathbb{R}^{n_x \times n_x}; \boldsymbol{R}^{\boldsymbol{GM}} \in \mathbb{R}^{n_y \times n_y}$$

where symbol $\odot$ indicates an element-wise multiplication.

2.2 EMHE combined with AGS-EKF framework

The EMHE problem considers $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{GM}}, \boldsymbol{Q}^{\boldsymbol{GM}})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}^{\boldsymbol{GM}}, \boldsymbol{R}^{\boldsymbol{GM}})$ determined from (6) to describe the process uncertainties and measurement noises affecting the operation, respectively. Therefore, the EMHE problem can be described as follows(Valipour and Ricardez-Sandoval, 2021c):

$$\min_{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=k-N}^{k-1}} \Psi + \Lambda + \varphi_{k-N} \tag{7}$$

s.t.

$$\Psi = \sum_{j=k-N}^{k-1} \left\| \boldsymbol{w}_j - \boldsymbol{\mu}^{\boldsymbol{GM}} \right\|^2_{\boldsymbol{Q}^{\boldsymbol{GM}^{-1}}}$$

$$\Lambda = \sum_{j=k-N+1}^{k} \left\| \boldsymbol{v}_j - \boldsymbol{\tau}^{\boldsymbol{GM}} \right\|^2_{\boldsymbol{R}^{\boldsymbol{GM}^{-1}}}$$

$$\boldsymbol{x}_{j+1} = f(\boldsymbol{x}_j, \boldsymbol{u}_j) + \boldsymbol{w}_j \; ; \forall j = k-N, \dots k-1$$

$$\boldsymbol{y}_j = h(\boldsymbol{x}_j, \boldsymbol{u}_j) + \boldsymbol{v}_j \; ; \; \forall j = k-N+1, \dots k$$

$$g(\boldsymbol{x}_j, \boldsymbol{u}_j, \boldsymbol{w}_j, \boldsymbol{y}_j) \leq 0; \; \forall j = k-N, \dots k$$

$$\boldsymbol{x}^{\boldsymbol{l}} \leq \boldsymbol{x}_j \leq \boldsymbol{x}^{\boldsymbol{u}} \; ; \qquad \forall j = k-N, \dots k$$

$$g: \mathbb{R}^{n_x \times n_u \times n_x \times n_y} \to \mathbb{R}^{n_g}; \; \boldsymbol{x}^{\boldsymbol{l}}, \boldsymbol{x}^{\boldsymbol{u}} \in \mathbb{R}^{n_x}$$

where the index j is the time interval within the estimation horizon N. Moreover, g represents a general set of the inequality constraints present in the system, excluding the bounds on the states and the inputs. The vectors $\boldsymbol{x}^{\boldsymbol{l}}$, and $\boldsymbol{x}^{\boldsymbol{u}}$ represent the lower and upper bounds on the states, respectively. As shown in (7), Ψ and Λ represent the ℓ^2-norm for the process uncertainties and measurement noises, respectively. Note that the standard MHE is a special case of EMHE that holds the zero-mean Gaussian assumption for the noises and uncertainties. That is, EMHE is equivalent to the standard MHE problem when the system follows Gaussian distributions for the process uncertainty and measurement noises and $\boldsymbol{\mu}^{\boldsymbol{GM}}$=$\boldsymbol{0}$ and $\boldsymbol{\tau}^{\boldsymbol{GM}} = \boldsymbol{0}$. The arrival cost ($\varphi_{k-N}$) in problem (7) is defined as follows:

$$\varphi_{k-N} = \left\| \boldsymbol{x}_{k-N} - \bar{\boldsymbol{x}}_{k-N} \right\|^2_{\boldsymbol{P}^{-1}_{k-N}} \tag{8}$$

where $\bar{\boldsymbol{x}}_{k-N}$ and $\boldsymbol{P}_{k-N}$ are the mean value and the covariance matrix of the arrival cost term that are typically estimated online using a standard state estimation scheme, e.g., EKF, Particle Filter (PF), Unscented Kalman Filter (UKF) (López-Negrete et al., 2011). The most common arrival cost estimator is EKF due to its efficiency and accuracy. However, both EKF and UKF consider that the random noises and uncertainties follow zero-mean Gaussian distributions. Although PF is a sampling-based approach that can deal with non-Gaussian distributions, it requires significant additional computational costs when compared to EKF. Thus, in present study, AGS-EKF is used as the arrival cost estimator to perform EMHE for applications featuring non-Gaussian process uncertainties and measurement noises. The AGS-EKF is presented next.

The following expressions describe the AGS-EKF method as the arrival cost estimator for non-Gaussian process uncertainties and measurement noises approximated using $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{GM}}, \boldsymbol{Q}^{\boldsymbol{GM}})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}^{\boldsymbol{GM}}, \boldsymbol{R}^{\boldsymbol{GM}})$ (Valipour and Ricardez-Sandoval, 2021a)(Valipour and Ricardez-Sandoval, 2021b):

$$\hat{\boldsymbol{x}}_{k-N|k-N-1} = \boldsymbol{A}_{k-N-1} \hat{\boldsymbol{x}}^{EMHE}_{k-N-1} + \boldsymbol{B}_{k-N-1} \boldsymbol{u}_{k-N-1} + \boldsymbol{\mu}^{\boldsymbol{GM}} \tag{9}$$

$$\boldsymbol{P}_{k-N|k-N-1} = \boldsymbol{A}_{k-N-1} \boldsymbol{P}_{k-N-1|k-N-1} \boldsymbol{A}^T_{k-N-1} + \boldsymbol{Q}^{\boldsymbol{GM}}$$

where $\boldsymbol{A}_{k-N-1} = \frac{\partial f(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{x}} | \boldsymbol{x} = \hat{\boldsymbol{x}}^{EMHE}_{k-N-1}, \boldsymbol{u} = \boldsymbol{u}_{k-N-1}$;
$\boldsymbol{B}_{k-N-1} = \frac{\partial f(\boldsymbol{x},\boldsymbol{u})}{\partial u} | \boldsymbol{x} = \hat{\boldsymbol{x}}^{EMHE}_{k-N-1}, \boldsymbol{u} = \boldsymbol{u}_{k-N-1}$

$$\boldsymbol{K}_{k-N} = \boldsymbol{P}_{k-N|k-N-1} \boldsymbol{H}^T_{k-N} / (\boldsymbol{H}_{k-N} \boldsymbol{P}_{k-N|k-N-1} \boldsymbol{H}^T_{k-N} + \boldsymbol{R}^{\boldsymbol{GM}}) \tag{10}$$

$$\bar{\boldsymbol{x}}_{k-N} = \hat{\boldsymbol{x}}_{k-N|k-N-1} + \boldsymbol{K}_{k-N} \left(\boldsymbol{y}_{k-N} - h(\hat{\boldsymbol{x}}_{k-N|k-N-1}, \boldsymbol{u}_{k-N}) - \boldsymbol{\tau}^{\boldsymbol{GM}} \right)$$

$$\boldsymbol{P}_{k-N} = (\boldsymbol{I} - \boldsymbol{K}_{k-N} \boldsymbol{H}_{k-N}) \boldsymbol{P}_{k-N|k-N-1}$$

where $\boldsymbol{H}_{k-N} = \frac{\partial h(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{x}} | \boldsymbol{x} = \hat{\boldsymbol{x}}^{EMHE}_{k-N-1}, \boldsymbol{u} = \boldsymbol{u}_{k-N}$

$$\boldsymbol{P} \in \mathbb{R}^{n_x \times n_x}, \boldsymbol{A} \in \mathbb{R}^{n_x \times n_x}, \boldsymbol{B} \in \mathbb{R}^{n_x \times n_u}, \boldsymbol{K} \in \mathbb{R}^{n_x \times n_y}, \boldsymbol{H} \in \mathbb{R}^{n_y \times n_x}$$

where $\boldsymbol{K}_{k-N}$ is the Kalman gain evaluated at the corresponding time interval $k-N$, which consists of elements with values between 0 and 1 that represent the reliability of the measurements, i.e., more accurate measurements make the Kalman gain elements closer to the unity. Moreover, $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{H}$ are the sensitivity matrices in the linearized state-space

model of the system around a nominal operating condition, i.e., they describe the relation between states with states, states with inputs, and states with measurements, respectively. Similar to EKF, these sensitivity matrices are updated at each time interval (i.e., when k is updated) based on the corresponding measurements, inputs, and estimated states available. Note that the AGS-EKF presented in (9)-(10) is a general formulation of the EKF, i.e., AGS-EKF and EKF are equivalent when $\boldsymbol{\mu^{GM}}=\mathbf{0}$ and $\boldsymbol{\tau^{GM}} = \mathbf{0}$.

Fig. 1 presents a schematic of the state estimation framework proposed in this work. Similar to MHE, the arrival cost (φ_{k-N}) in EMHE summarizes the historical information not considered in the estimation horizon, i.e., the information from the initial time interval 0 to the beginning of the estimation horizon $k-N$. That is, at the current time interval k, the arrival cost estimator uses the posterior estimations provided by EMHE at time $k-N-1$ (i.e., $\hat{\boldsymbol{x}}_{k-N-1}^{EMHE}$ in (9)) as well as the online measurements available from the sensors at the instant ($\boldsymbol{y}_{k-N}$) to provide an approximation of the posterior distribution of states at time $k-N$ to EMHE, as shown in (10).

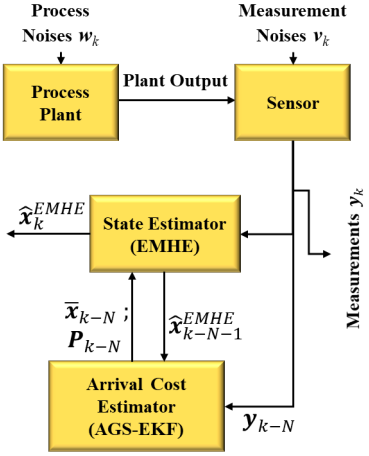

Fig. 1. A schematic of the state estimation framework

3. COMPUTATIONAL EXPERIMENTS

The computational experiments conducted to this work aim to show the benefits of the application of the combined EMHE and AGS-EKF estimation framework over the standard MHE and EKF under scenarios involving non-Gaussian process uncertainties and measurement noises. For this purpose, the Williams-Otto reactor as a highly nonlinear dynamic system has been considered as a case study. The Williams-Otto reactor is an open-loop unstable system widely used to assess the performance of control and estimation schemes proposed in the literature (Varshney et al., 2019)(Roberts, 1979)(Matias and Le Roux, 2018)(Marchetti, 2013). The Williams-Otto reactor considers the following three reactions:

$$
\begin{aligned}
A + B &\xrightarrow{k_1} C & k_1 &= 1.6599 \times 10^6 e^{-6666.7/TR}\ s^{-1} \\
B + C &\xrightarrow{k_2} P + E & k_2 &= 7.2117 \times 10^8 e^{-8333.3/TR}\ s^{-1} \\
C + P &\xrightarrow{k_3} G & k_3 &= 2.6745 \times 10^{12} e^{-11111/TR}\ s^{-1}
\end{aligned}
$$

where k_1, k_2, and k_3 denote the reaction rate constant for the corresponding reactions. The mechanistic model describing this process is as follows:

$$
\begin{aligned}
W\frac{dX_A}{at} &= F_A - (F_A + F_B)X_A - r_1 \\
W\frac{dX_B}{at} &= F_B - (F_A + F_B)X_B - r_1 - r_2 \\
W\frac{dX_C}{at} &= -(F_A + F_B)X_C + 2r_1 - 2r_2 - r_3 \\
W\frac{dX_E}{at} &= -(F_A + F_B)X_E + r_2 \\
W\frac{dX_G}{at} &= -(F_A + F_B)X_G + 1.5r_3 \\
W\frac{dX_P}{at} &= -(F_A + F_B)X_P + r_2 - 0.5r_3 \\
r_1 &= k_1 X_A X_B W;\ \ r_2 = k_2 X_B X_C W; \\
r_3 &= k_3 X_C X_P W
\end{aligned}
\tag{11}
$$

where $W = 2104.7kg$ represents the mass hold up of the reactor; r_1, r_2, and r_3 are the reaction rates. The reactant A and B represent the main reactants of this set of reactions; P is the main product in the Williams-Otto reactor, whereas E and G are the by-products of this process. The reactor temperature is set to $T_R = 366.05\ K$, and the mass flowrates for the reactant A and B are F_A and F_B, respectively. The nominal value for these flowrates are $F_A = 1.8\frac{kg}{s}$ and $F_B = 6.1\frac{kg}{s}$. Note that F_A is the main disturbance considered for this system, i.e., F_A is subjected to a normally distributed random disturbance with a standard deviation set to $1\ kg/s$. To test the performance of our estimation framework, the distributions of the process uncertainties and measurement noises are assumed to become non-Gaussian following such external disturbances. The Williams-Otto reactor consists of six states, i.e., X_A, X_B, X_C, X_E, X_G, and X_P, each representing the mass fractions of the corresponding chemical components. The nominal steady-state values of the states are reported in Table 1.

Table 1. Nominal values of states for the Williams-Otto reactor

Process Variables	Base case value
X_A	0.090
X_B	0.399
X_C	0.015
X_E	0.141
X_G	0.110
X_P	0.105

The mass fractions for the reactants A and B as well as the final product P can be measured online, i.e., X_A, X_B, and X_P. The linear observability of the process was confirmed, i.e., the linear observability matrix was full-rank at the nominal operating condition indicated in Table 1. In the present work, the Williams-Otto reactor is assumed to be subjected to uncorrelated bimodal process uncertainties and uniform measurement noises. Fig. 2 illustrates the process uncertainty and the measurement noise associated with X_A. The red solid

lines in this figure represents the GMM provided by the EM algorithm, which approximates the true non-Gaussian densities. The rest of process uncertainty variables and measurement noise variables considered in this case study follow similar distributions and are not shown here for brevity. Note that the number of Gaussian components considered in the mixture of each bimodal process uncertainty variable was set to two, i.e., $ngp = 2$ whereas three Gaussian components were considered in the mixture of each uniform process measurement noise, i.e., $ngm = 3$.

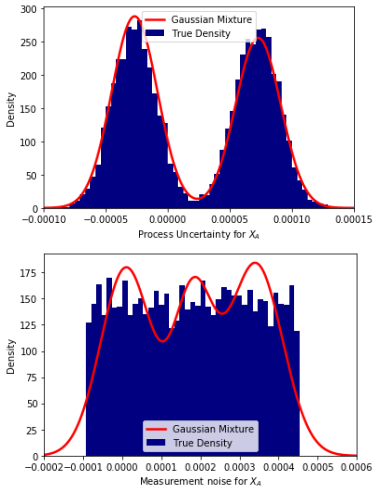

Fig. 2. Histogram for the true non-Gaussian distribution and the GMM approximation of the process uncertainty and measurement noise associated with X_A

As indicated above, the standard MHE (with EKF as the arrival cost) was implemented in this work to compare the performance of the proposed estimation scheme. The standard MHE and EKF assume that process uncertainties and measurement noises follow zero-mean Gaussian distributions. Hence, the standard deviation for each process uncertainty is set to 0.1% of the nominal steady-state value of the corresponding state variable. Likewise, the standard deviation for each measurement noise is set to 0.5% of the nominal steady-state value of their corresponding measurable state. Preliminary tests were performed to select adequate standard deviations for the random variables in favour of achieving a high performance in the standard MHE and EKF schemes. The sampling interval for this process is set to 1 s. The length of estimation horizon for both EMHE and MHE is set to 8 s, which is sufficiently long to show the impact of the arrival cost while requiring a relatively short CPU time to provide the point estimates online. The state estimation scheme shown in Fig. 1 was initialized assuming that the states are +1% away from the nominal true plant states reported in Table 1. As the state variables in the process are mass fractions, lower and upper bounds on the states are set to 0 and 1, respectively.

The computational experiments presented in this section were implemented in Python 3.7 on a computer running Microsoft Windows Server 2016 standard. The computer was equipped with 16 GB RAM and Intel(R) Core(TM) i7-9700K CPU @ 3.60GHz. The mean squared error (MSE) is used as the metric to compare the accuracy offered by each estimation framework. The MSE for the m^{th} state variable obtained by performing the n^{th} state estimation scheme can be estimated as follows:

$$MSE_{x_m}^{(n)} = \frac{1}{t_f}\sum_{k=0}^{t_f}\left(\hat{x}_{k,m}^{(n)} - x_{k,m}\right)^2 \quad (12)$$

where t_f is the final time interval considered in the experiments; $m \in \{1, 2, \ldots, n_x\}$; $n \in \{\text{MHE with EKF, EMHE with AGS} - \text{EKF}\}$; $\hat{x}_{k,m}^{(n)}$ and $x_{k,m}$ are scalars representing the estimated state and the plant output at time interval k, respectively. Note that the plant output is the output of the plant model involving additive process uncertainties, excluding the measurement noises, as shown in Fig. 1.

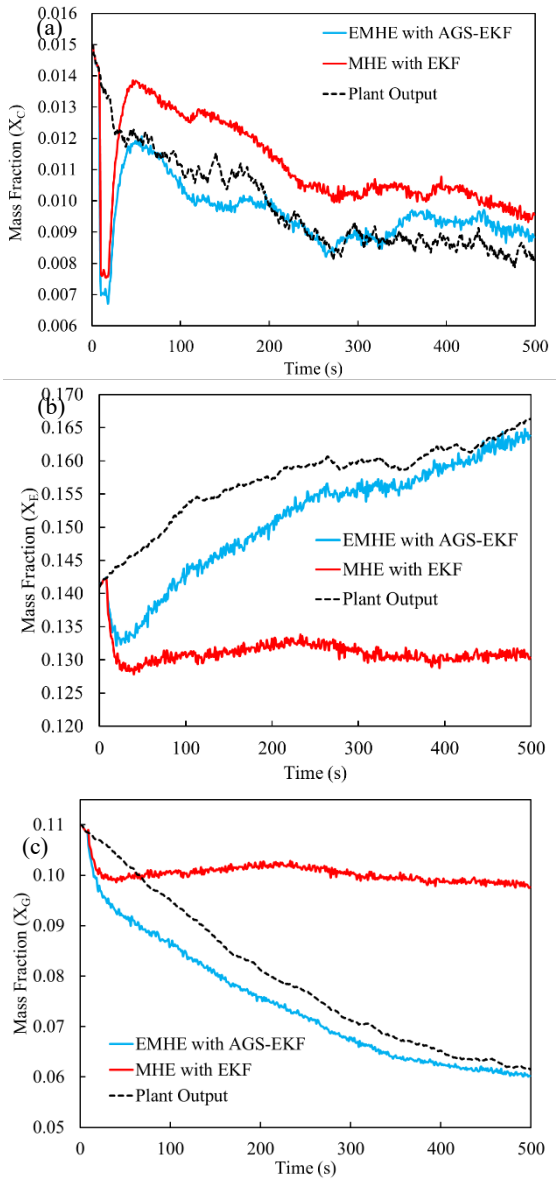

Fig. 3. Estimation provided by EMHE combined with AGS-EKF, and MHE combined with EKF for (a) X_C; (b) X_E; (c) X_G

Fig. 3 shows the estimations provided by EMHE combined with AGS-EKF and the standard MHE combined with EKF for the unknown states X_C, X_E, and X_G. Table 2 reports the error in the estimation of the unknown states using the two estimation frameworks considered in this work. According to

Fig. 3 and Table 2, the estimation accuracy improved significantly when using the EMHE combined with AGS-EKF framework. As shown in Table 2, the estimation error for X_G and X_E was reduced by an order of magnitude when performing EMHE combined with AGS-EKF instead of the standard MHE combined with the standard EKF. Moreover, the estimation provided by EMHE combined with AGS-EKF for X_C is approximately 50% more accurate than that provided by MHE combined with EKF. This is because EMHE and AGS-EKF make use of the main characteristics of the GMMs and are able to capture the non-Gaussianity in the distributions of the process uncertainties and measurement noises. On the other hand, neither standard MHE nor the standard EKF are capable of handling the non-Gaussianity present in the process thus resulting in an overall loss in the estimation performance.

Table 2. MSE for unknown states X_C, X_E, and X_G using different estimation schemes

Estimation method (n)	$MSE_{X_C}^{(n)}$	$MSE_{X_E}^{(n)}$	$MSE_{X_G}^{(n)}$
MHE with EKF	3.18e-6	7.11e-4	6.14e-4
EMHE with AGS-EKF	1.62e-6	4.55e-5	3.51e-5

In the present work, the non-Gaussian distributions of the process uncertainties and measurement noises are assumed to be known *a priori* and remain unchanged throughout the plant operation, which is a common assumption made when performing MHE and EKF (López-Negrete et al., 2011)(Haseltine and Rawlings, 2005). Thus, the proposed EMHE combined with AGS-EKF performs the EM algorithm offline, i.e., approximate the GMMs of the non-Gaussian distributions offline. As a result, the EMHE combined with AGS-EKF framework improves the estimation accuracy for the applications with non-Gaussian process uncertainties and measurement noises at no additional computational costs. For instance, for the current case study, the averaged CPU time required to perform the point estimation is 0.27 s in both EMHE combined with AGS-EKF and MHE combined with EKF. Note that under scenarios that the non-Gaussianity in the densities of these random variables is caused due to scheduled changes in the plant operation, the non-Gaussian distributions of these variables can be updated online as long as the distributions are known *a priori*. In that case, EMHE and AGS-EKF would need to be updated with the revised GMMs, which may result in a relatively small additional computational cost.

4. CONCLUSIONS

This work investigated the performance of the combined EMHE and AGS-EKF as the lead state estimator and the arrival cost estimator, respectively. The key advantage of these framework compared to the standard MHE combined with the standard EKF as the arrival cost estimator is under scenarios where the process uncertainties and measurement noises follow non-Gaussian distributions. The results of performing both the proposed and the standard frameworks for a highly nonlinear open-loop unstable process such as the Williams-Otto reactor revealed that the proposed EMHE combined with AGS-EKF framework was able to improve the estimation at no additional computational costs. Future work considers testing the performance of EMHE combined with AGS-EKF considering scheduled operational changes that lead to changes in the known non-Gaussian distributions of the process uncertainties and measurement noises. Moreover, future work also considers the present EMHE with AGS-EKF framework for scenarios featuring process uncertainties and measurement noises for which their non-Gaussian distributions are not known *a priori* due to unscheduled/sudden changes in the operation. In addition, systems involving structured model uncertainty will also be considered as part of the future work.

REFERENCES

Bilmes, J.A., 1998. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute*, 4, 126.

Dinov, I.D., 2008. Expectation maximization and mixture modeling tutorial.

Haseltine, E.L., Rawlings, J.B., 2005. Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial & engineering chemistry research*, 44, 2451–2460.

López-Negrete, R., Patwardhan, S.C., Biegler, L.T., 2011. Constrained particle filter approach to approximate the arrival cost in moving horizon estimation. *Journal of Process Control*, 21, 909–919.

Marchetti, A.G., 2013. A new dual modifier-adaptation approach for iterative process optimization with inaccurate models. *Computers & Chemical Engineering*, 59, 89–100.

Matias, J.O.A., Le Roux, G.A.C., 2018. Real-time Optimization with persistent parameter adaptation using online parameter estimation. *Journal of Process Control*, 68, 195–204.

Roberts, P.D., 1979. An algorithm for steady-state system optimization and parameter estimation. *International Journal of Systems Science*, 10, 719–734.

Valipour, M., Ricardez-Sandoval, L.A., 2021a. Abridged Gaussian sum extended Kalman filter for nonlinear state estimation under non-Gaussian process uncertainties. *Computers & Chemical Engineering*, 155, 107534.

Valipour, M., Ricardez-Sandoval, L.A., 2021b. Constrained abridged Gaussian sum extended Kalman filter: Constrained nonlinear systems with non-Gaussian noises and uncertainties. *Industrial & Engineering Chemistry Research*, 60 (47), 17110-17127.

Valipour, M., Ricardez-Sandoval, L.A., 2021c. Extended moving horizon estimation for chemical processes under non-Gaussian noises. *AIChE Journal*, 68 (3), e17545. https://doi.org/10.1002/aic.17545

Varshney, D., Bhushan, M., Patwardhan, S.C., 2019. State and parameter estimation using extended Kitanidis Kalman filter. *Journal of Process Control*, 76, 98–111.