Novel Methodologies in State Estimation for Constrained Nonlinear Systems under Non-Gaussian Measurement Noise & Process Uncertainty

by

Mahshad Valipour

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Jinfeng Liu Professor Department of Chemical and Materials Engineering, University of Alberta
Supervisor	Luis A. Ricardez-Sandoval Associate Professor Chemical Engineering, University of Waterloo
Internal Examiner	Peter L. Douglas Professor Chemical Engineering, University of Waterloo
	Hector Budman Professor Chemical Engineering, University of Waterloo
Internal-external Examiner	Ehsan Hashemi Adjunct assistant Professor Mechanical and Mechatronics Engineering, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapters 1, 2 and 9 were entirely written by me. My supervisor, Luis A. Ricardez-Sandoval, provided suggestions and recommendations regarding their contents.

The results in chapter 3 of this thesis have been published in *Control Engineering Practice* on December 18, 2020. Valipour, M., Toffolo, K. M., & Ricardez-Sandoval, L. A. (2021). "State estimation and sensor location for Entrained-Flow Gasification Systems using Kalman Filter." *Control Engineering Practice*, 108, 104702. https://doi.org/10.1016/j.conengprac.2020.104702 My Colleague MSc student Kathryn Toffolo collaborated to perform analytical evaluation of the sensitivity matrices (included in the state-space model) of the gasifier in python. My PhD advisor and I provided solutions to the challenges in the evaluation the sensitivity matrices. The validation of the linearized version of ROM, implementation of the state estimation and sensor location for the gasifier was performed by me under my PhD advisor's supervision. Kathryn Toffolo provided the gasifier model equations (which is shown in Table 3-2 of this thesis) and the nomenclature of the corresponding paper. The rest of that paper was written entirely by myself and edited by my supervisor.

The results in chapter 4 of this thesis have been published in *Computers & Chemical Engineering* on September 12, 2021. Valipour, M., & Ricardez-Sandoval, L. A. (2021). "Abridged Gaussian Sum Extended Kalman Filter for nonlinear state estimation under Non-Gaussian process uncertainties." *Computers & Chemical Engineering*, 107534. https://doi.org/10.1016/j.compchemeng.2021.107534 That paper was written entirely by myself and edited by my supervisor.

The results in chapter 5 of this thesis have been published in *Industrial & Engineering Chemistry Research* on November 17, 2021. Valipour, M., & Ricardez-Sandoval, L. A. (2021). "Constrained Abridged Gaussian Sum Extended Kalman Filter: Constrained nonlinear Systems with non-Gaussian noises and uncertainties." *Industrial & Engineering Chemistry Research*. https://pubs.acs.org/doi/pdf/10.1021/acs.iecr.1c02804 That paper was written entirely by myself and edited by my supervisor.

The results in chapter 6 of this thesis have been published in *Industrial & Engineering Chemistry Research* on February 15, 2021. Valipour, M., & Ricardez-Sandoval, L. A. (2021). "Assessing the Impact of EKF as the Arrival Cost in the Moving Horizon Estimation under Nonlinear Model Predictive Control." *Industrial & Engineering Chemistry Research*, 60(7), 2994-3012.

https://dx.doi.org/10.1021/acs.iecr.0c06095?ref=pdf That paper was written entirely by myself and edited by my supervisor.

The results in chapter 7 of this thesis have been published in *AIChE Journal* on December 8, 2021. Valipour, M., & Ricardez-Sandoval, L. A. (2021). "Extended Moving Horizon Estimation for Chemical Processes under Non-Gaussian Noises." *AIChE Journal*. https://doi.org/10.1002/aic.17545 That paper was written entirely by myself and edited by my supervisor.

The results in chapter 8 of this thesis have been published in *Computers & Chemical Engineering* on Dec 3, 2021. Valipour, M., & Ricardez-Sandoval, L. A. (2021). "A Robust Moving Horizon Estimation under Unknown Distributions of Process or Measurement Noises." *Computers & Chemical Engineering*, 107620. https://doi.org/10.1016/j.compchemeng.2021.107620 That paper was written entirely by myself and edited by my supervisor.

Abstract

Chemical processes often involve scheduled/unscheduled changes in the operating conditions that may lead to non-zero mean non-Gaussian (e.g., uniform, multimodal) process uncertainties and measurement noises. Moreover, the distribution of the variables of a system subjected to process constraints may not often follow Gaussian distributions. It is essential that the state estimation schemes can properly capture the non-Gaussianity in the system to successfully monitor and control chemical plants. Kalman Filter (KF) and its extension, i.e., Extended Kalman Filter (EKF), are well-known model-driven state estimation schemes for unconstrained applications. The present thesis initially performed state estimation using this approach for an unconstrained large-scale gasifier that supports the efficiency and accuracy offered by KF. However, the underlying assumption considered in KF/EKF is that all state variables, input variables, process uncertainties, and measurement noises follow Gaussian distributions. The existing EKF-based approaches that consider constraints on the states and/or non-Gaussian uncertainties and noises requires significantly larger computational costs than those observed in EKF applications. The current research aims to introduce an efficient EKF-based scheme, referred to as *constrained Abridged Gaussian Sum Extended Kalman Filter* (constrained AGS-EKF), that can generalize EKF to perform state estimation for constrained nonlinear applications featuring non-zero mean non-Gaussian distributions. Constrained AGS-EFK uses Gaussian mixture models to approximate the non-Gaussian distributions of the constrained states, process uncertainties, and measurement noises. In the present *abridged Gaussian sum* framework, the main characteristics of the overall Gaussian mixture models are used to represent the distributions of the corresponding non-Gaussian variable. Constrained AGS-EKF includes new modifications in both prior and posterior estimation steps of the standard EKF to capture the non-zero mean distribution of the process uncertainties and measurement noises, respectively. These modified prior and posterior steps require the same computational costs as in EKF. Moreover, an intermediate step is considered in the constrained AGS-EKF framework that explicitly applies the constraints on the priori estimation of the distributions of the states. The additional computational costs to perform this intermediate step is relatively small when compared to the conventional approaches such as Gaussian Sum Filter (GSF). Note that the constrained AGS-EKF performs the modified EKF (consists of modified prior, intermediate, and posterior estimation steps) only once and thus, avoids additional computational costs and biased estimations often observed in GSFs.

Moving Horizon Estimation (MHE) is an optimization-based state estimation approach that provides the optimal estimations of the states. Although MHE increases the required computation costs when compared to EKF, MHE is best known for the constrained applications as it can take into account all the process constraints. This PhD thesis initially provided an error analysis that shows that EKF can provide accurate estimates if it is constantly initialized by a constrained estimation scheme such as MHE (even though EKF is unconstrained state estimator). Despite the benefits provided by MHE for constrained applications, this framework assumes that the distributions the process uncertainties and measurement noises are zero-mean Gaussian, known *a priori*, and remain unchanged throughout the operation, i.e., known time-independent distributions, which may not be accurate set of assumptions for the real-world applications. Performing a set of MHEs (one MHE per each Gaussian component in the mixture model) more likely become computationally taxing and hence, is discouraged. Instead, the abridged Gaussian sum approach introduced in this thesis for AGS-EKF framework can be used to improve the MHE performance for the applications involving non-Gaussian random noises and uncertainties. Thus, a new extended version of MHE, i.e., referred to as *Extended Moving Horizon Estimation* (EMHE), is presented that makes use of the Gaussian mixture models to capture the known time-dependent non-Gaussian distributions of the process uncertainties and measurement noises use of the abridged Gaussian sum approach. This framework updates the Gaussian mixture models to represent the new characteristics of the known time-dependent distribution of noises/uncertainties upon scheduled changes in the process operation. These updates require a relatively small additional CPU time; thus making it an attractive estimation scheme for online applications in chemical engineering. Similar to the standard MHE and despite the accuracy and efficiency offered by the EMHE scheme, the application of EMHE is limited to the scenarios where the changes in the distribution of noises and uncertainties are known *a priori*. However, the knowledge of the distributions of measurement noises or process uncertainties may not be available *a priori* if any unscheduled operating changes occur during the plant operation. Motivated by this aspect, a novel robust version of MHE, referred to as *Robust Moving Horizon Estimation* (RMHE), is introduced that improves the robustness and accuracy of the estimation by modelling online the unknown distributions of the measurement noises or process uncertainties. The RMHE problem involves additional constraints and decision variables than the standard MHE and EMHE problems to provide optimal Gaussian mixture models that represent the unknown distributions of the random noises or uncertainties along with the optimal estimated states. The additional constraints in the RMHE problem does not considerably increase the required computational costs that that needed in the standard MHE and consequently, both the present RMHE

and the standard MHE require somewhat similar CPU time on average to provide the point estimates. The methodologies developed through this PhD thesis offers efficient MHE-based and EKF-based frameworks that significantly improve the performance of these state estimation schemes for practical chemical engineering applications.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Luis A. Ricardez-Sandoval for his endless help and support, patience, and encouragement. I greatly appreciate his invaluable inputs and advice during my PhD study that made this work happen and brought up the researcher one inside me.

I would like to extend my sincere thanks to my examining committee for taking time, reviewing my thesis, and providing valuable comments: Prof. Jinfeng Liu, Prof. Peter L. Douglas, Prof. Hector Budman, Prof. Ehsan Hashemi.

I wish to acknowledge the financial assistance provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Waterloo.

I want to express my heartfelt thanks to my dear friends Kimia Aghasadeghi, Sepideh Karkouti, Arian Koranlou, and Babak Feizifar for their friendship and immense support whenever I most needed during my journey towards this milestone. I want to extend my appreciation to my friends and colleges from the University of Waterloo who made my PhD memorable. Dear “Luis’ awesome research group”, many thanks for all the encouragements, coffees, and camping lifestyle that we shared in office. In particular, I wish to thank Dr. Mina Rafiei-Shishavan, Dr. Kavitha G. Menon, Dr. Yue Yuan, Dr. Grigoriy Kimaev, Kimia Entezari, Oscar Palma Flores, Yael Valdez-Navarro, Donovan Chaffart, Ilse Cerrillo-Briones, Huabei You, Gabriel Patron, Kathryn Toffolo, Sepehr Ghadami, Dr. Amir Mowla, Dr. Rasool Nasseri, Dr. Sepehr Khazraei, Dr. Javad Noroozi, and Dr. Navid Bizmark.

Last but certainly not least, I would like to express my wholehearted appreciation to my parents and dearest siblings, Shahab, Shaghayegh, and Shafagh for their continuous love, encouragement, and patience all the way from thirteen thousand kilometers away that got me through this milestone. My jaanan, I need a thesis long letter to begin to thank you. Hopefully, your patience battery is back on full-charged over past few months as I have new journeys to begin.

Dedication

To my Mom, the strongest, bravest, smartest and yet the kindest and the most compassionate human being I have ever known.

Jaan e del mamani! you are my mom, my dad after dad, my best friend after Nowrouz 1372, my ultimate mentor of all time, my rock, and sure my most spoiled girl. No word can possibly carry the weight of love and gratitude that I have for you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AGS-EKF	Abridged Gaussian Sum Extended Kalman Filter
AGSF	Adaptive Gaussian Sum Filter
AC	Arrival Cost
CFD	Computational Fluid Dynamics
C-PF	Constrained Particle Filter
CSTR	Continuous Stirred-Tank Reactors
CV	Controlled Variable
DSZ	Down Stream Zone
EKF	Extended Kalman Filter
EM	Expectation-Maximization algorithm
EMHE	Extended Moving Horizon Estimation
ERZ	External Recirculation Zone
GSF	Gaussian Sum Filter
GS-PF	Gaussian Sum Particle Filter
GS-UKF	Gaussian Sum Unscented Kalman Filter
HIPS	High Impact Polystyrene Process
IGCC	Integrated Gasification Combined Cycle
JEZ	Jet Expansion Zone
KF	Kalman Filter
MHE	Moving Horizon Estimation
MSE	Mean Squared Error
MV	Manipulated Variable
NMPC	Nonlinear Model Predictive Control
NLP	Nonlinear Optimization Problem
PDF	Probability Density Function
PF	Particle Filter
PFR	Plug-Flow Reactor
RMHE	Robust Moving Horizon Estimation
ROM	Reduced-Order Model
SSE	Sum of Squared Error
UKF	Unscented Kalman Filter
WTP	Wastewater Treatment Plant

List of Symbols

c	Index for a specific measurement noise in the vector of measurement noises
$c_{pg}, c_{pp}, c_{pslag}$	Heat capacity of gas phase, particle phase, and slag (J/kg/K)
c_s	Oxygen specific saturation
c_w	Dissolved oxygen concentration(mg/L)
$c_{w_{sp}}$	Dissolved oxygen set-point in the reactor (mg/L)
$\boldsymbol{e}$	Estimation error
$e_{k\|k}$	Estimation error in EKF at time interval k
d	dimension of the variable $\boldsymbol{\varpi}$
f	Process model function
f_d	Fraction of death biomass
f_k	Turbine speed
g	Function that summarizes the set of inequality constraints present in the process
h	Measurement model function
h_p	Convection coefficient of the particle phase (W/m^2/K)
i	Index for the Gaussian component in a general Gaussian mixture model
is, ip, im	Indexes for the Gaussian component in the Gaussian mixture model of states, process uncertainties, and measurement noises, respectively
j	Index of time interval within the estimation horizon
k	Index for the current time interval
$k_{g,eff}, k_{slag}$	Effective thermal conductivity of gas phase and thermal conductivity of slag (W/m/K)
k_r, k_1, k_2, k_3	Reaction constant
k_s	Saturation constant
k_d	Biomass death rate
k_c	Specific Cellular activity
k_{la}	Oxygen transfer into the water constant
k_{01}	Oxygen demand constant
l	Index for a specific process uncertainty in the vector of uncertainties
$lr_d,\ lr_b, lr_r$	Depth of the first, second and bottom layers in the decanter (m)
m	Subscript to show the index of the state variable in the state vector $\boldsymbol{x}$
m_p, m_{slag}	Mass of particle phase and slag (kg)

$m'_{slagging}$	Mass flux of slag (kg/m^2/s)
mw_{is}	Membership weight in EM algorithm
n	Superscript to denote the n^{th} estimation method
ng	Number of Gaussian components presented in a general Gaussian mixture model
ngr	Number of Gaussian components presented in the Gaussian mixture model of RMHE
ngs, ngp, ngm	Number of Gaussian components presented in the Gaussian mixture model of states, each process uncertainty, each measurement noise
ngt	Total number of EKFs present in the set of GSF
n_x, n_y, n_w, n_v	Number of state variables, measurements, process uncertainties, measurement noises
p	Probability density function
p_A, p_B	Partial pressures of species *A* and *B*
p^l, p^u	Lower and upper bounds on the state variables in the Gas-phase reactor
q, q_i, q_p, q_r, q_1, q_2	Flow rate (m^3/hr)
r	Index for the Gaussian component in the Gaussian mixture model
$r_{gasifier}$	Radius of gasifier (m)
s	Index for the samples drawn from non-Gaussian distribution
s_i	Inlet organic substrate concentration(mg/L)
s_{ir}	organic substrate concentrations entering the bioreactor (mg/L)
s_w	Organic substrate concentrations in the bioreactor (mg/L)
$s_{w_{sp}}$	Substrate set-point in the reactor (mg/L)
s_w^{up}	upper bund (saturation limit) on the substrate concentration
t	time
t_f	Final time interval considered in the numerical experiments
u_g, u_p	Velocity of gas phase and particle phase (m/s)
$\boldsymbol{u}_k$	Inputs to system at time interval *k*
$\boldsymbol{u}^l, \boldsymbol{u}^u$	Bounds on the inputs
vsd, vsb, vsr	Rate of settling for the activated sludge at first, second and bottom layers in the decanter
$\boldsymbol{v}_k$	Random measurement noise signals at time interval *k*
$\boldsymbol{w}_k$	Random process uncertainties at time interval *k*
$\boldsymbol{wp}$	Matrix of process uncertainties for all realizations in the C-PF algorithm

$\boldsymbol{wgt}$	Weight function in C-PF
$\hat{x}_{k,m}^{(n)}$	estimation provided by n^{th} estimation scheme for m^{th} state variable at k^{th} sampling time
$x_d\,, x_b\,, x_r$	Biomass concentrations at first, second and bottom layers in the decanter (mg/L)
x_{in}	Inlet biomass concentration(mg/L)
x_{ir}	biomass concentrations entering the bioreactor (mg/L)
x_w	Biomass concentrations in the bioreactor (mg/L)
$\boldsymbol{x}_0$	nominal (initial) condition
$\boldsymbol{x}_k$	State variables at time interval k
$\widehat{\boldsymbol{x}}_{k\|k}$	Posterior estimation of states at current time interval k
$\widehat{\boldsymbol{x}}_{k+1\|k+1}^{\boldsymbol{GSF}}$	Point estimates provided by GSF at time interval $k+1$
$\widehat{\boldsymbol{x}}_{k+1\|k+1}^{i}$	Posterior estimation of states provided by i^{th} EKF in the set of GSF
$\overline{\boldsymbol{x}}_{k-N}$	Expected value of the approximated posterior distribution of the states at time $k-N$
$\boldsymbol{x}^l$, $\boldsymbol{x}^u$	bounds on the states
$\boldsymbol{x'}_z$	State of the system at the z^{th} time interval within the time horizon L
$\boldsymbol{y}_k$	Available measurement at time interval k
$\boldsymbol{y}'_z$	Output of the system at the z^{th} time interval within the time horizon L
$\boldsymbol{y}_j$	Measurement of the system at the j^{th} time interval within the time horizon N
z	Index of time interval within the prediction horizon
A_d	cross-sectional area of the decanter (m^2)
A_H	Heat transfer area (m^2)
A	Amplitude of the sinusoidal changes in the inlet flowrate (m^3/hr)
$\boldsymbol{A}_k$	Jacobian matrix at time interval k
$\boldsymbol{B}_k$	Sensitivity matrix at time interval k
C	Length of control horizon
C_{A_1},C_{A_2},C_{A_3},C_{B_1},C_{B_2},C_{B_3}	Concentration of the chemical components in the corresponding vessel ($kmol/m^3$)
$C_i, C_s, C_m, C_b, C_r, C_{br}$	Concentration of initiator, solvent, monomer, butadiene, radicals, branched radicals in HIPS (mol/L)
$C_i^s, C_m^s, C_b^s, C_r^s, C_{br}^s$	Scaling factor (mol/L)
C_m^f	Monomer feed concentration (mol/L)
C_b^f	Butadiene feed concentration (mol/L)

C_{ps_w}	Water heat capacity (J/kg.K)
C_{ps}	Monomer heat capacity (J/kg.K)
C_{total}	Total concentration (mole/m^3)
$\boldsymbol{C}_k$	Sensitivity matrix at time interval k
$\boldsymbol{Cov}_k^{\boldsymbol{S-M}}$	Cross-covariance matrix at time interval k
$D_{g,eff}$	Effective diffusivity of the gas phase (m^2/s)
$\boldsymbol{D}_k$	Sensitivity matrix at time interval k
EM	Expectation-maximization function
F_A and F_B	Mass flowrates of the reactants A and B
$F'_{g\rightarrow wall}$	Volumetric force between the gas and the wall (N/m^3)
$F'_{g\rightarrow p}$	Volumetric force between the gas and the particle phase (N/m^3)
$\boldsymbol{G}_k$	Sensitivity matrix at time interval k
H_g, H_p	Enthalpy of the gas phase and the particle phase (J/kg)
$\boldsymbol{H}_k$	Sensitivity matrix at time interval k
$\boldsymbol{K}_k$	Kalman gain at time interval k
K_{i0}, K_{i1}, K_{i2}, K_{i3}, K_d, K_{fb}, K_{fs}, K_P, K_t	Kinetic coefficients in HIPS model
L	Length of prediction horizon in NMPC
M_{slag}	Slag mass flow (kg/s)
M_{solid}	Mass flow of solid phase (kg/s)
$\boldsymbol{M}^{\boldsymbol{GM}}$	Covariance matrix of the Gaussian mixture model of the non-Gaussian variable $\boldsymbol{\varpi}$
$\boldsymbol{M}^i$	Covariance matrix of the i^{th} Gaussian component in the mixture
M_n	Number of average molecular weight (kmol/m^3)
M_w	Weight average molecular weight (kg^2/kmol.m^3)
N	Length of estimation horizon in the MHE-based estimation schemes
NP	Number of particles in C-PF
NS	Number of samples drawn from the original non-Gaussian distribution
N_p	Number density of the particle phase (m^{-3})
$\boldsymbol{P}_{k-N\|k-N}$	Covariance matrix of the approximated posterior distribution of the states at time $k-N$
$\boldsymbol{P}_{k\|k}^i$	Covariance of posterior estimation of states given by i^{th} EKF in the set at time k

$\boldsymbol{P}_{k+1\|k}^{is}$	Covariance matrix for the is^{th} Gaussian component in the multivariate Gaussian mixture model corresponded to the constrained prior estimation of states
$\boldsymbol{P}_{k+1\|k}^{\boldsymbol{GM}}$	Covariance matrix of the constrained prior estimation of the states' distribution
Q	Outlet flow rate (L/s)
$Q'_{conv\rightarrow g}$	Heat flux from gas convection (W/m^2)
$Q'_{convg\rightarrow p}$	Heat flux from gas convection to particle phase (W/m^2)
$Q'_{convg\rightarrow slag}$	Heat flux from gas convection to slag phase (W/m^2)
$Q'_{convslag\rightarrow surrounding}$	Heat flux from slag convection to surroundings (W/m^2)
$\boldsymbol{Q}$	Covariance matrix of process uncertainties in the standard MHE and EKF
$\boldsymbol{Q}_k$	Covariance matrix of process uncertainties at time interval k
$\boldsymbol{Q}_k^{\boldsymbol{GM}}$	Covariance matrix of the process uncertainty Gaussian mixture at the current time interval k
$\boldsymbol{Q}_k^r$	Covariance matrix of the r^{th} Gaussian component in the process uncertainty Gaussian mixture at the current time interval k in RMHE framework
$\boldsymbol{Q_{in}}$	Weights on the manipulated variables
$\boldsymbol{Q_{out}}$	Weights on the controlled variables
Q_{cw}	Cooling water flow rate (L/s)
Q_i	Initiator flow rate (L/s)
Q^s, Q_{cw}^s, Q_i^s	Scaling factor (L/s)
R	Ideal gas constant cal/(mol.K)
R_{het}, R_{hom}	Heterogeneous and homogeneous reactions
$\boldsymbol{R}$	Covariance matrix of measurement noises in the standard MHE and EKF
$\boldsymbol{R}_k$	Covariance matrix of measurement noises at time interval k
$\boldsymbol{R}_k^{\boldsymbol{GM}}$	Covariance matrix of the measurement noise Gaussian mixture at the current time interval k
$\boldsymbol{R}_k^r$	Covariance matrix of the r^{th} Gaussian component in the measurement noise Gaussian mixture at the current time interval k in RMHE framework
$\boldsymbol{S}_k$	Estimation of the measurement's covariance at time interval k
T_1, T_2, T_3	Temperature inside the CSTR1, CSTR2, and the separator (K)
T_j	Jacket temperature (K)
T^f	Feed steam temperature (K)
T_j^f	Inlet water temperature (K)
T_g, T_p, T_{slag}	Temperature of gas phase, particle phase, and slag (K)

T^s	Scaling factor (K)
T_R	Reactor temperature (K)
V	Reactor volume (L)
V_c	Jacket volume (L)
V_r	The volume of the aeration tank (m^3)
W	Mass hold up
$X_A, X_B, X_C, X_E, X_G, X_P$	Mass fractions of the corresponding chemical components
Greek Symbols	
α^i	Corresponding weight for the i^{th} Gaussian component of the Gaussian mixture
$\alpha s, \alpha m, \alpha p$	Weights assigned to corresponding component in the Gaussian mixture model of states, measurement noise, and process uncertainty variable
$\boldsymbol{\alpha m}_k^r$	Corresponding weight for the r^{th} Gaussian component of the noise Gaussian mixture
$\boldsymbol{\alpha p}_k^r$	Corresponding weight for the r^{th} Gaussian component of the uncertainty Gaussian mixture
β^i	Weight (probability) that the i^{th} EKF in the set of GSF presents the posterior estimation of the states at each time interval
δ_{slag}	Slag thickness (m)
$\varepsilon_g, \varepsilon_p$	Volume fraction of gas and particle phases
$\boldsymbol{\eta}^{\boldsymbol{GM}}$	Mean-vector of the overall Gaussian mixture model of the non-Gaussian variable
$\boldsymbol{\eta}^i$	Mean-vector of the i^{th} Gaussian component in the mixture
λ_1	First moment of dead polymer molecular weight (kg/m^3)
$\boldsymbol{\mu}$	Mean-vector of the non-Gaussian process uncertainties
$\boldsymbol{\mu}_k^{\boldsymbol{GM}}$	Mean-vector of the process uncertainty Gaussian mixture at time interval k
$\boldsymbol{\mu}_k^r$	Mean-vector of the r^{th} Gaussian component in the process uncertainties mixture
μ_r^0	Zeroth moment dead polymer
μ_{slag}	Slag's viscosity (Pa s)
μ_b^0	Zeroth moment of the chain-length distribution for growing polymer chains with a terminal butadiene unit
$\mu_r^{0^s}, \mu_b^{0^s}$	Scaling factor (mol/L)
$\boldsymbol{\xi}_k$	Non-zero mean Gaussian random process uncertainties at time interval k
ρ_{cw}	Water density (Kg/L)
ρ_g, ρ_p	Density of gas and particle phases (kg/m^3)

ρ_s	Monomer density (Kg/L)
σ	Standard deviation
$\boldsymbol{\tau}$	Mean-vector of the non-Gaussian measurement noises
$\boldsymbol{\tau}_k^{\boldsymbol{GM}}$	Mean-vector of the measurement noises Gaussian mixture at time interval k
$\boldsymbol{\tau}_k^r$	Mean-vector of the r^{th} Gaussian component in the measurement noises mixture
φ_{k-N}	Arrival cost term at time $k-N$
$\boldsymbol{\varpi}_k$	Non-zero mean non-Gaussian variable at time interval k
Λ	Weight correction term in GSF algorithm
$\boldsymbol{\Pi}$	Error factor
$\Upsilon_{\varrho}, \Upsilon_{mn}, \Upsilon_{pu}$	Binary parameters in the RMHE framework
$\boldsymbol{\psi}$	Process uncertainty factor

Chapter 1
Introduction

Online measurement technologies are not often available for key state variables that are critical for online monitoring and control of major industrial applications. Thus, it is essential to develop state estimation methods to provide online estimation of the unmeasured state variables for the purpose of online monitoring and control. Hence, developing techniques that can improve the performance (i.e., accuracy and computational efficiency) of the state estimation schemes for both model-driven and data-driven approaches is a topic of interest that is receiving increasing attention in the field of process control and estimation. Kalman filters and moving horizon estimation (MHE) are the cornerstones of model-driven estimation schemes due to their efficiency and accuracy. High performance of MHE requires an accurate and efficient approximation of the arrival cost (AC) parameters as well as an appropriate assumption of the specifications of the process uncertainties and measurement noises. Similarly, proper initialization of the prior distribution of the process variables (states, process uncertainties, measurement noises) and the evaluation of the sensitivity (Jacobian) matrix of the process are the key steps in the Kalman filter and the extended Kalman filter (EKF) methods. Thus, the focus of this thesis is to develop novel methodologies to improve the EKF and MHE performances. The details on these methodologies are presented below.

Kalman Filters are standard state estimation methods for linear and nonlinear dynamic systems. Extended Kalman filter[1,2], unscented Kalman filter (UKF)[3,4,5], the sampling-based group of filters referred to as particle filters (PFs)[6,7,8] have been widely used to deal with the nonlinearity of the systems. Both UKF and PFs require additional computational costs when compared to EKF, which makes EKF a superior estimation scheme for large-scale applications. However, EKF requires online estimation of the sensitivity (Jacobian) matrix of the process, which may be a complex and computationally taxing task for such applications. This mathematical complexity along with the observability challenges and plant-model mismatch commonly associated with large-scale systems limits the application of KF/EKF for such processes. To the author's knowledge, studiefs exploring the performance of KF/EKF for industrial processes featuring more than 200 states is quite limited in the open literature. Moreover, Kalman filter assumes that all the state variables, input variables, process uncertainties and the measurement noises are normally distributed and can be described using a Gaussian probability density function (PDF). However, most real-world state estimation problems involve nonlinear systems with

inequality constraints on the states, and process uncertainties (and measurement noises) that often follow non-zero mean non-Gaussian distributions, i.e., multi-modal, uniform, Gamma, etc[9,10,11]. Current approaches to improve Kalman filter performance for applications considering constrained states or zero-mean non-Gaussian noises and uncertainties often require online solution of optimization problems, implementation of sampling-based approaches or implementation of multiple EKFs (Gaussian sum filters (GSF)), which come at the cost of a significant increase in the computational effort[10,12,13,14,15,16,17]. Moreover, to the author's knowledge, a study that considers EKF applications featuring non-zero mean non-Gaussian noises and uncertainties is not available in the literature. Therefore, an efficient EKF-based estimation scheme is required to fill-in this gap and improve the estimation of EKF for general chemical engineering applications, which are often subject to constraints and a general class of non-Gaussian uncertainties/noises[18,19,20,21,22].

For the constrained applications, MHE is a well-known state estimation scheme that solves an optimization problem involving all the process constraints[23,24,25,26]. The MHE problem aims to seek for the optimal estimated states by minimizing the historical errors present in the process. These errors consist of process uncertainties and measurement noises over a finite horizon previous sampling intervals (N) in addition to *arrival cost*. Arrival cost is a term presents in the objective function of the MHE problem that summarizes the past information of the process (from the initial time until $(N\text{-}1)^{th}$ sampling interval) that are discarded from the estimation horizon. A standard state estimation method is required to approximate the AC parameters online in order to achieve high performance in MHE. In addition to the popularity of EKF as a state estimation scheme, it is also a standard approach to approximate the AC parameters in the MHE framework. The accuracy and efficiency of the AC estimator is a key to achieve high MHE performance. Initially, the main motivation of using EKF as the AC estimator in the MHE framework was the estimation accuracy and efficiency offered by EKF. However, the Gaussian assumption in EKF cannot hold under the scenarios involving constraints in the process model, which raises the question on the use of EKF as the AC estimator for constrained applications. Previous studies showed through small-scale examples that EKF coupled with MHE results in an inappropriate estimation, instability, or even failure in the estimation scheme. On the other hand, studies have shown that the main reason of the divergence in EKF is the lack of a proper initialization[27]; in fact, EKF is expected to fail under a poor initialization in the distribution of the states[27,28]. Given that EKF as the AC estimator is initialized by the state estimates provided by MHE, the poor initialization would no longer be a concern. To the extent of the author's knowledge, no mathematical analysis has been presented that disqualify EKF as the AC estimator for the constrained

nonlinear applications. Motivated by this, investigating the performance of EKF coupled with MHE and the impact of this estimation approach on the closed-loop system is a gap in knowledge that is worth exploring as EKF efficiency makes this approach favorable for the purpose of online control.

Moreover, the MHE problem has been designed to accommodate different distributions for the process uncertainties, measurement noises, and state variables. For instance, previous studies have illustrated the ability of MHE in handling bounded Gaussian noises [29,30,31]. However, the standard MHE formulation assumes that process uncertainties and measurement noises follow zero-mean Gaussian distributions[32,33]. In the standard MHE scheme, these distributions are known *a priori* and remain unchanged throughout the process, while chemical processes often involve distributions of the process uncertainties and measurement noises that may change throughout the operation due to changes in the operating conditions, i.e., sessional changes or switching to different product grades. Note that the non-Gaussianity and changes in the noises/uncertainties distributions can be known *a priori* under the scheduled events, i.e., a scheduled replacement of the measurement devices. To date, the performance of MHE under those conditions has not been investigated. Therefore, it is critical to improve MHE scheme to capture both non-Gaussianity and expected (scheduled) changes in the plant that may lead to changes in the distributions of the process uncertainty and measurement noises. On the other hand, these changes in the distribution of noises/uncertainties may happen due to the unscheduled events such as an unexpected measurement device failure. A more realistic scenario that is likely to happen in practice is that the distribution of the noises/uncertainties cannot be known *a priori* in the estimation scheme. Hence, it is also essential to develop an MHE-based framework that is robust against unexpected noises or uncertainties (with unknown non-Gaussian distributions) that may occur due to unexpected changes such as sensor failures or sudden changes in the operation of the plant.

1.1 Research objectives:

To address the issues mentioned above, the purpose of this thesis is to provide new insights on the capabilities of EKF and MHE and present novel techniques and developments that can improve the performance of these widely used estimation techniques in Chemical Engineering. To pursue this goal, the current PhD study focuses on the following research objectives:

- Investigate the benefits and limitations of KF/EKF for large-scale systems involving a large number of states.
- Develop an efficient scheme to improve EKF performance for general processes featuring non-Gaussian states (i.e., constrained states) and non-zero mean non-Gaussian noises/ uncertainties.

- Assess the impact of EKF as the AC estimator in the MHE framework under open-loop and closed-loop operation.
- Develop a new MHE formulation that takes into account known non-Gaussian distributions of the uncertainties and noises, which can be updated online based on scheduled changes in the distribution of these random variables.
- Develop a new robust framework to improve the MHE performance under scenarios considering unexpected noises (e.g., due to sudden measurement device failure) or process uncertainties (e.g., due to unscheduled changes in the operating conditions) that follow unknown non-Gaussian probability density functions.

1.2 Expected contributions

The work conducted in this PhD thesis is expected to result in the following contributions:

- Provide insights on the KF/EKF abilities, limitations, and challenges in the state estimation and sensor location for an actual entrained-flow gasification system consisting of a large number of states (i.e., >200 states) under practical scenarios, i.e., plant-model mismatch, load-following, additive uncertainties.
- Present a new modification to the standard EKF formulation that makes EKF capable of capturing non-zero mean Gaussian process uncertainty and measurement noise. The key contribution is that the new EKF developed in this thesis would require the same computational costs as the standard EKF while improving the accuracy in the estimation under the conditions mentioned above.
- Present a novel state estimation framework, referred to as *constrained Abridged Gaussian Sum Extended Kalman Filter* (constrained AGS-EKF) to generalize EKF for constrained nonlinear systems under general non-Gaussian noises and uncertainties. AGSF-EKF not only reduces the computational costs incurred when using conventional estimation methods such as GSF and PF, but it also avoids biased estimations that may occur in the traditional approaches.
- Provide new insights on the capabilities of EKF as the AC estimator (when coupled with MHE and NMPC) through conducting an error analysis as well as performing comprehensive implementations.
- Present a novel extended version of MHE, referred to as *Extended Moving Horizon Estimation* (EMHE), that relaxes the zero-mean Gaussian assumption considered for the process uncertainties and measurement noises in the standard MHE. The key novelty in the proposed EMHE formulation

is that it requires the same computational costs as the standard MHE while improving the estimation under the operating scenarios describe before.

- An EMHE-based estimation scheme is introduced to update online the characteristics of the non-Gaussian distributions of the noises/uncertainties that may change due to scheduled changes in the operating conditions. Gaussian mixture models are introduced to the proposed estimation scheme to represent an accurate approximation of the known non-Gaussian densities of these random variables online, which requires a relatively small additional CPU time and hence the efficiency of the scheme. Note that EMHE focuses on applications where the distributions of the process uncertainties and measurement noises are known *a priori* over the entire operating horizon.
- Present a novel robust estimation scheme, referred to as *Robust Moving Horizon Estimation* (RMHE) that uses EMHE and the Gaussian mixture model to improve the robustness of the MHE estimation in the presence of unexpected measurement noises or process uncertainties with unknown densities. The RMHE problem involves additional constraints and decision variables to find the optimal Gaussian mixture model to describe the unknown densities of the unexpected noises/uncertainties. The additional computational costs in the RMHE framework (compared to MHE) is not significant, hence is efficient.

In summary, the research conducted in this thesis presents new estimation technologies that can be implemented to consider more realistic scenarios that occur during operation in chemical plants.

1.3 Structure of the Thesis

This PhD thesis is organized as follows:

Chapter 2 presents a literature review on the application of EKF for the large-scale system, current constrained EKFs, GSFs, MHE and the common AC estimators used in the MHE framework, MHE for applications featuring time-dependent non-Gaussian measurement noises/ process uncertainties, and robust MHE methods. The gaps in knowledge that motivate this research are explicitly discussed in this chapter.

Chapter 3 presents the application of KF/EKF for an actual entrained-flow gasification system. The challenges, features, and limitations of this estimation scheme for large-scale systems are highlighted and discussed in detail. The work presented in this chapter has been published in Control Engineering Practice[34].

Chapter 4 introduces abridged Gaussian sum extended Kalman filter (AGS-EKF) for applications involving non-Gaussian process uncertainties. A new modification in the prior estimation step of EKF

has been presented in this chapter that is critical to capture non-zero mean non-Gaussian process uncertainties. The work has been published in Computers & Chemical Engineering[35].

Chapter 5 presents a novel constrained AGS-EKF that improves estimation for a general constrained nonlinear application featuring non-Gaussian process uncertainties and measurement noises. This chapter provides the new modifications considered in the posterior estimation step (to capture non-zero mean non-Gaussian measurement noises) and the intermediate estimation step that explicitly takes into account the constraints on the state variables. This work has been published in Industrial & Engineering Chemistry Research[36].

Chapter 6 provides a new insight on EKF capabilities as an AC estimator when coupled with MHE in both open-loop and closed-loop. This chapter compares the performance of open-loop and closed-loop in the case of using different AC estimators through extensive simulations and various practical scenarios. An error analysis is presented to make this study more comprehensive. This work has been published in the Industrial & Engineering Chemistry Research[37].

Chapter 7 presents the proposed Extended MHE (EMHE) that accommodates the known non-zero mean non-Gaussian distributions of the process uncertainties and measurement noises to the estimation scheme. The derivation of EMHE as well as the recursive estimation scheme to capture the scheduled changes in the distribution of the measurement noises/ process uncertainties online are discussed in detail. The underlying assumption in EMHE is that the non-Gaussian distribution of the noises/uncertainties are known *a priori* to the estimation scheme during the plant operation, which is a common assumption made when implementing MHE. This work has been published in AIChE Journal[38].

Chapter 8 introduces the Robust MHE (RMHE) formulation along with the underlying assumptions in this framework. The RMHE problem focuses on the scenarios that the non-Gaussian distributions of the process uncertainties or measurement noises are unknown *a priori* due to the unscheduled operational changes. The additional constraints in the MHE framework to model the unknown distributions of noises (or uncertainties), the performance, challenges, and limitations of the proposed RMHE are discussed in this chapter. This work has been published in Computers & Chemical Engineering[39].

Chapter 9 presents the concluding remarks and contributions achieved by the studies conducted in this research. This chapter also provides recommendations for potential future work in this area.

Chapter 2
Literature Review

Over the past decades, several new developments have been performed to the field of state estimation to improve the accuracy and efficiency of these approaches for real-world applications under practical but challenging scenarios that occur during plant operation. Although these modifications have been conducted for both model-driven and data-driven estimation schemes, the main focus of this thesis is on the core model-driven estimation schemes, i.e., KF/EKF and MHE, and their extensions. Numerous efforts involving accurate mathematical developments have been presented in the literature with the aim of improving the efficiency, accuracy, and robustness of both EKF and MHE frameworks. Each of these modifications have offered different levels of improvement in the conventional estimation schemes for the application at hand. Nevertheless, there are still gaps that limits the online monitoring and control of chemical plants under certain operating conditions. This chapter presents a literature review and the knowledge gaps that serve the motivation for the studies performed in this thesis. This chapter is organized as follows: section 2.1 provides a literature review on KF and EKF. Section 2.2 provides an overview on Gaussian Sum Filter (GSF) that is an EKF-based approach to deal with non-Gaussianity present in the process. An overview on MHE is presented in section 2.3. A summary of this chapter is presented in the end.

2.1 An Overview on Kalman Filters

Kalman filter (KF) is a standard approach to perform state estimation for unconstrained linear applications (linear process and measurement models) with zero-mean Gaussian process uncertainties and measurement noises[40,41,42]. The optimal estimated states provided by KF for such processes also follow Gaussian distributions. KF uses the linearized version of the true process models, i.e., the state-space model, to provide *a priori* estimation of the distribution of the states. The sensitivity matrices present in the state-space model describe the relation between states and states, states and inputs, states and process uncertainties, states and measurements, and measurements and measurement noises. These sensitivity matrices remain unchanged throughout the process. Extended Kalman filter (EKF) is widely used as an efficient extended version of KF to perform state estimation for cases where the process and/or measurement models are described by unconstrained nonlinear functions while the process uncertainties and measurement noises follow zero-mean Gaussian distributions. To capture the nonlinearity of the dynamic systems, EKF updates the sensitivity matrices of the state-space model

based on the posterior estimations of the states at the last time interval[43,44]. The formal formulation of the KF/EKF is presented in chapter 3 (Equations (3-4) and (3-5)). Note that these updates in EKF increase the required computational costs to perform the point estimates when compared to KF. Nevertheless, the efficiency and accuracy offered by KF and EKF approaches has motivated the application of these scheme to perform state estimation for the unconstrained large-scale applications. An overview on this topic is presented in section 2.1.1. Moreover, section 2.1.2 provides an overview on the benefits and limitations offered by the existing developments in EKF aimed at improving this estimator for applications involving constrained states and/or non-Gaussian process uncertainties and measurement noises.

2.1.1 KF/EKF and large-scale applications:

In principle, unmeasured states can be inferred from historical data and causal relations between the states, which can be predicted from a dynamic process model and the available process measurements. Thus, an accurate dynamic process model that can predict the behaviour of the actual plant, combined with an adequate number of hardware sensors, are essential to achieve an accurate estimation of the unmeasured states. Efficient frameworks that investigate the minimum number of sensors required and their potential locations to assess the observability of the system are available[45,46]. As mentioned earlier, KF as well as its extensions, such as EKF have become standard tools in the industry and academia to estimate key process states that cannot be measured online[47]. For instance, literature presents studies performing KF/EKF based state estimations for large scale systems, e.g., a version of Tennessee Eastman process[48,49], a reactive distillation process[50], a real large-scale agriculture field[51], heavy oil hydro processing reactors[52], etc.

In the case of large and intensive systems, state estimation turns out to be quite challenging due to high process nonlinearities and the limited number of sensors available for process measurement. The energy sector is a fairly good example for such a case because processes in this sector operate at harsh conditions and with a very limited number of sensors. Thus, assessing the observability of intensive systems such as a gasifier become challenging since they often involve a significantly large variety of states that cannot be measured online. In addition, the application of certain estimation methods such as EKF require online estimation of the sensitivity (Jacobian) matrix of the process, which may be computationally taxing, particularly for large-scale and complex applications. Consequently, the application of state estimation techniques in the energy sector is quite limited, e.g.,[53,52,54,55]. The main

interest of this research is to perform state estimation for a large-scale gasifier as a part of integrated gasification combined cycle (IGCC) plant. Thus, an overview on this topic is presented next.

The interest in using coal-fired power generation systems, which provide almost 40% of worldwide power generation, has increased significantly over the last decade[56]. Nowadays, due to environmental concerns, process intensification, computer aided materials design and optimization of energy systems have become topics of interest to further improve the efficiency of these systems[57,58,59,60,61,62]. Available technologies for integrated gasification combined cycle (IGCC) are classified based on the available gasifier configurations since this is the main unit in IGCC. Further details on IGCC can be found elsewhere[63]. Gasification is a thermochemical process that converts heavy liquid and solid fuels into gaseous fuels that contain useful heating value[64]. The operating temperature (ranging from 400 °C to above 2000 °C)[65,66] and pressure (from atmospheric pressure up to 70 bar)[67,68,69], as well as the flow geometry and oxidation agent, can vary in different gasification units. Entrained-flow gasifiers are the most common commercial gasifiers due to their short residence time (up to 5 s) and high throughput and conversion when compared to other available technologies[70]. Finely pulverized coal is injected into the entrained flow gasifier to ensure a high carbon conversion. An undesirable side product in the gasification process is ash that forms a slag layer throughout the inner wall of the gasifier. Therefore, it is critical to keep the operating temperature and pressure higher than they would be in ash slagging conditions to allow for molten ash removal from the gasifier[71]. The operating temperature and pressure are up to and over 2000 °C and 15 bar, respectively, which may vary based on the feedstock composition, the flow regime within the gasifier, and the extent of mixing[65,70].

Based on the above, to avoid a loss of plant efficiency due to the formation of this slag, the slag thickness within the gasifier can be properly controlled by introducing limestone to the gasifier and by accurately monitoring and controlling the peak temperature within the gasifier. Therefore, online monitoring of the temperature and the slag thickness inside of the gasifier is essential to keep the efficiency of this unit at acceptable levels, particularly under operating conditions that often occur in gasification, such as load-following and co-firing[66]. However, measuring the slag thickness in real time is not feasible as the slag is porous and thin and reacts with refractory[72]. Moreover, very few types of thermocouples can provide online measurements for the extreme temperatures reached during operation[73,74]. Neither the usual thermocouples nor advanced temperature measurement devices, i.e., laser-based temperature sensors, can provide reliable online measurements of the corresponding wall temperatures at the critical locations within the gasifier for extended periods of operation. Moreover,

most of these thermocouples are still expensive thus imposing a high tax on the budget[75,76]. In addition, several techniques such as gas chromatography are commonly used to quantify different gas concentrations, i.e., CO, CO_2, H_2, and CH_4 in industrial plants; however, none of these techniques are able to provide measurements of the species concentration in real-time. Thus, online estimation of the gasifier's temperatures, slag thicknesses and species concentrations become critical for the safe and successful operation of this essential process in IGCC.

Based on above, performing state estimation plays an important role in online monitoring and control of IGCC process. However, the application of state estimation techniques on IGCC plants (or some sections of IGCC) is still an open issue. Bhattacharyya and et al.[77] employed an adaptive KF method for state estimation of an acid gas removal unit, which is a part of an IGCC plant with CO_2 capture. The results showed that their proposed adaptive KF method improves the accuracy of the estimation. Carrasco and et al.[78] applied EKF-based neural network training for the char reduction zone of a solid fuel gasification process. EKF was used to estimate the six weights in the neural network while using the output of the neural network model as the measurements in EKF framework. Their study considered 6 states variables and did not include the wall temperature or the slag thickness in the estimation algorithm. In addition, that study assumed that the online measurements were available for all the states. Huang and et al. [79] used KF to estimate the distribution of the wall temperature for an entrained-flow gasifier at steady state and for different extents of slag penetration. In addition, a dynamic estimation of the slag thickness while considering a fixed inlet gas temperature was provided in that study. The highest temperature reported in that work is 1400°C. In that study, a mechanistic thermal model and a slag penetration model were proposed and validated using experimental data. States related to the reaction kinetics of the gasification process were not considered in that study. To this date, the application of a practical conventional method such as KF and EKF for an actual pilot-scale coal-fire gasification unit is lacking from the literature, which is worth exploring.

2.1.2 EKF under Non-Gaussian densities

State estimators aim to provide adequate estimates that can be used to properly initialize a control system such as model predictive controller (MPC); a poor initialization of any model-based controller may lead to a loss performance or unstable operation[37,80,81,82]. Although Kalman filters are well-known for unconstrained applications, these methods do not take into account the process constraints. This is because one of the key underlying assumptions in Kalman filters is that all variables in the system (i.e., states, inputs, process uncertainties, measurement noises) are assumed to follow Gaussian distributions.

However, this assumption may not hold in the presence of constraints on the states, process uncertainties and measurement noise signals. Rao et al. showed that the lack of knowledge of the process constraints in Kalman filter leads to biased/infeasible estimates[83]. Motivated by the crucial role of process constraints and given this shortcoming in conventional state estimation frameworks, the optimization-based state estimation schemes such as MHE, as well as the constrained versions of Kalman filters and sampling-based approaches (e.g., constrained PF) have been developed to improve the estimation performance and consequently, the control performance[83,84]. MHE as an optimization-based estimation method takes into account all the constraints in the dynamic system, which is the main reason of its popularity in the field of estimation[84]. However, this method requires a relatively high computational costs and CPU time when compared to EKF. In addition, MHE considers that process uncertainties and measurement noises follow zero-mean Gaussian distributions. A detailed discussion on MHE is presented in section 2.3. A large number of studies have been conducted to develop Kalman filter-based approaches to improve the performance of this filter for real-world applications featuring constrained states and non-Gaussian noises/uncertainties. These modifications include a constrained version of well-known Kalman filters, i.e., constrained EKF[14,85], constrained UKF[86,87]and constrained PF[13,37,88] that have been proposed in the literature to overcome the nonlinearity caused by the bounded states and other process feasibility constraints in the system. In general, EKFs are the most computational efficient methods among the aforementioned filters. However, most of the existing constrained-EKF approaches often require to solve an additional/internal optimization problem to meet the constraints and bounds on the state variables[14]. This increases both the complexity and computational costs of the filters, which may limit the application of these approaches for large-scale constrained nonlinear chemical systems. Prakash et al. proposed an optimization-free constrained EKF framework that considers a truncated distribution of the state to explicitly take into account the constraints on the prior and posterior distributions of the estimated states[89]. The existing constrained-EKFs and constrained-UKFs assume that the process uncertainty and measurement noise follow zero-mean Gaussian distributions[14,89,90]; whereas, PFs have shown better performance for the case of non-Gaussian distributed noises and uncertainties, with the main drawback of large computational demands[91,92].

In most of the engineering and science applications, dynamic processes are nonlinear, while the process uncertainties and measurement noise signals are represented by bounded Gaussian distributions or assumed to follow a non-Gaussian distribution, e.g., multi-modal, Gamma, etc. Several studies have proposed recursive algorithms to overcome the challenges of estimation for such systems[10,12,15]. Those

studies provided an appropriate estimation of the states when tested for engineering applications involving a very limited number of states, inputs and measurements. However, those algorithms involve second-order gradient information (second-order Taylor series expansion)[10,12] and solving convex optimization problems[10]. Thus, the required computational costs to deal with the convex optimization problem as well as difficulties to evaluate the second-order information for large-scale systems may make these algorithms computationally intractable for practical processes.

Previous studies have considered non-Gaussian distributions for the process uncertainties and measurement noises[16,17]. However, the Gaussian mixture models in their approach consist of zero-mean Gaussian distributions, which may not always be an adequate assumption, e.g., when the non-Gaussianity is described by multi-modal, non-symmetric bounded Gaussian, uniform distributions. If the process uncertainty/measurement noise present in a process follows a non-zero mean distribution, the estimation scheme would require a modification to adopt non-zero mean process uncertainties/measurement noises to keep consistency with the process model and be able to provide accurate estimations. To the author's knowledge, such modifications are absent from the literature. In addition to the existing developments on EKF discussed above, Gaussian sum filter (GSF) is an EKF-based framework introduced in the literature that aims to capture the non-Gaussianity in the process variables. As this approach shares some features that are relevant to the novel methodologies presented in this thesis, an overview on the features and limitations of this estimation scheme is presented next.

2.2 An Overview on Gaussian Sum Filter (GSF) and Non-Gaussian Applications

Gaussian sum filters (GSF) have been introduced by Sorenson et al.[93] as a state estimation approach that withdraws the Gaussian assumption for the distributions of states, measurement noises, and process uncertainties. The main idea in GSF is that any non-Gaussian distribution can be approximated to a *Gaussian mixture model*, which consists of a proper finite number of Gaussian distributions, also referred to as *Gaussian components*[93]. Therefore, GSF performs a set of EKFs to provide the point estimates, i.e., one EKF to project ahead (in time) each of the Gaussian components in the Gaussian mixture model of the original non-Gaussian distribution[94]. In principle, GSF can be applied for a wide range of practical applications, regardless of the level of nonlinearity/non-Gaussianity of the probability density functions that describe the states, measurements, process uncertainties and measurement noises[93]. In practice, GSF and its extensions (i.e., Gaussian sum unscented Kalman filter (GS-UKF) and Gaussian sum particle filter (GS-PF)) have been proposed only for constrained linear/nonlinear

systems involving a relatively small number of state variables where the constraints on the states were not active[17,95,96,97]. Moreover, the application of GSF in the field of chemical engineering is quite limited. To the author's knowledge, only two studies have considered performing state estimation using the Gaussian sum-based filters for chemical processes[97,98]. In those studies, the process uncertainties and the measurement noise signals are assumed to follow a Gaussian distribution. In general, the performance of GSFs for applications featuring non-zero mean non-Gaussian process uncertainties and measurement noises has not been reported by the literature. This limitation is more likely due to a few drawbacks associated with this method; for instance,

I. The number of required EKFs in the set of GSF increases exponentially as the number of variables with non-Gaussian distributions increases[99], which impacts the computational costs associated with this method.
II. Performing individual EKFs based on solely one Gaussian component may result in biased estimations and/or divergence of EKF[35]. For instance, it can drive the system to its feasibility limits and eventually lead to instability or failure in the estimation scheme[35].
III. Previous studies have shown that the Gaussian components in the Gaussian mixture models may exhibit large covariance matrices, which lead to an inaccurate estimation in EKFs (UKFs/PFs) in the set of GSF (GS-UKF/GS-PF)[100,101,102].

To circumvent the limitation highlighted in item (III) above, Psiaki[103] considered an upper limit for the covariance of the Gaussian components. This upper limit is user-defined, which can make its application challenging in practice. Their proposed approach also considers a re-approximation of the Gaussian mixture model in the case that the covariance of the Gaussian components violate their corresponding upper bounds[103]. As a result, that approach improves the GSF performance by eliminating the limitation outlined in item III. However, the limitations described in items (I) and (II) above remained unsolved. Moreover, in that framework, the constraints on the states were not involved in the re-approximation of the Gaussian mixture models[103].

Regarding the limitations indicated in items (I) and (II) above, no practical solution has been presented in the literature to address these issues for a general application involving constraints on the states of non-zero mean non-Gaussian measurement noises. Bhushan et al.[97] presented a GSF-based approach (*Unscented Gaussian sum filter* (UGSF)) for constrained systems that requires the same computational costs as that needed by UKF, which is yet significantly higher than that required in EKF. Although UGSF was presented as a general method to deal with the constrained state estimation, their case studies

were tested under scenarios where the process was operating far from the feasibility limits, i.e., states were not close to their bounds. Moreover, one of the underlying assumptions in UFSG is that the noises and uncertainties follow a Gaussian distribution[97]. Therefore, there is a great incentive to develop tools that can address these aspects in EKF/KF estimation.

2.3 An Overview on Moving Horizon Estimation

Online optimization is becoming a standard tool for solving control and estimation problems due to the computational power that is becoming available for industrial applications. Contrary to standard EKF, optimization-based state estimation techniques are capable of solving complex constrained problems online. The full information estimator is an optimization-based state estimation framework that can takes into account the process constraints[104,105]. The full information estimator solves an optimization problem subject to equality and inequality process constraints and considering all past measurements from the initial time up to the current time step at each time interval. The goal of this optimization problem is to find the optimal estimated states so that it minimizes the errors in the process and measurement models. The successful solution to a full information estimator optimization problem is able to satisfy all the constraints and bounds on the estimated states and disturbances, which distinguishes full information estimators from conventional Kalman filters. However, since this method solves the problem over a growing time horizon, it is in general a computationally intractable approach to consider[106]. To overcome this obstacle, moving horizon estimation (MHE) has been proposed as an alternative algorithm to the full information problem for online applications[107].

MHE can be formulated as a constrained nonlinear optimization problem that contrary to the full information, takes into account a fixed-length horizon of N past measurements and dynamic system updates. All the discarded previous measurements which are not included in the current horizon are summarized in a term referred to as the *arrival cost* (AC), which appears as a penalty term in the objective function. The moving horizon slides with time while partially accounting for the AC term[28]. Figure 2-1 shows the difference between the full information problem and MHE. Note that the formal optimization formulation of the MHE is presented in chapter 6 (Equation (6-1)). MHE aims to address the main limitation of the full information estimator, i.e. computational costs; hence, this method has received attention in recent years[108,109,110,111,112,113,114,115].

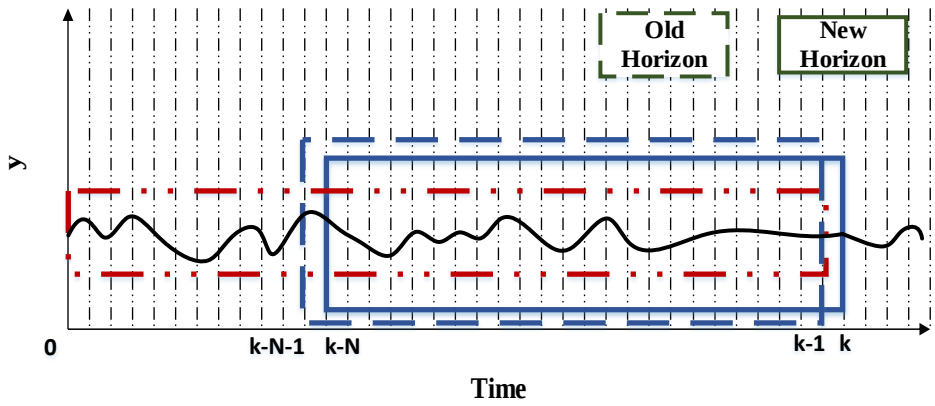


Figure 2-1: Full information (red windows) problem vs MHE (blue windows)

There are many factors that need to be considered when choosing the length of horizon (N), e.g., computational budget, system observability and model accuracy. Higher estimation accuracy may be obtained by either long horizon lengths or determining accurate AC approximations[108,116]. An MHE with a long horizon length approximates to the full information estimation problem, and therefore can easily become computationally intractable, particularly for large-scale online applications. Therefore, one of the open issues in applying MHE method is how to estimate the AC to summarize comprehensively and accurately all the past information of the system.

In addition to the importance of AC in the MHE performance, it is essential to provide proper models to the MHE formulation to accurately describe the distribution of the measurement noises and process uncertainties. Similar to EKF, both full information estimator and MHE consider that process uncertainties and measurement noises follow zero-mean Gaussian distributions, which may not be always a valid assumption. This thesis is focused on the techniques that can improve the MHE performance under cases involving non-Gaussian noises and uncertainties.

Based on above, an inappropriate initialization of the three terms in the objective function of the MHE problem, i.e., AC, summation of the measurement noises, and summation of the process uncertainties, may result in biased state estimation. Thus, an overview on the impact of AC in the MHE performance is presented next. The gaps that motivated the research performed in this thesis (presented in chapter 6) is outlined in section 2.3.1. Later, a literature review concerning the performance of MHE-based schemes while taking into account the non-Gaussianity in the distribution of noises and uncertainties is presented in section 2.3.2.

2.3.1 Arrival Cost Estimator in the Moving Horizon Estimation coupled with Nonlinear Model Predictive Control

MHE is a well-known optimization-based estimation for the constrained application, as discussed above. The stability of MHE, the required CPU time, and the accuracy of its solution is highly dependent on the AC since it is needed to identify the initial states of the plant in the moving horizon framework. Explicit AC expressions can only be specified for unconstrained linear systems; thus, for most of the constrained nonlinear dynamic systems, it is essential to find ways to approximate the AC[117]. A poor approximation of the AC leads to the necessity of choosing a long estimation horizon to reduce the impact of AC on the MHE framework at the expense of additional computational costs[108]. However, the computational cost of full-information optimization problems becomes significant due to the increase in the estimation horizon and can limit the application of MHE as an online estimator for highly nonlinear large-scale systems[118]. This is particularly critical for MHE applications in closed-loop control systems. Hence, the AC approximation has a critical role to reduce the computational costs in the MHE estimation and thus improve the performance of advanced process control systems such as nonlinear model predictive control (NMPC) [119]. Note that NMPC as an optimization problem relies on the initialization provided by MHE. A proper initialization of the NMPC optimization problem is key to maintain closed-loop operation on target[120]. In addition to the requirement of the approximation of AC for the purpose of initializing the MHE, and consequently the NMPC problem, an AC estimator requires to receive a prior estimation of the states as well as the current control actions computed by the NMPC, respectively. That is, there exists an interaction between MHE and AC as well as NMPC and AC. Thus, the effect of AC approximation on the MHE performance, and therefore on the NMPC performance, becomes a topic of interest that has not been widely investigated in the literature.

Efforts to improve AC approximations can be found in the literature. The most common approaches for approximating the AC are KF and EKF[30]. For instance, Rao et al. employed KF and Kalman Smoother methods to update the covariance matrix in the AC for a linear system that follows a Gaussian distribution[121]. As for unconstrained nonlinear systems, MHE may provide stable state estimates that are closer to the true states when EKF is used as the AC estimation method[106]. According to the literature, the stability conditions can be derived when an unconstrained nonlinear MHE framework is considered[106,122]. Since the errors in the AC approximation are propagated throughout the estimation horizon, EKF may exhibit instability. Moreover, the main assumption of EKF is that the conditional probability density of the states is normally distributed. Although process uncertainties and states of

the system may follow a normal distribution, the prior and posterior distributions may not be Gaussian in the presence of bounds on the states or other process variables (e.g., saturation limits in the manipulated variables). Furthermore, if a normally distributed random variable propagates through a nonlinear model, its distribution may not necessarily remain Gaussian[88]. On the other hand, according to the previous studies, EKF divergence is mainly due to the poor initialization in the distribution of the states[27,28]. Moreover, previous studies have shown that Kalman filter framework is an adequate estimation scheme for large-scale applications if proper initial guesses of states and a reliable mechanistic model are available[34,123].Qu and Hahn[117] employed the UKF method to update only the covariance matrix in the AC estimation. Although the performance of MHE improves slightly, the issue of inconsistency between the unconstrained UKF to provide the AC estimation for constrained MHE still remains an open issue. To circumvent this problem, the use of a constrained version of UKF, PF and cell filtering (CF) estimators has been proposed[124]. However, constrained UKF often results in suboptimal solutions at best. On the other hand, PF has been identified as a standard nonlinear state estimation method that can explicitly handle constrained state estimation problems (non-Gaussian state-space models) and makes PF more attractive and practical than EKF for industrial applications[125]. Lang et al.[126] developed a constrained PF (C-PF) that truncates or adjusts the suitable constrained distribution (prior or likelihood) to satisfy the process constraints, thus ensuring that the posterior distribution also satisfies constraints. However, that study did not solve the challenges of choosing the importance function in C-PF[127]. On the other hand, reports involving small-scale case studies revealed that among all the methods used in that study, CF is the most adequate choice for estimating the AC. However, the high computational cost associated with CF limits its application for industrial and/or complex systems[124]. Despite the limitations mentioned above, C-PF, and in particular EKF, remain as the most widely used filters to deal with nonlinear constrained state estimation problems.

In the literature, only very few studies have assessed the impact of AC in the MHE framework[88,124]. For instance, López-Negrete et al.[88] analyzed the benefits of using different sampling-based methods to estimate the AC parameters. Since particles are being propagated through the nonlinear system, the normally distributed assumption of the state estimation error in the EKF estimator can be dropped, as well as the difficulty associated with the calculation of the Jacobian matrix at each time interval. Those studies have only considered relatively simple case studies, e.g., an isothermal batch reactor (i.e., includes two states and one measurement), a non-isothermal Continuously Stirred Tank Reactor (CSTR) (i.e., involves three states and one measurement)[88] and a synthetic univariate time-varying

nonlinear system (i.e., involves one state with a single measurement)[124]. In addition, studies addressing the effect of using different AC estimation methods for chemical systems under conditions that may rise during operation has not been reported, e.g., bounds on the process uncertainty. Moreover, studies assessing the impact of AC on the quality of the performance in closed-loop using advanced controllers such as NMPC are not available.

2.3.2 Moving Horizon Estimation under Non-Gaussian Noises and Uncertainties

As mentioned above, chemical plants often involve process uncertainties and measurement noises that follow non-zero mean non-Gaussian distributions. The studies on the MHE performance under the non-Gaussian noises/uncertainties can be categorized into two major types, i.e., non-Gaussian noises/uncertainties with *Type I*-known priori distributions and with *Type II*- unknown distributions. A literature review for each type is presented next.

2.3.2.1 *Type I*: MHE under noise/uncertainty with known non-Gaussian distribution

Contrary to the underlying assumption in the standard MHE formulation[32,33], random measurement noises and process uncertainties associated with the chemical processes often follow non-zero mean non-Gaussian distributions[9,10,11]. This inconsistency between the MHE framework and the actual process operating conditions may lead to an inaccurate estimation. Note that if the non-Gaussianity in the distribution of the noises is only due to bounds on the noises, i.e., noises follow truncated zero-mean Gaussian distribution, standard MHE can accurately estimate the states because the Gaussian distribution assumption holds with meaningful statistics and therefore, the MHE formulation would take into account the bounds on the noises as constraints[128,29,30,31,37,129]. However, if the process uncertainties and measurement noises belong to a non-Gaussian density, i.e., uniform, multi-modal, etc., the Gaussian distribution assumption in the standard MHE framework is no longer valid. Therefore, a modification in the MHE framework is needed to overcome the challenges caused by presence of such random variables in the system. Studies investigating the effect of non-Gaussian process uncertainties and measurement noises (i.e., multi-modal, uniform, Gama, etc. distribution) on the MHE performance are scarce[130,131]. Bae and et al. showed that for humanoid state estimation involving non-Gaussian noises, standard MHE provides more accurate estimates than KF[130]. Yin and et al. presented a l-2 regularized MHE that involves an additional weighted l2-norm term to improve the estimation accuracy against outlier measurement noises[131]. That study showed a reduction in the

estimation error using their proposed MHE when compared to KF for an application involving non-Gaussian noises, because of the robustness offered by the additional term. Nevertheless, their approach did not take into account the non-Gaussianity of the noises[131].

Moreover, the distribution of the process uncertainties and measurement noises may change over time from a known arbitrary distribution to another distribution, i.e., from a zero-mean Gaussian distribution to a non-zero mean non-Gaussian distribution. This may happen due to the scheduled changes in the operational conditions and set-points. Therefore, it is essential to make the MHE framework flexible to adjust the state estimation accordingly to the current distribution (i.e., the shape and the modelling characteristics of the distribution) of the random variables present in the process. Although the state-space model used in the standard MHE and other state estimation algorithms considers process uncertainties and measurement noises that randomly changes at each sampling time, these random variables are assumed to belong to known distributions that remain unchanged throughout the process. For the remainder of this work, the distribution that remains unchanged during operation is referred to as *time-independent distribution*; likewise, the distribution that changes due to changes in the operation is referred to as *time-dependent distribution*. Studies in state estimation that consider scenarios with time-dependent distribution for the process uncertainty and measurement noise are limited[21]. Xu and et al. proposed a Fixed-point Iteration Gaussian Sum Filter to improve the estimation for the case of time-dependent non-zero mean non-Gaussian distribution of measurement noise[21]. The performance of their approach has not been investigated for cases involving process uncertainties with time-dependent non-Gaussian distributions. Moreover, similar to conventional Gaussian sum filter-based approaches, the approach proposed by Xu requires to perform several parallel filters[21], i.e., one for each of the Gaussian component in the Gaussian mixture, which increase the computational load exponentially depending on the level of non-Gaussianity of the system. To the author's knowledge, no study has considered online adaptation on the distribution of the process uncertainty (and measurement noise) within the context of the MHE framework. Also, MHE-based estimation schemes that explicitly take into account the non-Gaussian process uncertainties and measurement noises have not been reported in the literature.

In general, non-Gaussian probability density functions can be approximated to a Gaussian mixture model, i.e., a summation of multiple Gaussian distributions. Thus, in theory, a similar procedure as in GSF may help to improve the MHE performance under non-Gaussian noises and uncertainties. While performing Gaussian sum to the MHE framework can improve the estimation accuracy for the applications involving the non-Gaussian distributed variables, this combination can become

computationally demanding. That is, performing a set of MHEs to evaluate the point estimates means solving a set of optimization problems thus making this approach computationally taxing for the purpose of online estimation and control. To the author's knowledge, no study has considered the application of Gaussian sum (or Gaussian mixture models) in the context of the MHE framework.

2.3.2.2 *Type II*: Robust MHE under unexpected noise/uncertainty with unknown distribution

One of the underlying assumptions in the standard Moving Horizon Estimation (MHE) framework is that the process uncertainties and measurement noises can be described by zero-mean Gaussian distributions where the covariance matrix of their corresponding distributions are known *a priori* [3233]. However, the process uncertainties' densities may not follow their ordinary distribution during operation [1819]; for instance, a sudden extreme weather condition may impose unexpected random process uncertainties and start to follow a non-Gaussian distribution (e.g., bimodal distribution). Likewise, the unexpected measurement noises caused by occasional failures of the measurement devices may lead to gross measurement noises, i.e., data outliers, drifting data, etc. [202122]. The distribution of the corresponding process uncertainty (or measurement noise) may become non-Gaussian upon these expected random variables interfere the operation, up until the effect of these operational changes vanishes. These aspects often arising in chemical plant operations motivate the need to develop robust MHE schemes.

The conventional robust MHE schemes presented in literature have focused on replacing the ℓ^2-norm function by either the *Fair*, *Lorentzian*, *Logistic*, *Huber*, and *Welsch* functions[132,133,134,135,136] to increase the robustness of the state estimation against gross measurement noises[137,138]. Despite these changes in the MHE framework, the performance of those schemes for applications involving unexpected process uncertainties have not been investigated. Another MHE-based robust scheme has been presented in[115] that can handle outliers in the measurement noise for linear applications. That robust MHE framework performs a set of optimization problems based on a set of least-squares cost functions[115]. Each cost function considers that one of the measurements present in the horizon is contaminated by an outlier and needs to be left out from the state estimation[115]. The lowest cost function is retained to provide the point estimate and to propagate to the next time interval[115]. That approach requires a significant amount of computational effort to determine the point estimates when compared to the standard MHE, which may limit its application for nonlinear industrial-scale applications. Ji and et al.[113] developed an MHE-based framework that involves an additional max term to the standard MHE objective function to improve the robustness in the estimation. That work assumed that the process uncertainties and

measurement noises are bounded[113]. Moreover, robust MHE frameworks have been presented in[30,139] that involve an auxiliary nonlinear observer combined with the standard MHE framework (under the assumption of bounded process uncertainty). The deterministic observer is used to provide a confidence region and the reference estimates that limit the robust MHE to find optimal estimations within these confidence regions for nonlinear applications involving bounded uncertainties and noises[30,139,140]. Each of these robust MHE approaches have shown different level of success in improving the MHE performance. Nevertheless, the bounded distribution assumption on the random variables may not necessarily hold for real-world applications. In addition, the current robust MHE schemes share the same assumption that the distribution of the process uncertainties and measurement noises are known *a priori*, i.e., Gaussian zero-mean distribution with a known covariance matrix[30,137,138,115,113,139]. Hence, frameworks involving non-Gaussian noises or uncertainties with unknown densities are of practical interest in the context of state estimation. Gaussian mixture models are well-known for their ability to approximate arbitrary non-Gaussian distributions. Despite this attractive feature, the application of Gaussian mixture models to improve the accuracy and robustness in MHE estimation has not been investigated to the author's knowledge.

2.4 Summary

In summary, this chapter presented a review on KF, EKF, and MHE, which are the core state estimation schemes for the research performed in this PhD thesis. Moreover, an overview on GSF was presented to provide an insight on this method. As discussed in this chapter, KF/EKF are efficient approaches to perform state estimation for unconstrained large-scale systems. Due to the observability and stability challenges associated with the energy sector, the performance of these state estimation schemes for unconstrained large-scale applications such as entrained-flow gasifier has not been evaluated to date. As for the applications featuring constrained states, non-Gaussian process uncertainties, and non-Gaussian noises, an efficient (optimization-free and sampling-free) EKF-based framework that can provide appropriate estimations of the states is still an open issue in the field of state estimation.

Based on the review conducted on MHE, the key to success when performing state estimation using this method is to provide accurate approximations of the arrival cost (AC), process uncertainties, and measurement noises. Based on the multiple contradicting studies presented in the literature on the reason as to why EKF tend to fail under certain scenarios (e.g., process constraints). An error analysis on the EKF performance under different estimation frameworks can provide new insights on the capabilities of EKF as a single state estimator or when it is combined with a constrained estimation

scheme such as MHE. With regards to the role of process uncertainties and measurement noises in the MHE problem, the literature review revealed that this is a fairly new area in the context of MHE that is worth exploring. To the best of author's knowledge, a development on MHE that makes this scheme suitable for the applications featuring non-Gaussian or/and unknown or/and time-dependent distributions of the process uncertainties or measurement noises are absent in the literature.

Chapter 3

State estimation and sensor location for Entrained-Flow Gasification Systems using Kalman Filter

Performing KF/EKF for large-scale systems poses challenges such as observability and computational complexities in the evaluation of the Jacobian matrix. Consequently, exploring the performance of this estimation scheme under practical scenarios occurring in large-scale power plants such as IGCC systems is a knowledge gap in the literature. This chapter presents the application of a practical method such as KF and EKF for an actual pilot-scale coal-fire gasification unit. The gasifier model employed in this study is a dynamic reduced-order model (ROM) of an actual pilot-scale entrained-flow gasifier unit, which includes 479 states describing the reaction kinetics and the mass, heat and momentum transport phenomena taking place inside this unit. To the author's knowledge, the estimation of such a large-scale intensive system in the energy sector that involves that many states across a gasifier has not been attempted before, hence the novelty of this research. The estimation has been performed under different scenarios involving a different arrangement of the sensors available for online estimation of the states. In addition, the estimation for the pilot-scale gasification unit was also tested in the presence of additive uncertainty in the prior estimation, plant-model mismatch, and load-following, the latter being a typical operation often performed in IGCC gasification units.

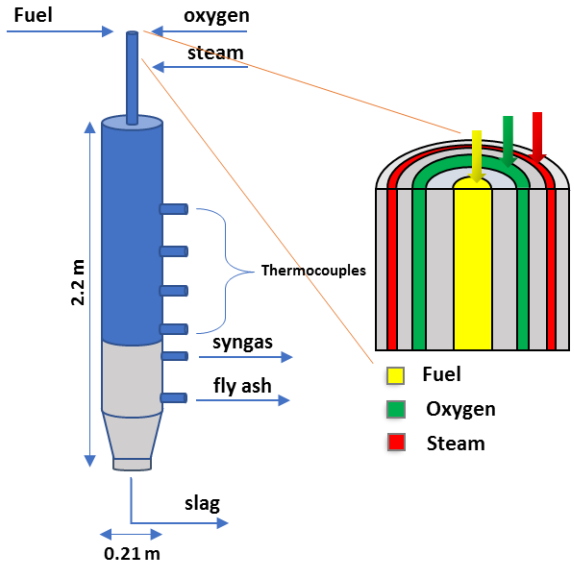
This chapter is organized as follows: the next section elaborates on the gasifier unit and the model used in this study to simulate the transient operation of this unit. Section 3.2 provides an overview of the Kalman filter estimation. Section 3.3 presents the results, detailed discussions, and underlying assumptions for each of the scenarios considered in this study. A summary of this chapter is presented in section 3.4.

3.1 Entrained-flow Slagging Gasifier

The gasification unit is responsible for the transformation of the solid fuels into the so-called *syngas* gaseous fuel. As shown in Figure 3-1(a), solid fuel, oxygen, and steam enter into the gasifier. At the top of the gasifier, the fuel, steam and oxygen mix with each other, which leads to vaporization of the moisture content (H_2O) and devolatilization of the volatile components within the solid fuel (e.g., hydrogen H_2, tar, methane CH_4, ethane C_2H_6, etc.). This process is called drying and pyrolysis (see Figure 3-1(b)). The solid residue remaining after this process is known as char, which contains carbon and ash. The major part of the ash content of the fuel melts and makes a liquid slag layer on the gasifier's

refractory, which needs to be removed from the gasifier to avoid losing efficiency in the unit. The char and the combustible species within the volatile product react with oxygen. Figure 3-1(b) presents the homogeneous reactions within the unit (see *Volatile and Combustion reaction* section in Figure 3-1(b)). This set of reactions aim to provide the required energy for the four heterogeneous char gasification reactions to produce the syngas as well as the heat needed for the drying and pyrolysis. Moreover, two sulfur (homogeneous) reactions convert the sulfur content of the fuel into hydrogen sulfide (H_2S). The product species of all the reactions are listed in Figure 3-1(b), including syngas (i.e., combination of CO and H_2), which represents the main product used for power production. As shown in Figure 3-1(a), the syngas together with the remaining ash and molten slag represent the main outlet products obtained from this unit[141]. Table 3-1 presents the fuel composition and the nominal flowrate of these inlet streams as well as some of the main design parameters of the pilot-scale gasifier considered in this study.

(a)



(b)

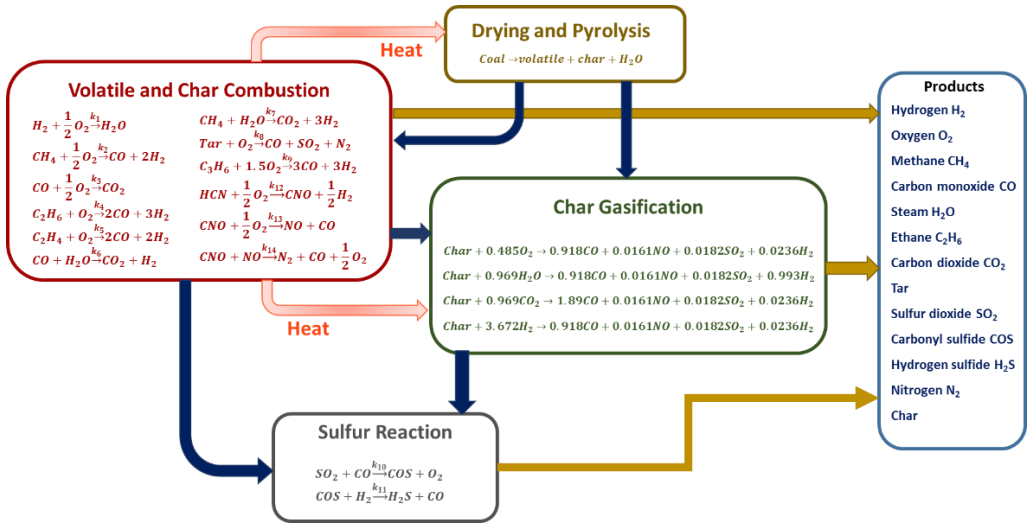


Figure 3-1: (a) Entrained-flow gasifier[142]; (b) reaction system of gasification process[141]

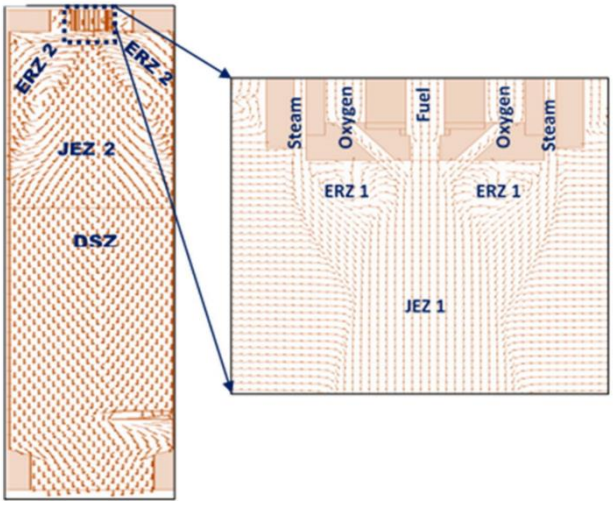
Table 3-1: Operating conditions and design specifications for the pilot-scale gasifier[72,142]

Parameter	Value
Inner diameter	0.21 m
Length	2.20 m
Operating pressure	15 bar
Inlet H_2 mole fraction	0.1883
Inlet O_2 mole fraction	0.4156
Inlet CH_4 mole fraction	0.0040
Inlet CO mole fraction	0.0085
Inlet H_2O mole fraction	0.2166
Inlet C_2H_6 mole fraction	0.0024
Inlet CO_2 mole fraction	0.0002
Inlet Tar mole fraction	0.0029
Inlet SO_2 mole fraction	0
Inlet COS mole fraction	0
Inlet H_2S mole fraction	0.0061
Inlet N_2 mole fraction	0.1545
Inlet temperature of entering gas	461.904 K
Inlet char flowrate	0.7839 mol/s
Inlet particle temperature	390 K
Inlet slag thickness	1e-05 m
Nominal fuel flowrate into gasifier	41.2 kg/h
Nominal oxygen flowrate	37.2 kg/h
Nominal steam flowrate	10.7 kg/h

The models currently available to represent the dynamic behaviour of compact entrained-flow gasification units are limited due to their complexity[66]. Although computational fluid dynamics (CFD) can be used to simulate the steady-state behaviour of the multi-phase transport phenomena of the gasification process as well as its reactions, this modelling approach would not be computationally efficient for the purpose of modelling the transient operation of the process[143]. Recently, multiple studies have revealed that a suitable and computationally attractive dynamic model for the gasification process can be obtained using one-dimensional reduced-order models (ROMs)[144,145,146,147,148]. In previous studies, Sahraei et al.[72,142] developed a dynamic ROM for the pilot-scale unit considered in the present study in this chapter. That model was validated using both CFD simulation data and transient experimental data. Hence, the present study used that dynamic ROM as an adequate representation of the operation of the pilot-scale gasification unit.

The dynamic ROM used in this work was designed based on the streamlines obtained from CFD simulations of the pilot-scale unit (Figure 3-2(a)) around a nominal operation condition specified in Table 3-1. The CFD data was used to design a reactor network model, which aims to capture the different flow zones taking place inside the gasification unit. As shown in Figure 3-2(b), the reactor network consists of three plug-flow reactors (PFRs) that model the laminar jet-flow configuration as well as two continuous stirred-tank reactors (CSTR) to capture the mixing flow regime of the pilot-scale's gasifier. As shown in Figure 3-2(b), the downstream zone (DSZ) is represented using a PFR, however, two PFRs and two CSTRs are required to express the jet expansion zones (JEZs) and external recirculation zones (ERZs), respectively [72].

(a)



(b)

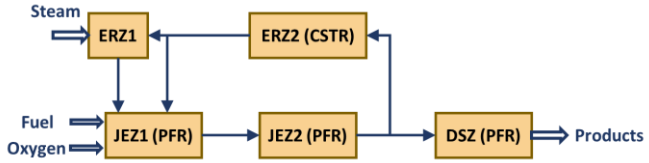


Figure 3-2: (a): Reactor network model for the gasifier; (b). Streamlines of CFD simulation; JEZ: jet expansion zone; ERZ: external recirculation zone; DSZ: downstream zone [72]

The dynamic ROM consists of over 80 algebraic equations that represent operability constraints and empirical correlations that are needed to estimate some of the model parameters, e.g., heats of reaction, reaction rates, recirculation ratio, slag thickness, slag viscosity, etc. The partial differential equations (PDEs) representing the momentum, mass, and energy balances for the gasifier are listed in Table 3-2 [72,142]. The momentum balance consists of 3 PDEs representing conservation equations for gas velocity, particle velocity, and particle number density. The particle number density is calculated throughout the reactor and over time as it changes due to particle deposition on the walls. The particle number density PDE is a critical equation in the calculation of the particle velocity and solving the momentum balance. The outcome of the momentum balance is used in the mass and heat balances where the particles devolatize and the solid residue is modeled as char. The mass and heat balances consist of 16 PDEs and involve mole fractions of all the species, mass, temperature, and slag thickness across the gasifier. Moreover, the CSTR model describing the behaviour in the ERZ zones (see Figure 3-2(b)) contains 15 PDEs that describe the mass and heat balance conservation equations. As shown in Figure 3-2(b), the gasifier can be split into two sections, i.e., the top section and the DSZ section. In each zone, 14 nodes (each describing the behaviour of the process at a particular height in the gasifier) are considered. Accordingly, a set of 284 PDEs and more than 80 algebraic equations are solved simultaneously for the top section of the gasifier to evaluate the velocity, pressure, composition, slag thickness, and temperature profiles of the system at any time interval. Note that the notation for the model equations and the parameters of the model are presented in the nomenclature. Moreover, sub-model descriptions, i.e., slag modeling and the modified correlations such as the recirculation ratio of the reactor network correlation, can be found elsewhere[72,142]. Please note that A_{cs} represents the cross-sectional area within each reactor in the network. As described in Sahraei et al.[142], the JEZ sections of the gasifier consider changes in A_{cs} at different heights (z) within the top section of the reactor, i.e., the cross-sectional area is determined as a function of the location in the gasifier.

Table 3-2: Partial different equations, dynamic ROM [72,142]

Gas phase	
Velocity	$\frac{\partial\left(A_{cs}\varepsilon_g\rho_g u_g\right)}{\partial t} = -\frac{\partial}{\partial z}\left(A_{cs}\varepsilon_g\rho_g u_g^2\right) + A_{cs}\left(-\frac{dp}{dz} + \varepsilon_g\rho_g g - F'_{g\to wall} - F'_{g\to p}\right)$
Molar composition	$\frac{\partial\left(A_{cs}\varepsilon_g C_{total}x_i\right)}{\partial t} = \frac{\partial}{\partial z}\left(A_{cs}D_{g,eff}\frac{\partial\left(\varepsilon_g C_{total}x_i\right)}{\partial z}\right) - \frac{\partial\left(A_{cs}\varepsilon_g u_g C_{total}x_i\right)}{\partial z} + A_{cs}(R_{hom} + R_{het});$ $where\ i = H_2, O_2, CH_4, CO, H_2O, C_2H_6, CO_2, SO_2, COS, H_2S, N_2$
Temperature	$\frac{\partial\left(A_{cs}\varepsilon_g C_{total}c_{P_g}T_g\right)}{\partial t} =$ $\frac{\partial}{\partial z}\left(A_{cs}k_{g,eff}\frac{\partial T_g}{\partial z}\right) - \frac{\partial\left(A_{cs}\varepsilon_g u_g C_{total}c_{P_g}T_g\right)}{\partial z} + A_{cs}(R_{hom}H_R) + Q'_{convg\to p} - Q'_{convg\to slag}$
Solid phase	
Velocity	$\frac{\partial\left(A_{cs}\varepsilon_p\rho_p u_p\right)}{\partial t} = -\frac{\partial}{\partial z}\left(A_{cs}\varepsilon_p\rho_p u_p^2\right) + A_{cs}\left(\varepsilon_p\rho_p g + F'_{g\to p}\right)$
Mass flow rate	$\frac{\partial\left(M_{solid}/u_p\right)}{\partial t} = -\frac{\partial M_{solid}}{\partial z} + A_{cs}(R_{het}) - m'_{slagging}$
temperature	$\frac{\partial\left(A_{cs}\varepsilon_p\rho_{solid}c_{P_p}T_p\right)}{\partial t} = -\frac{\partial}{\partial z}\left(A_{cs}\varepsilon_p u_p\rho_p c_{P_p}T_p\right) + A_{cs}(R_{het}H_R) - Q'_{conv\to g} - m'_{slagging}h_p$
Particle number density	$\frac{\partial\left(A_{cs}N_p\right)}{\partial t} = \frac{\partial}{\partial z}\left(A_{cs}N_p u_p\right) + \frac{m'_{slagging}}{m_p}$
Liquid phase	
Slag mass	$2\pi r_{gasifier}\rho_{slag}\frac{\partial\left(\delta_{slag}\right)}{\partial t} = -\frac{\partial m_{slag}}{\partial z} + m'_{slagging}$
Slag temperature	$2\pi r_{gasifier}\rho_{slag}c_{P_{slag}}\frac{\partial\left(T_{slag}\delta_{slag}\right)}{\partial t} =$ $2\pi r_{gasifier}k_{slag}\frac{\partial}{\partial z}\left(\delta_{slag}\frac{\partial T_{slag}}{\partial z}\right) - \frac{\partial}{\partial z}\left(m_{slag}c_{P_{slag}}T_{slag}\right) + Q'_{convg\to slag}$ $+ Q'_{convslag\to surronding} + m'_{slagging}H_p$

As shown in Figure 3-3, the dynamic ROM considers 14 uniformly distributed nodes across the axial domain for each section of the gasifier, i.e., the top section and DSZ section. Previous studies on this model have shown that 14 nodes are sufficient to accurately predict the operation of the gasification unit[72]. The distance between adjacent nodes is approximately 0.040 m. Note that these nodes are located in the JEZ1 and JEZ2 sections for the top section of the gasifier; the DSZ section of the gasifier is also divided into 14 nodes. Note that the JEZ1, JEZ2 and DSZ are modelled using PFRs as shown in Figure 3-2(a). Each of these nodes represents a potential sensor location to monitor the state variables of this process in real-time. As mentioned above, the dynamic ROM discretizes all of the PDEs based on the node specification to predict the temporal evolution of each state variable at a particular height in the gasifier. Although the discretization of the PDEs using the nodes significantly increases the number of

differential equations to be solved by the model, and therefore the corresponding computational costs, this significantly improves the accuracy of the dynamic ROM to capture the dynamic and nonlinear behaviour of the gasification process within the gasifier. Figure 3-3 shows a schematic of the potential sensor locations of the state variables within the top section of the gasifier (denoted as JEZ1 and JEZ2 in Figure 3-2). At each node, all of the mole fractions (e.g., H_2, O_2, CH_4 etc.), char flow rate, slag thickness, and gas and solid temperatures are identified as states of the system and calculated from the dynamic ROM. That is, at every location in the axial domain of the gasifier, there are 16 state variables that describe the temporal evolution of the system. Moreover, the same state variables (except for the slag thickness) represent the behaviour of the ERZ sections of the gasifier. As the ERZ sections in the top section of the gasifier have been modeled using two CSTRs (see Figure 3-2(b)), the predictions of all 15 states are assumed to be constant within these zones. In addition, the DSZ of the gasifier also contains similar state variables and potential sensor locations to those considered for the top section. However, this section does not include any CSTR unit, as shown in Figure 3-2(b).

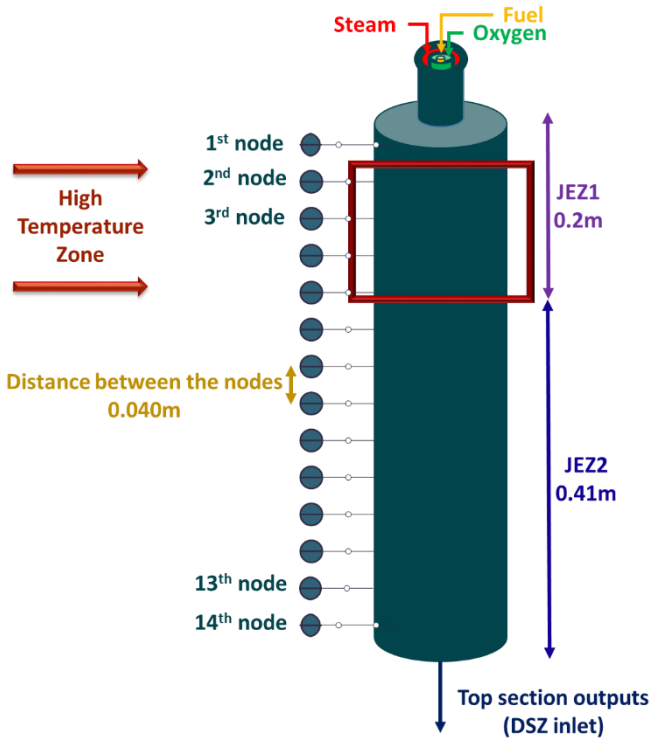


Figure 3-3: A schematic of the potential location of sensors for the top section of the gasifier

Based on the above, the total number of states is 239 and 240 for the top and bottom sections of the gasifier, respectively. Due to the size and complexity of the problem, the current study focuses on the

top section of the gasifier because the most critical states relevant to the operation of this unit, namely the peak temperature and slag thickness, are located in this section.

3.2 State Estimation (Kalman Filter)

Given the complexity of the system considered in this study, it is critical to use an estimation tool that can predict the states accurately and in short simulation times so that it can be adequately used for online monitoring. However, in the case of large-scale applications such as the gasifier unit, high simulations costs are expected due to the complexity and high nonlinearity of the system. This study considers the use of KF (and EKF) as a practical tool that can provide online estimation of the key states in reasonable computational times.

Consider a nonlinear model with states $\boldsymbol{x}$, inputs $\boldsymbol{u}$, and measurements $\boldsymbol{y}$ can be described as follows:

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k \tag{3-1}$$

$$\boldsymbol{y}_k = h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{v}_k \tag{3-2}$$

$$\mathbb{E}\left[\boldsymbol{w}_\varsigma \boldsymbol{w}_k^T\right] = \begin{cases} \boldsymbol{Q}, & \varsigma = k \\ 0, & \varsigma \neq k \end{cases} \; ; \; \mathbb{E}\left[\boldsymbol{v}_\varsigma \boldsymbol{v}_k^T\right] = \begin{cases} \boldsymbol{R}, & \varsigma = k \\ 0, & \varsigma \neq k \end{cases} \tag{3-3}$$

$\boldsymbol{w} \in \mathbb{R}^{n_x}$, $\boldsymbol{v} \in \mathbb{R}^{n_y}$, $\boldsymbol{x} \in \mathbb{R}^{n_x}$, $\boldsymbol{y} \in \mathbb{R}^{n_y}$, $\boldsymbol{u} \in \mathbb{R}^{n_u}$, $\boldsymbol{Q} \in \mathbb{R}^{n_x \times n_x}$, $\boldsymbol{R} \in \mathbb{R}^{n_y \times n_y}$,
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 constant. Moreover, $\boldsymbol{w}$ and $\boldsymbol{v}$ are mutually uncorrelated, zero-mean Gaussian random process uncertainties and measurement noise signals, respectively. $\boldsymbol{Q}$ and $\boldsymbol{R}$ are diagonal covariance matrices for the process uncertainties and the measurement noise signals, respectively. Note that $\mathbb{E}\left[\boldsymbol{w}_\varsigma \boldsymbol{x}_0^T\right] = 0$ for all k where $\boldsymbol{x}_0$ denotes the vector of initial states. Moreover, the function f describes the process model, whereas the function h represents the measurement model. For the nonlinear dynamic system presented in Equations (3-1)-(3-3), the formulation of the KF/EKF is as follows[149]:

Prior estimation

$$\begin{aligned} &\hat{\boldsymbol{x}}_{k+1|k} = f\left(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k\right) \approx \boldsymbol{A}_{k+1}\hat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_{k+1}\boldsymbol{u}_k \\ &\boldsymbol{P}_{k+1|k} = \boldsymbol{A}_k \boldsymbol{P}_{k|k} \boldsymbol{A}_k^{\boldsymbol{T}} + \boldsymbol{Q} \\ &\text{where } \boldsymbol{A}_k = \frac{\partial f(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{x}} \left| \boldsymbol{x} = \hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u} = \boldsymbol{u}_k \right. \end{aligned} \tag{3-4}$$

Posterior estimation

$$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1|k}\boldsymbol{H}_{k+1}^{T}/(\boldsymbol{H}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{H}_{k+1}^{T} + \boldsymbol{R})$$
$$\widehat{\boldsymbol{x}}_{k+1|k+1} = \widehat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1}\left(\boldsymbol{y}_{k+1} - h\left(\widehat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k\right)\right) \quad (3\text{-}5)$$
$$\boldsymbol{P}_{k+1|k+1} = (\boldsymbol{I} - \boldsymbol{K}_{k+1}\boldsymbol{H}_{k+1})\boldsymbol{P}_{k+1|k}$$
where $\boldsymbol{H}_k = \frac{\partial h(\boldsymbol{x},\boldsymbol{u})}{\partial x}|\boldsymbol{x} = \widehat{\boldsymbol{x}}_{k|k}, \boldsymbol{u} = \boldsymbol{u}_k$

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

where the Jacobian matrix $\boldsymbol{A}$ denotes the correlation between the system states; the sensitivity matrix $\boldsymbol{B}$ describes the correlation between the states and the inputs. This estimation scheme performs two major steps of estimation, namely the prior and the posterior estimation. The former is also known as the time update step as it projects the probability density function (PDF) of the states at the current time interval k ahead to estimate the prior distribution of the states in the next time interval $k+1$, i.e., $\widehat{\boldsymbol{x}}_{k+1|k}$ and $\boldsymbol{P}_{k+1|k}$. The last step evaluates the Kalman gain $\boldsymbol{K}$ at the time interval $k+1$ and updates both the expected value $\widehat{\boldsymbol{x}}_{k+1|k+1}$ and the covariance $\boldsymbol{P}_{k+1|k+1}$ of the posterior distribution. Note that in the current study, $\boldsymbol{R}$ and $\boldsymbol{Q}$ are assumed to be constant over time.

KF uses a linear dynamic model with constant $\boldsymbol{A}$ and $\boldsymbol{B}$; these are evaluated only at the nominal value of the states ($\boldsymbol{x}_{\mathbf{nominal}}$) and the inputs ($\boldsymbol{u}_{\mathbf{nominal}}$). On the other hand, EKF updates both matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ at each time interval k based on the latest inputs and posterior estimates. Thus, EKF requires additional computational effort to perform these updates, particularly for large-scale applications such as the pilot-scale gasifier unit. The main challenges in the estimation, as well as the outcome of performing KF and EKF for the pilot-scale gasification unit considered in this study, are discussed next.

3.3 Results

The current study performs state estimation using Kalman Filters for a pilot-scale gasifier under different scenarios, which aim to investigate the capability of this technique in the presence of additive uncertainty in the prior estimation, common changes in the input streams of the gasifier, and in the case of plant-model mismatch. In all of these scenarios, one of the main challenges of performing KF and EKF for the gasification unit is to adequately estimate matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in order to capture the dynamic behaviour of the process accurately. As the current case study includes 239 states, the Jacobian matrix $\boldsymbol{A}$ is a 239x239 square matrix whereas $\boldsymbol{B}$ is a 239x81 matrix. The 81 inputs to the Kalman filter framework represent the characteristics of inlet streams (e.g., molar fractions for the species,

temperatures, slag thickness, and char flowrate), inlet flowrates (e.g., fuel, steam, and oxygen), and the velocity and pressure profiles. The numerical evaluation of the Jacobian matrix for such a large-scale system is computationally expensive. The required CPU time for the analytical evaluation of these matrices at the nominal operating condition is approximately 19.30 minutes. Once the analytical solution of the sensitivity matrices is available, the evaluation of the sensitivity matrices is relatively fast (~0.6 s). The analytical calculation of the $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices has been conducted using the Python function *lambdify*. The dynamic ROM presented in the previous section was implemented in Python 3.7. Similarly, all the computational experiments performed in this study were conducted using Python 3.7 on a computer equipped with 32 GB RAM and Intel(R) Core(TM) i7-4770 CPU @ 3.40 GHz. Note that the linear state space model used in this study has been validated using the dynamic ROM for the gasifier. In this work, the dynamic ROM was considered as the true plant model of the gasification unit. Figures 3-4 (a)-(d) compare the predictions of the linear model to the states estimated by the dynamic ROM (denoted as "True State") for temperature, slag thickness, H_2 mole fraction, and the CO mole fraction at the 14th node (end of the top section). As shown in this figure, the linear state space model captures the transient operation of the gasification unit. Note that this validation was performed under the load-following condition (see section 3.3.2).

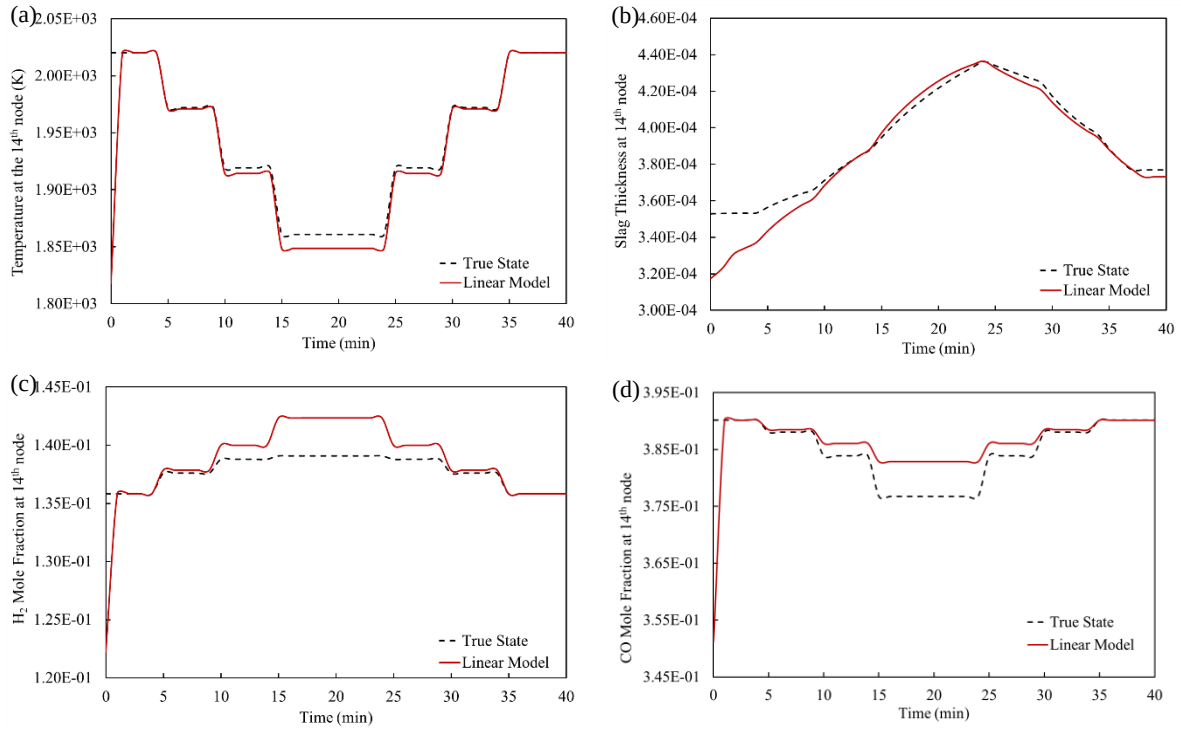


Figure 3-4: Accuracy of the linear state space model with respect to the true value of the states at the 14th node for (a) solid temperature; (b) slag thickness; (c) H_2 model fraction; (d) CO mole fraction

A key feature in the current work is associated with the specification of an adequate time interval in the state estimation framework such that the proposed dynamic model can be solved with sufficient accuracy and avoid numerical instabilities. That is, the time interval to extrapolate the linear model and estimate the prior distribution of states $\mathcal{N}\left(\hat{\mathbf{x}}_{k+1|k}, \boldsymbol{P}_{k+1|k}\right)$ is key to KF accuracy. In addition, the gasifier unit considered in this study involves states that evolve at different scales, e.g., the mole fraction of H_2 across the gasifier is on average six orders of magnitude smaller than the gasifier's average temperature in the top section. To circumvent this issue, relatively small time steps, i.e., 1e-5 s, are needed to maintain numerical stability in the prediction of the prior distribution of the estimated states. Note that the sampling time interval of the actual system is one minute. That is, the prior estimation is updated at very small time steps (i.e., every 1e-5s) while the posterior estimation is updated once new measurements become available (i.e., every minute). Although the proposed approach to deal with the complexity of the current system increases the CPU costs, it allows the KF to maintain stable estimations of the prior states and covariance matrix.

A few alternatives were considered to increase the time interval (1e-5s). The maximum absolute eigenvalue of the Jacobian matrix evaluated at the nominal point is 1.18582e+5. Note that the required time step in the prior KF estimation to maintain numerical stability is proportional to the inverse of the maximum absolute eigenvalue of the Jacobian matrix [150]. Thus, the time-step considered in this study (1e-5s) is the largest time-step that can be considered to achieve numerical stability of the linear state space model. The Jacobian conditioning (currently at 4.4081e+11 at the nominal operating point) was also considered to improve the computational efficiency in estimation. One potential alternative to deal with the time scale problem was to normalize the system. However, even in the case of normalization, there is no guarantee that the normalized state-space linear model may not contain large eigenvalues or improve conditioning of the Jacobian matrix. Also, normalization may add additional numerical errors to the system that may result in numerical intractability. An alternative considered was to assume that states with fast dynamics would only be considered using their steady-state information, i.e., the dynamic behavior of those states would be neglected. To perform this task, all the states associated with large eigenvalues (>1,000) were removed from the linear state space model and only their steady-state information would be considered to update those states in the KF estimation. A total of 56 states were removed from the linear state space model and were only considered as steady-state inputs to the model. Most of the states with fast dynamics involved the mole fractions of H_2, O_2, CH_4, CO along the axial

domain of the gasifier, which are important states that need to be estimated accurately for the purpose of monitoring and control of the gasification process. The resulting reduced linear state space model was used for prior estimation in the KF framework using larger time steps, i.e., 1e-3s and 1e-4s. Results from this implementation showed that the KF estimation diverged, mostly because the maximum absolute eigenvalue of the corresponding Jacobian matrix in the reduced linear state space model was still in the order of 1e+5. In addition, this study also considered other reduced linear state space models that involved different combinations between states with fast dynamics, e.g., two linear reduced state space models with total number of states 225 and 28. The 225 states-ROM involved all the states except the model fractions of H_2, whereas the 28 states-ROM only involved the wall temperature and the slag thickness along the gasifier. The maximum eigenvalue in the 225 states-ROM and 28 states-ROM was 1.18582e+5 and 1.0251e+4, respectively, which limit the use of time steps larger than 1e-5 s. As a result, the KF estimation also diverged when larger time-steps were used, mostly due to inaccurate prior estimations. These tests indicate that small time intervals are needed to maintain numerical stability of the linear state space model.

Moreover, the number of states that can be measured online may be limited in an actual setting. The type of states that can be measured online as well as the location of the sensors inside the gasifier can affect the quality of the estimation, particularly in the case that the linear model does not capture the process behaviour due to the presence of process uncertainties, additive uncertainty in the prior estimation, or plant-model mismatch. To gain insight on the impact of sensor availability and their location within the gasifier, the present study analyzed the performance of the estimation under different numbers, types and locations of the sensors around the gasifier. Furthermore, this study assumed that the standard deviation of the process uncertainty associated with each of the plant states is 2% of the nominal steady-state values of the states, $\boldsymbol{x}_{\mathbf{nominal}}$ (see Table 3-1), whereas the standard deviation of the measurement noise is negligible and set to 1e-6$\boldsymbol{x}_{\mathbf{nominal}}$. This study assumed that the dynamic ROM is the true plant and provides the true states of the gasification unit. Thus, the actual states of the system (i.e., Plant Outputs) are the outputs of the dynamic ROM (i.e., True States) complemented with the added process uncertainties. The additional underlying assumptions of each of the scenarios and the detailed discussion of the results are presented next.

3.3.1 *Scenario I*: Additive uncertainty in the prior estimation

In order to analyze the impact of sensor location in the KF performance, this scenario aims to validate the estimation of KF under three different sets of the partial states considered as *Case A*, *Case B*, and

Case C. Table 3-3 shows the detailed information of the partial states known for each case, i.e., the types and locations of the measurable states. Note that T_g and T_w represent the temperature of the gas and the gasifier's wall, respectively; C_i is the mole fraction of the species i shown in Figure 3-1(b), and *ST* denotes the slag thickness. As mentioned above, having access to the measurements of the states within the gasifier is difficult due to the extreme operating conditions, particularly around the zone where the wall temperature is at its highest. As the peak temperature often takes place in the second-highest node in the gasifier (see Figure 3-3), it is more likely that sensor hardware can only be employed far away from the peak temperature zone in the gasifier due to the extremely high temperatures (>2000°C). Thus, all of the states within the first 5 nodes are assumed to be unknown in all of the three cases. Moreover, *Case A* assumes that online measurements are available every minute for all of the states within the last 9 nodes (within the last 0.406 m of the top section) along the axial domain of the gasifier, i.e., the gas temperature, wall temperature (using thermocouples), concentration of all of the components (using gas chromatography), and the slag thickness (using viscosity-based correlations). Note that the viscosity is assumed to be measurable through the dynamic ROM model. The following equations describe the viscosity as a function of the silica ratio of slag and temperature of the slag layer (notations are provided in the nomenclature section)[72]:

$$S = \frac{W_{SO_2}}{W_{SiO_2} + W_{Fe_2O_3} + W_{CaO} + W_{MgO}} \tag{3-6}$$

$$\log(\mu_{slag}) = 4.468S^2 + 1.2659\left(\frac{10^4}{T}\right) - 8.44 \tag{3-7}$$

$$M_{slag} = \frac{2\pi r_{gasifier}\delta_{slag}^3\rho_{slag}^2 g}{3\mu_{slag}} \tag{3-8}$$

In addition, *Case A* assumes that all the states are measurable in the ERZ sections (CSTR). In other words, out of the 239 state variables of the system, only 80 of the states are assumed to be unknown in *Case A*.

Based on above, *Case A* corresponds to an idealistic condition because access to online measurement of concentrations of species is not often performed in an actual setting. *Case B* assumes that only 9 thermocouples can be placed in the last 9 nodes of the top section of the gasifier, reporting the temperature of the wall every minute. Note that *Case B* assumes that all of the concentrations are unknown; thus, 230 out of the 239 states are not measured in real time. According to the profile of the temperature and the molar fractions for the species, the variability of the states becomes larger in the

adjacent nodes to the hot spot of the gasifier. However, the changes in the true states in the last 8 nodes of the gasifier, which are far from the gasifier's hot spot, are relatively small. Moreover, as mentioned in the chapter 2, the expenses associated with sensors impose a restriction on the number of sensors that can be located in the unit. Hence, *Case C* has been defined as a more realistic case that considers only three thermocouples that are located four nodes apart. Note that these three thermocouples provide the wall temperature and are located at nodes 6, 10 and 14 (i.e., 0.366 m, 0.203 m, and 0.040 m) away from the end of top section of the gasifier, respectively.

Observability tests for *Cases A-C* have been performed using the linear state space model. A linear state space system is observable if and only if [151]:

$$rank(H, A - \lambda_i I) = n_x; \qquad i = 1, 2, 3, \dots, n_x \tag{3-9}$$

where n_x is the number of states variables of the system and λ_i denotes the i^{th} eigenvalue in the eigenvector. Moreover, the sensitivity matrix $\boldsymbol{H}$ describes the correlation between the measurements and the system states. This observability test was satisfied for *Cases A-C* considered in this study. Note the time-dependent linear state-space model of the gasifier is not an explicit function of time, i.e., the system is time-invariant[152]. An observability analysis using the actual nonlinear model is part of the future work.

Table 3-3: Sensors type and location in (a) *Case A*; (b) *Case B*; (c) *Case C*

PFR Node	Distance from top of the gasifier (m)	*Case A*				*Case B*				*Case C*			
		T_g	T_w	C_i	ST	T_g	T_w	C_i	ST	T_g	T_w	C_i	ST
6	0.244	✓	✓	✓	✓	×	✓	×	×	×	✓	×	×
7	0.284	✓	✓	✓	✓	×	✓	×	×	×	×	×	×
8	0.325	✓	✓	✓	✓	×	✓	×	×	×	×	×	×
9	0.366	✓	✓	✓	✓	×	✓	×	×	×	×	×	×
10	0.406	✓	✓	✓	✓	×	✓	×	×	×	✓	×	×
11	0.447	✓	✓	✓	✓	×	✓	×	×	×	×	×	×
12	0.488	✓	✓	✓	✓	×	✓	×	×	×	×	×	×
13	0.528	✓	✓	✓	✓	×	✓	×	×	×	×	×	×
14	0.569	✓	✓	✓	✓	×	✓	×	×	×	✓	×	×
CSTR	0-0.61	✓	✓	✓		×	×	×		×	×	×	

At each time point (i.e., every 1 min), a time-varying variable has been randomly chosen around $\pm 0.01\boldsymbol{x}_{\mathbf{nominal}}$ using a binomial distribution as shown in Figure 3-5. These time-varying random variables have been added to the prior estimation of states at time $k+1$ as a way to consider additive uncertainty in the linear model predictions. This was done to consider cases when there exist errors

(mismatch) in the numerical identification of linear state space models used for prior estimation purposes. Therefore, in the presence of both additive uncertainties to the prior estimation (1% of $\boldsymbol{x}_{\mathbf{nominal}}$) and process uncertainties (2% of $\boldsymbol{x}_{\mathbf{nominal}}$), the available measurements considered in the posterior estimation of the states may improve the accuracy of the state estimation.

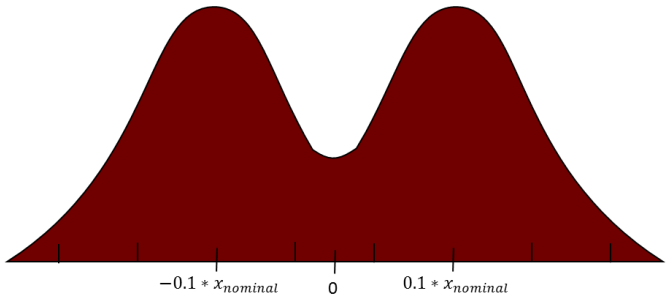


Figure 3-5: Distribution of additive uncertainty considered in the prior estimation in *Scenario I*

Figure 3-6(a) and 3-6(b) present the results of the KF for two of the most critical states in the gasification process, i.e., the peak temperature (wall temperature at the 2nd node) and the slag thickness at the 14th node of the top section of the gasifier, respectively. Note that the peak temperature cannot be measured for none of the *Cases A-C*. On the other hand, the slag thickness at the 14th node can only be measured online in *Case A* (see Figure 3-6(b)).The results presented in Figure 6 for the current scenario include the outcome of KF for *Cases A-C*, the output from the dynamic ROM with the added process uncertainties (denoted as "Plant Output"), and the true values of the states provided by the dynamic ROM without considering process uncertainties (denoted as "True State"). Note that in all three cases (*Cases A-C*), an initial condition that is different from the nominal steady-state condition has been considered to initialize the prior states in KF framework, i.e., 0.9*(the true steady-state value). Consequently, the system is expected to respond to this initialization and move the states from their initial values to their true (steady-state) values. Note that the system involves states with fast and slow dynamics, i.e., different time-scales. For instance, the wall temperature presents fast dynamics (see Figures 3-4(a) and 3-6(a)), whereas the slag thickness presents slow responses (see Figures 3-4(b) and 3-6(b)). Thus, as shown in Figure 3-6(b), the estimation of the slag thickness at the last node of the top section is slowly approaching the true state in both *Case B* and *Case C* since the KF estimation mainly relies on the linear state space model. Moreover, note that *Case A* assumes that the online measurements are available for this particular state. This assumption for *Case A* aids KF to recover from the inaccurate initial condition in shorter times when compared to *Cases B* and *C*. Moreover, the small variations in the estimations for *Case A* is due to the small but random measurement noise and process uncertainty

associated with the measurements considered in *Case A*. On the other hand, the limited number of measurements considered in *Cases B-C* forces the estimation for this state to rely on the prior KF estimations, which do not capture the variation in the "Plant Output". Nevertheless, based on Figure 3-6, the results of *Scenario I* show that KF provides estimations of the unknown states with similar accuracy for the three cases. This result is expected because the gasifier was assumed to be operating at the steady-state point around which the linear state space model was developed. Thus, KF can still rely on the linear state space model even though there are small process uncertainties and added noises considered in the prior estimation. The same conclusion can be drawn by calculating the mean square error (MSE) between the estimated states and the plant output. The MSE for all *Cases A-C* and for both slag thickness and peak temperature is approximately the same (3-4% different), which suggest that the quality of estimation using KF is adequate, even in *Case C* where three thermocouples are the only sensors available for monitoring the operation of the gasifier.

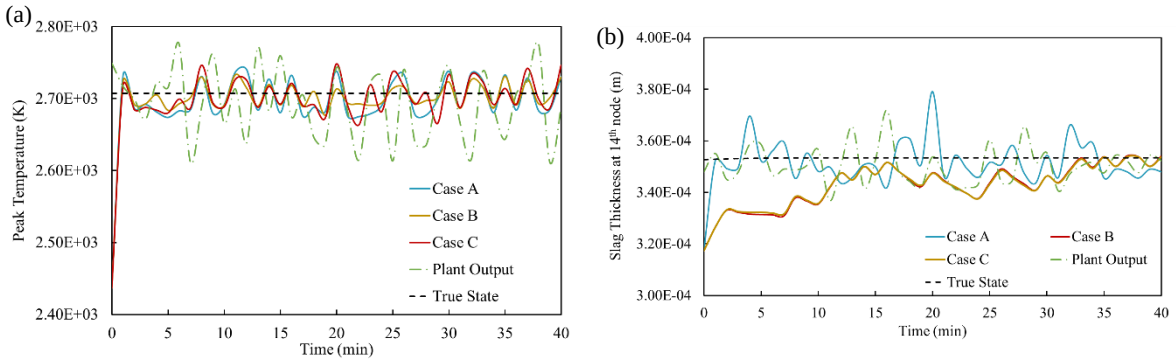


Figure 3-6: KF performance in the presence of additive uncertainty in the prior estimation: (a) peak temperature inside the gasifier; (b) slag thickness located in the 14th node of the top section of the gasifier

A comparison between KF and EKF was performed next. To maintain a successful EKF implementation for the pilot-scale gasifier, the Jacobian matrix $\boldsymbol{A}$ needs to be updated more frequently than the actual sampling time of the plant so that it remains numerically stable. Therefore, the implementation of EKF under the assumptions mentioned for *Case C* has been considered. However, for this comparison, the new measurements from the process are assumed to become available every second instead of every minute. This was done with the sole purpose of reducing the computational costs in the EKF estimation. This assumption holds for both the EKF and the KF implementation. Figure 3-7(a) and 3-7(b) shows the performance of both estimation methods for the peak temperature and the slag thickness at the 4th node of the gasifier, respectively. Note that online measurements are not available for these two states in any of the *Cases A-C*. The estimation provided by both the EKF

and KF for the slag thickness is not accurate over the period of simulation considered for this comparison. This was expected as the slag thickness has slow dynamics, i.e., it takes several minutes to adjust to the dynamic changes in the model and reach to its steady-state condition, as shown in Figure 3-6(b). Moreover, as shown in Figure 3-7(b), both EKF and KF present similar predictions for the slag thickness at the 4th node for the first few seconds. Note that this test has been performed considering that the time interval is 1 s, whereas the results presented in Figure 3-6 have been obtained using a 1 min time interval. Thus, the combination of the small time scales and the slow slag thickness dynamics produces an overall slow dynamic response in this variable. The results presented in Figure 3-7 show that there is no improvement in the accuracy of the estimation provided by EKF when compared to the estimation from KF. However, due to the complexity and different time scales associated with the gasification unit, the computational time required by EKF to update matrices ***A*** and ***B*** at every time interval is approximately three orders of magnitude larger than that required by KF; this is mostly due to the large number of function evaluations that need to be performed (approximately 1e+5 function evaluations per 1 s of simulation). Hence, the state estimation using EKF is computationally intractable while using the present gasification model. Methods that can circumvent the scaling of matrices ***A*** and ***B*** may help in reducing the computational costs but this is beyond the scope of this work. Therefore, only KF estimation is considered for *Scenario II* and *Scenario III*. Note that this test has been performed only for 15 seconds of the simulation. Results may differ when larger simulation times using other means to perform the estimation can be employed.

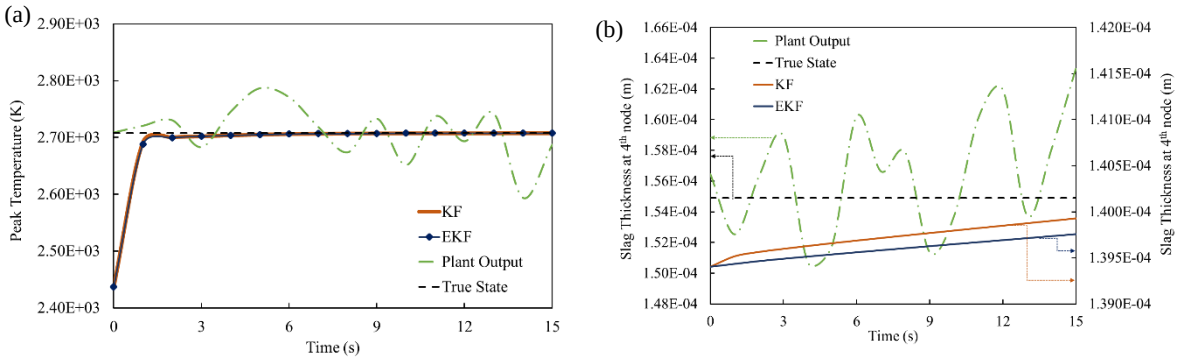


Figure 3-7: A Comparison of the EKF performance and the KF performance for (a) peak temperature inside the gasifier; (b) slag thickness located in the 4th node of the top section of the gasifier; right-scale for the KF and EKF; left-scale for the Plant Output and True States

3.3.2 *Scenario II*: Load-following scenario

This scenario aims to assess the capability of KF to deal with the dynamic behaviour of the system caused by significant changes to the inlet stream of the gasifier. One of the most common changes in the operating conditions of the gasifier occurs under load-following conditions. In an IGCC power plant, a load-following scenario is often applied to tune the operating conditions of the gasification process so that the power generation capacity of the IGCC power plant can be adjusted according to daily changes in the energy demands[153]. A load-following scenario consists of one or more ramp or step changes in the fuel flowrate to adjust the amount of syngas fed to the IGCC plant. To perform this test, several step changes have been considered to reduce the fuel flowrate from its nominal value (see Table 3-1) up until it reached 60% of its nominal value. Step changes of a 10% change in magnitude with respect to the nominal fuel flowrate were applied every 5 minutes. Once the load was at 60% of the nominal flowrate reported in Table 3-1, the flowrate was increased with a rate of 10% every 5 minutes to return it to the nominal capacity of the power generation. Note that a single load-following cycle takes approximately 40 minutes. Moreover, it is important to mention that during the load-following scenario, both the steam and oxygen flowrates entering the gasifier are adjusted accordingly so that the ratio of fuel to steam and the ratio of fuel to oxygen remain constant to satisfy the thermal constraint of the gasifier refractory wall[154].

The results of the implementation of KF for *Scenario II* are presented in Figure 3-8. As in *Scenario I*, three cases have been considered for the present scenario to explore the impact of the number and the location of the sensors inside the gasifier. The results show that KF is capable of providing adequately accurate estimation of the unknown states in for *Cases A-C*. Figure 3-8(a) and Figure 3-8(b) represent the estimation of peak temperature and slag thickness at the 4th node, respectively. Note that these are the critical states for the purpose of monitoring the gasification unit in an IGCC power plant; these states are not measurable in any of the *Cases A-C*. Figure 3-8(a) shows that the accuracy of the state estimations provided by KF for the peak temperature in all three cases remains the same, despite the fact that *Case A* assumes a larger number of sensors in the plant than those considered by *Case B* and *Case C*. The MSE of the estimations for all *Cases A-C* are approximately the same (less than 5% different). These results show that the linear state space model properly represents the dynamics of the process so that the prior KF estimation of the peak temperature can be performed successfully, even

though there is a very limited number of sensors in the unit in *Case C*. In addition, according to Figure 3-8(b), estimation of the slag thickness (at the 4th node) in *Case B* and *Case C* is smoother compared to *Case A*. That is, the MSE for the estimation of the slag thickness in *Case A* is 25% larger than the evaluated MSE in *Case B* as well as in *Case C*. As expected, in *Case B* and *Case C*, KF relies more on the prior estimation rather than on the posterior estimation due to the limited number of measurements, whereas KF has access to more sensors and consequently can rely on the posterior estimation for *Case A*. As prior estimation uses the linear model that excludes any process uncertainties or noises, the estimations obtained in *Case A* seem noisier when compared to *Case B* and *Case C*. That is the estimation provided by KF for the measurable states follow their plant outputs for this scenario. For instance, Figure 3-8(c) shows the KF performance for the slag thickness in the 14th node (the last node in the top section of the gasifier). As shown in this figure, the estimation provided by KF in *Case A* follows the measurements while there is a small error in the estimation of this state for *Case B* and *Case C* where the measurements of the slag thickness are not available. As a result, the estimation in *Cases B* and *C* mainly relies on the linear state space model and thus follows the "True States" rather than the "Plant Output".

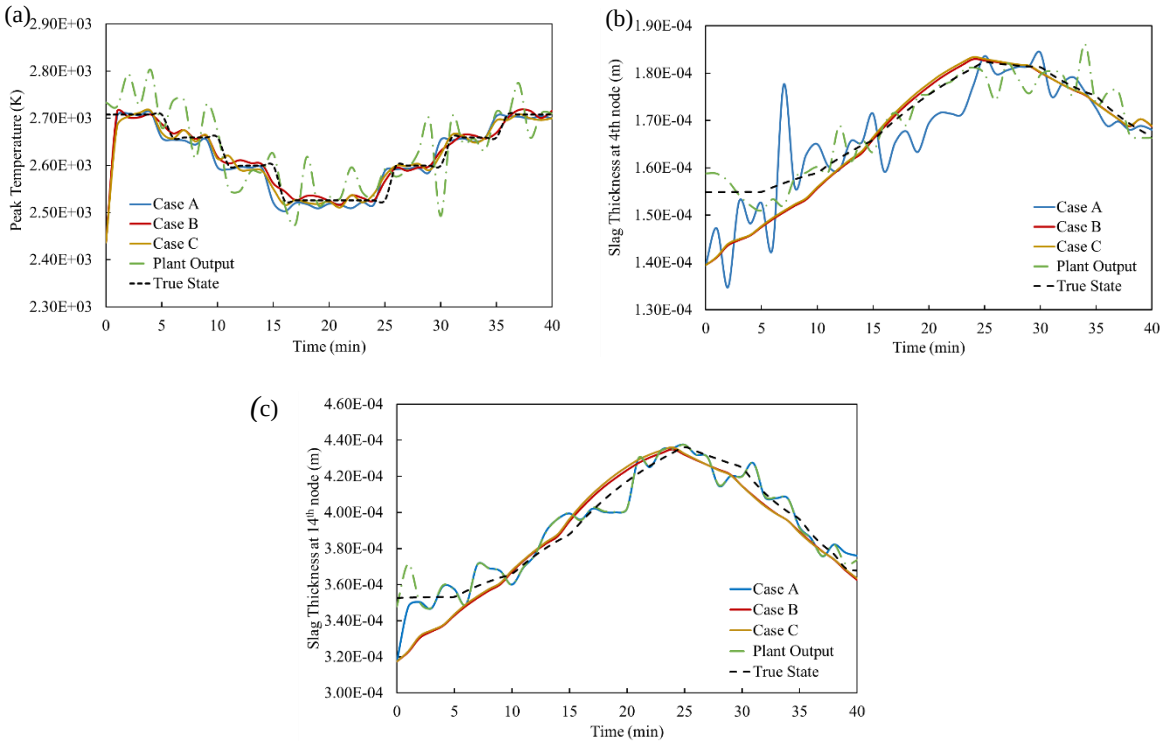


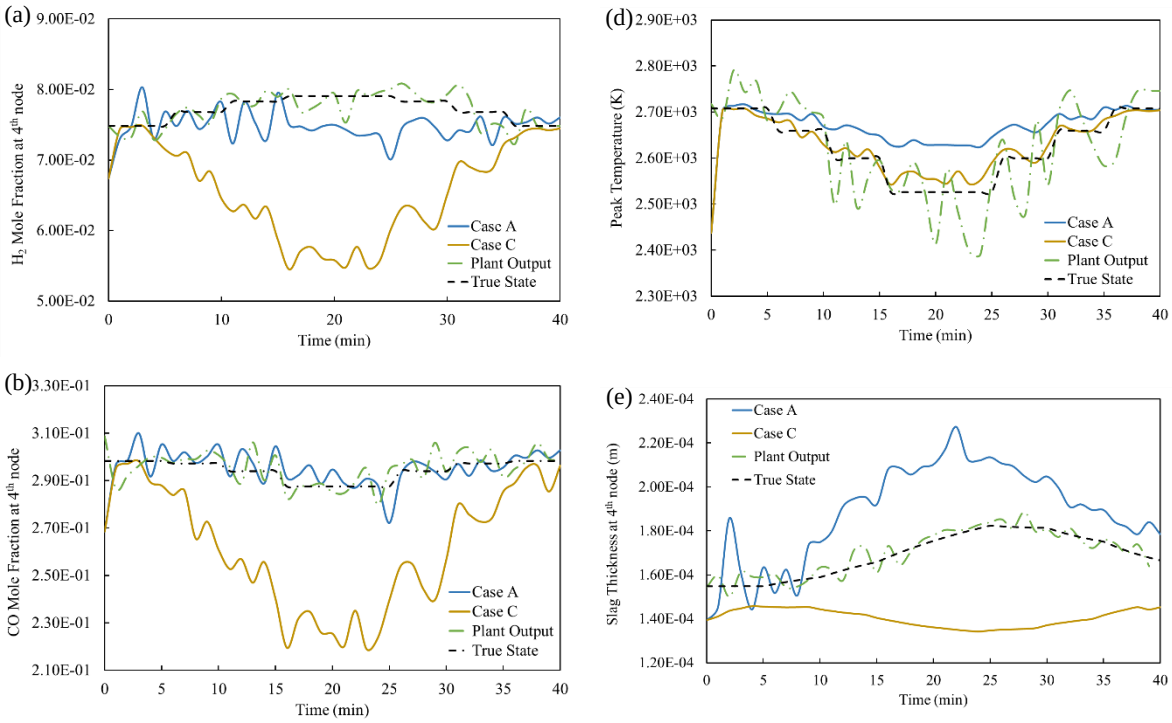
Figure 3-8: Estimations of unmeasurable states in the load-following condition for (a) peak temperature; (b) slag thickness at 4th node; (c) slag thickness at 14th node

3.3.3 *Scenario III*: KF performance under model structural error and load-following

The current scenario aims to investigate the performance of KF in the presence of plant-model mismatch for the pilot-scale gasifier. Therefore, this scenario considers a mismatch between the linearized version of the gasifier model and the dynamic ROM under load-following conditions. One challenge on the implementation of the KF algorithm on the gasifier under load-following conditions is to evaluate the sensitivity $\boldsymbol{B}$ matrix such that the linear state space model shown in Equation (3-4) is able to simulate the dynamic behaviour of the system caused by changes in the inputs of the plant. Given the complexity of the current pilot-scale gasifier and the nonlinear correlation between the momentum, mass, and energy balances of the system, a reasonable evaluation of the sensitivity matrices of the system is needed to adequately capture the transient operation of the gasifier, as presented in *scenario II*. However, in the presence of model mismatch, i.e., inaccurate $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices, KF is expected to heavily rely on the sensors available for the posterior estimation.

Based on the above, a similar load-following condition as described in *Scenario II* has been considered with the exception that the momentum balance equations were not included in the evaluation of the sensitivity $\boldsymbol{B}$ matrix. As a result, the linear state space model was not able to adequately capture the dynamic behaviour of the gasification process. As shown in Figures 3-9(a)-(c), the quality of the estimation for some of the unknown states in the vicinity of the peak temperature (located at the 4th node), i.e., hydrogen (H_2), carbon monoxide (CO), and carbon dioxide (CO_2) mole fractions, improved for *Case A* compared to *Case C* (see Table 3-3). This is due to the correction step in the KF algorithm where the online measurements coming from the sensors improve the posterior estimation of the states. Therefore, the quality of the estimation of unknown states improves as more measurements become available (*Case A*). However, according to Figure 3-9(d)-(f), in the case of some of the other unknown states, i.e., the peak temperature of the gasifier's inner wall, the slag thickness at the high temperature zone (at the 4th node), and the oxygen (O_2) mole fraction at the 4th node, KF was not able to adequately track the true states. As it shown in Figures 3-9(d)-(f), KF was not able to estimate these states properly in either *Case A* or *Case C*, which highlights the significance of plant-model mismatch in the estimation framework considered for *Scenario III*. Note that the KF estimation for *Case C* was able to track the changes in the actual peak temperature more closely than *Case A*, as shown in Figure 3-9(d). Note that *Scenario III* was performed under a model structural error (i.e., some element in the $\boldsymbol{B}$ matrix are

missing). This creates a plant-model mismatch in the linear state space model used in the KF framework. Due to this error, the prior estimation of the states provided by the linear state space model is not accurate and generates error at each time interval. Note that this error may vary for different states of the system, i.e., the missing elements in $\boldsymbol{B}$ affect in a different way the prior estimation of the different states. As shown in Equation (3-5), the term "$\boldsymbol{K}_{k+1}\left(\boldsymbol{y}_{k+1} - h\left(\hat{\boldsymbol{x}}_{k+1|k}, k+1\right)\right)$" depends on the prior estimation and therefore propagates an error between the measurements and the inaccurate prior estimation ($\hat{\boldsymbol{x}}_{k+1|k}$). Although more measurements improve the KF estimation, as it is the case for the states shown in Figure 3-9(a), the propagation of this error in the posterior calculation does not improve the posterior estimation for some of the state variables such as the peak temperature, as shown in Figure 3-9(d). That is, the more states available in the function h, the more errors that may be propagated through the posterior estimation due to the model structural error included in $\hat{\boldsymbol{x}}_{k+1|k}$.



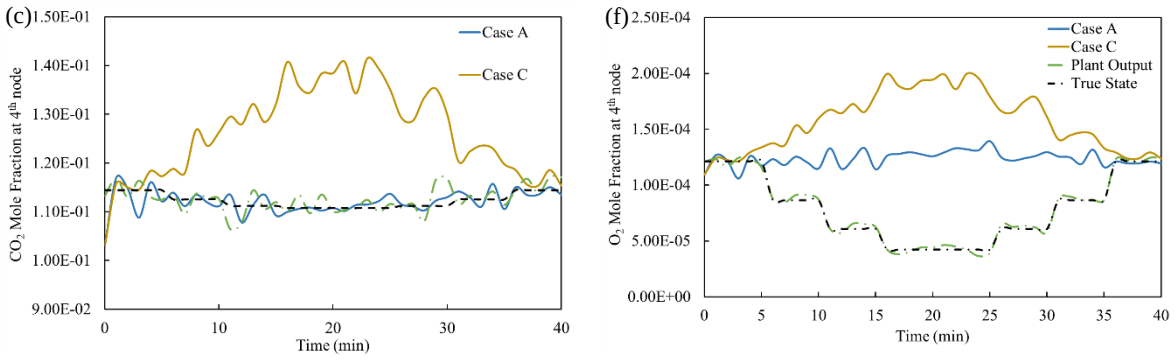


Figure 3-9: Estimations of unmeasurable states in the presence of model mismatch for (a) H2 model fraction at 4th node; (b) CO mole fraction at 4th node; (c) CO2 mole fraction at 4th node; (d) peak temperature; (e) slag thickness at 4th node; (f) O2 mole fraction at 4th node

Nevertheless, KF is expected to provide an accurate posterior estimation of the states for which online measurements are available. Figures 3-10(a)-(c), show the estimation results for wall temperature, slag thickness, and O_2 mole fraction at the 8^{th} node. Note that these three states are assumed to be measured online for *Case A*, while *Case C* does not consider online sensors for these states. As a result, there is a high-quality estimation provided in *Case A*, whereas the estimation in *Case C* was not able to recover from the poor prior estimation caused by the plant-model mismatch. Moreover, three thermocouples are responsible for reporting the online measurement of the wall temperature at the 6^{th}, 10^{th}, and 14^{th} nodes in both *Case A* and *Case C* (see Table 3-3). Therefore, according to the results shown in Figures 3-10(d)-(f), KF exhibited a reasonable recovery from the negative impact of model mismatch on the prior estimation of measurable states and provided an adequate estimation of the temperatures at these locations of the gasifier. Figures 3-10(d)-(f) shows that the estimation provided by KF in *Case A* and *C* are acceptable since they both follow the plant outputs. Note that the plant outputs shown in these plots correspond to those obtained for *Case A*. The plant outputs obtained for *Case C* are not shown for brevity but they exactly follow the estimations obtained for this case. In addition to the model structural error, both *Case C* and *Case A* include random process uncertainty and measurement noise thus making the plant outputs for each case slightly different. Note that measurements and the "Plant Outputs" are almost the same as this study assumed the measurement noise is very small.

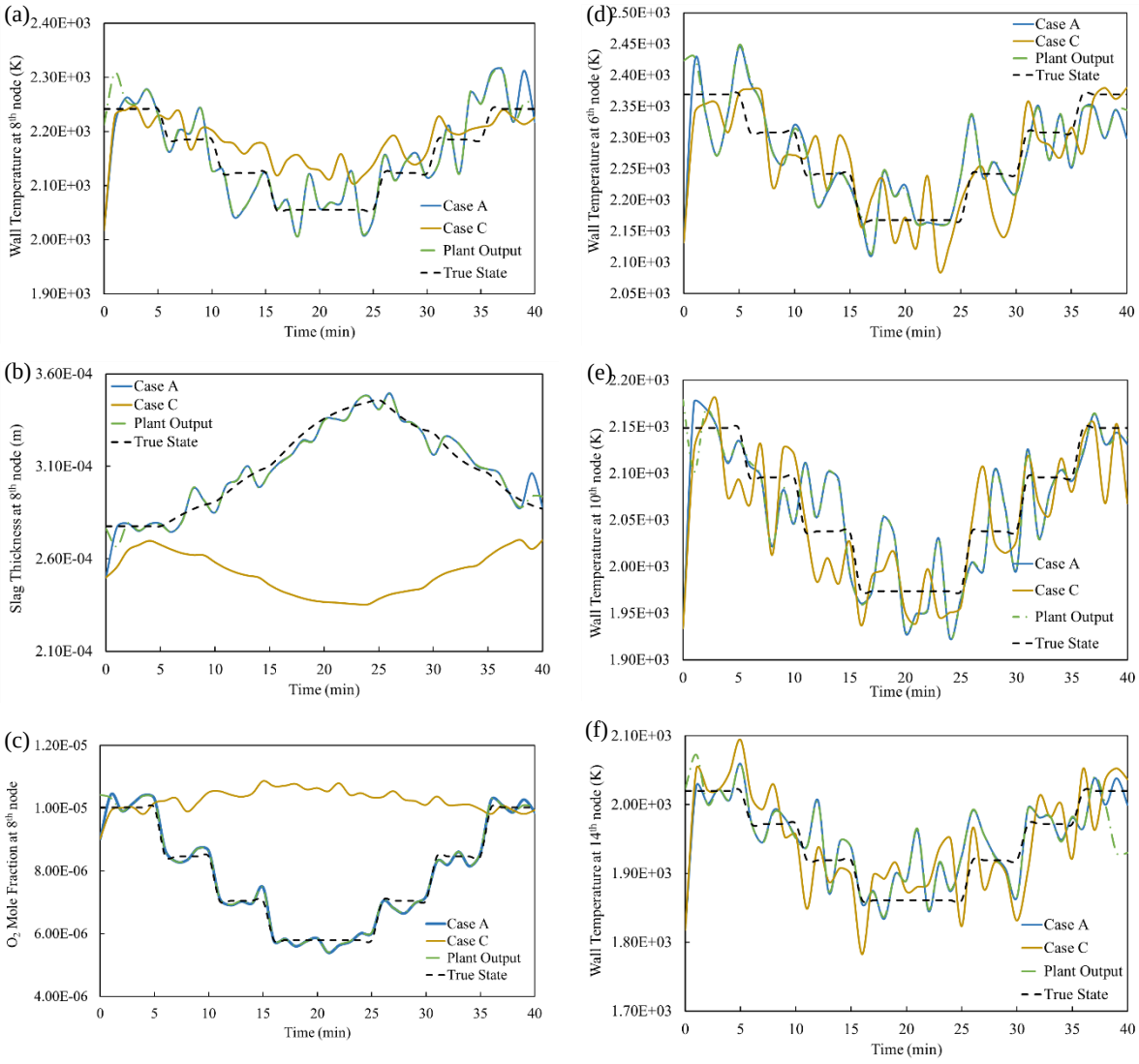


Figure 3-10: Estimations of measurable states in Case A in the presence of model mismatch for (a) wall temperature at 8th node; (b) slag thickness at 8th node; (c) O_2 mole fraction at 8th node; Estimations of measurable states in both Case A and Case C in the presence of model mismatch for (d) wall temperature at 6th node; (e) wall temperature at 10th node; (f) wall temperature at 14th node

3.4 Summary

In this chapter, the performance of KF for a pilot-scale gasifier system was performed. The transient behaviour of a pilot-scale gasification unit is represented using a dynamic reduced order model (see section 3-1). This model consists of 479 state variables including molar fractions for the species, temperature, and slag thickness across the gasifier. The quality of the state estimation provided by KF

has been evaluated under multiple arrangements of the number and the location of the sensors available for the top section of the gasifier. Also, the performance of KF for such a complex and highly nonlinear system has been explored under different scenarios. The outcome of *Scenarios I* and *II* showed that, if KF has access to an accurate linear state-space model, it is capable of providing acceptable state estimations for the gasification unit even in the presence of a realistic range of additive uncertainty in the prior estimation of the states. However, under *Scenario III*, KF was not able to provide acceptable state estimation for several states of the gasification system including the peak temperature. This was mostly due to the propagation of the error between the measurements and the inaccurate prior estimation provided by the linear state-space model corrupted with a model structural error. The poor performance of the KF under the plant-model mismatch scenario indicate the critical role of the mechanistic linear model in the KF framework. In addition, the current study assessed the impact of sensor location on KF performance, which is a critical factor when there is a significant plant-model mismatch. In general, the quality of the KF estimation can be improved as the number of sensors located within the process plant increases and an acceptable linear state space model is available.

The results show that as long as KF is provided with an accurate mechanistic model, it is capable of estimating the unknown states for a large variety of changes in the gasifier's inputs, even though online temperature sensors are only available in limited locations across the gasifier. Nevertheless, the estimation performed in this chapter was under the Gaussianity assumption typically considered in Kalman filters. The high performance of Kalman filter observed in the current study along with the proven track record in the literature for this approach is the main motivation to develop new KF-based schemes that can perform state estimation for general processes featuring constraints on the states, and non-Gaussian process uncertainties and measurement noises.

Chapter 4

Abridged Gaussian Sum Extended Kalman Filter (AGS-EKF)

Despite the outstanding capabilities of KF/EKF highlighted in the literature review (Chapter 2) and Chapter 3, the Gaussian assumption considered in this framework restricts the use of this approach for the non-Gaussian applications. The non-Gaussianity often appears in different variables in the industrial process plants and unit operations. This may happen due to the presence of bounds on the system's variables (states, inputs, measurements, process uncertainties, measurement noises) to ensure feasibility of the process model, or/and due to the scheduled/unscheduled changes during the plant operation. Following these events, KF/EKF may lose performance.

GSF is an EKF-based estimation scheme that is best known for the non-Gaussian applications. As shown in section 2.2, the development of GSFs for applications in which non-zero mean and non-Gaussian process uncertainty is involved poses challenges that have not been widely explored in the open literature. In addition, GSF requires to perform a series of EKFs at every time interval in the online estimation scheme, i.e., one EKF for each of ng Gaussian components included in the Gaussian mixture where ng denotes the number of Gaussian components needed to form the Gaussian mixture model. Therefore, the required number of evaluations in GSF is at least ng times larger than in EKF. This may make GSF intensive for complex nonlinear dynamic chemical engineering systems that may include a large number of states to estimate.

Based on the above, the non-Gaussian uncertainty scenario is lacking for online estimation by GSF and other constrained-EKF methods presented in the literature. Motivated by this fact, the focus of this chapter is on improving the estimation of EKF when the plant model is subject to process uncertainties with non-zero mean, non-Gaussian distributions. This chapter pursues the following objectives:

- A modification to the standard EKF framework is developed to improve the accuracy in the estimation for systems that involve non-zero mean Gaussian process uncertainty. The modified EKF requires the same computational costs as the standard EKF. The modified EKF formulation will be used as a basis to design the estimation schemes used in this work to deal with non-zero mean non-Gaussian process uncertainties.
- A new state estimation framework, referred to as *Abridged Gaussian Sum Extended Kalman Filter* (AGS-EKF) is presented. Contrary to the conventional estimation methods such as GSF, the proposed AGS-EKF does not consider different scenarios that limits the estimation scheme to draw

the process uncertainty variables only from one of the Gaussian components in the mixture. Instead, the proposed approach draws the process uncertainties from the overall Gaussian mixture involving all the Gaussian components. Thus, AGS-EKF does not required to perform a set of EKFs thus avoiding the additional computational costs incurred when using GSF. That is, AGS-EKF performs the modified EKF (mentioned above) only once to compute the estimation.

- In addition, this study presents an adaptation of the conventional GSF, referred to as *adaptive GSF* (AGSF). This approach makes use of the overall point estimates as a priori information to improve the estimations provided by each of the EKFs in the set of GSF. Note that both GSF and AGSF use the modified EKF introduced in this work to handle systems with non-zero mean process uncertainties.
- The performance of the Abridged GS-EKF, GSF, and Adaptive GSF have been assessed and compared using case studies emerging in the chemical engineering literature that have been widely used in the field of state estimation, i.e., an synthetic illustrative example [106], a gas-phase reactor [3,84,97,98,155], the Williams-Otto reactor [156,157], and a wastewater treatment plant[158,159,37].

The next section presents the proposed framework for the adaptive GSF (AGSF) and AGS-EKF. Section 4.2 presents the results and detailed discussions on the performance of AGS-EKF, GSF, AGSF, and standard EKF for each of the case studies considered in this research. A chapter summary is presented in the end.

4.1 Filters for non-Gaussian process uncertainty

This section presents the proposed AGS-EKF approach and other estimations schemes used in this study to evaluate the performance of AGS-EKF. As mentioned in the introduction of this chapter, AGS-EKF uses both a modified version of EKF and the Gaussian sum technique to improve the performance of EKF. Thus, this study first presents the modified EKF and the main features of Gaussian sum. To pursue this goal, section 4.1.1 presents a modification in the priori estimation step of EKF algorithm for the case of nonlinear applications with non-zero mean Gaussian process uncertainty. A brief review on Gaussian Sum is presented in section 4.1.2. Using the information provided in sections 4.1.1 and 4.1.2, the GSF adopted to non-Gaussian process uncertainty scenarios and an adaptive version of GSF are presented in sections 4.1.3 and 4.1.4, respectively. Section 4.1.5 introduces the proposed estimation scheme, i.e., AGS-EKF.

4.1.1 Modified EKF for the application with non-zero mean Gaussian process uncertainty

Considering the nonlinear model described by Equations (3-1)-(3-3), Equations (3-4) and (3-5) represent the standard EKF for a nonlinear system with zero-mean Gaussian process uncertainties as shown in Equation (3-3). A modification in the prior estimation step in EKF is needed for the case when the nonlinear system is subjected to non-zero mean, Gaussian random process uncertainties, i.e., $\boldsymbol{\xi}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{Q})$. A nonlinear process model involving non-zero mean process uncertainties $\boldsymbol{\xi}$ can be described as follows:

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{\xi}_k, k) \approx \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{G}_k \boldsymbol{\xi}_k$$
$$\text{where } \boldsymbol{G}_k = \frac{\partial f(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \Big| \boldsymbol{x} = \boldsymbol{x}_k, \boldsymbol{u} = \boldsymbol{u}_k, \boldsymbol{\xi} = \boldsymbol{\xi}_k \tag{4-1}$$

$$\mathbb{E}\left[(\boldsymbol{\xi}_\varsigma - \boldsymbol{\mu})(\boldsymbol{\xi}_\varsigma - \boldsymbol{\mu})^T\right] = \begin{cases} \boldsymbol{Q}, & \varsigma = k \\ \boldsymbol{0}, & \varsigma \neq k \end{cases}; \tag{4-2}$$

$$\boldsymbol{\xi} \in \mathbb{R}^{n_w}, \boldsymbol{x} \in \mathbb{R}^{n_x}, \boldsymbol{u} \in \mathbb{R}^{n_u}, \boldsymbol{Q} \in \mathbb{R}^{n_w \times n_w}, \boldsymbol{G}_k \in \mathbb{R}^{n_x \times n_w}, f: \mathbb{R}^{n_x \times n_u \times n_w} \to \mathbb{R}^{n_x}$$

where n_w represents the number of process uncertainty variables considered in the system. Matrix $\boldsymbol{G}_k$ is the sensitivity matrix at time interval k that relates the process uncertainties to the states of the system. The prior estimation of the states and their corresponding covariance matrix is as follows:

$$\widehat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{G}_k \boldsymbol{\mu}$$
$$\boldsymbol{P}_{k+1|k} = \boldsymbol{A}_k \boldsymbol{P}_{k|k} \boldsymbol{A}_k^T + \boldsymbol{G}_k \boldsymbol{Q} {\boldsymbol{G}_k}^T \tag{4-3}$$

The measurement model in this system is same as the model presented in Equation (3-2). Likewise, the posterior estimation step in the modified EKF is given by Equation (3-5). Similar to the standard EKF, $\mathbb{E}\left[\boldsymbol{\xi}_\varsigma \boldsymbol{x}_0^T\right] = 0$ for all k.

The current work presented in this chapter considers that each state variable of the system is associated with one process uncertainty variable. That is, $n_w = n_x$ and matrix $\boldsymbol{G}_k$ can be expressed by an identity matrix of proper dimensions, i.e., $\boldsymbol{G}_k = \boldsymbol{I}_{n_x \times n_x}$. Given this assumption, Equation (4-3) can be simplified to Equation (4-4), as follows:

$$\widehat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{\mu} \tag{4-4}$$

$$\boldsymbol{P}_{k+1|k} = \boldsymbol{A}_k \boldsymbol{P}_{k|k} \boldsymbol{A}_k^T + \boldsymbol{Q}$$

Note that Equation (4-4) is a more general representation of Equation (3-4), i.e., Equation (4-4) reduces to Equation (3-4) when $\boldsymbol{\mu} = \boldsymbol{0}$. The derivation of prior estimation presented in Equation (4-3) is presented next.

Derivation of the prior estimation: The prior estimation of the states, i.e., the expected value of the prior PDF, can be expressed as follows:

$$\widehat{\boldsymbol{x}}_{k+1|k} = \mathbb{E}[f(\widehat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k, \boldsymbol{\xi}_k)|\boldsymbol{Y}_k] \tag{4-5}$$

where: $\boldsymbol{Y}_k = [\boldsymbol{y}_0, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k]$

Equation (4-5) can be re-written as follows:

$$\begin{aligned} \widehat{\boldsymbol{x}}_{k+1|k} &= \mathbb{E}[\boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{G}_k \boldsymbol{\xi}_k | \boldsymbol{Y}_k] \\ &= \boldsymbol{A}_k \mathbb{E}[\widehat{\boldsymbol{x}}_{k|k}|\boldsymbol{Y}_k] + \boldsymbol{B}_k \mathbb{E}[\boldsymbol{u}_k|\boldsymbol{Y}_k] + \boldsymbol{G}_k \mathbb{E}[\boldsymbol{\xi}_k|\boldsymbol{Y}_k] \end{aligned} \tag{4-6}$$

Sensitivity matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and $\boldsymbol{G}$ are independent from the previous measurements. Moreover, input variables and process uncertainties are uncorrelated with the measurements. Hence $\widehat{\boldsymbol{x}}_{k+1|k}$ is obtained as shown in Equation (4-3). To derive the prior covariance matrix $\boldsymbol{P}_{k+1|k}$, the expected value of the error in the estimation needs to be evaluated. Let $\boldsymbol{e}_{k+1|k}$ represent the prior estimation error:

$$\boldsymbol{e}_{k+1|k} = \boldsymbol{x}_{k+1} - \widehat{\boldsymbol{x}}_{k+1|k} \tag{4-7}$$

By substituting Equations (4-1) and (4-6) in Equation (4-7), the prior estimation error is as follows:

$$\begin{aligned} \boldsymbol{e}_{k+1|k} &= \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{G}_k \boldsymbol{\xi}_k - \left(\boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{G}_k \boldsymbol{\mu}\right) \\ &= \boldsymbol{A}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\right) + \boldsymbol{G}_k (\boldsymbol{\xi}_k - \boldsymbol{\mu}) \end{aligned} \tag{4-8}$$

Therefore, the prior estimation of the states covariance is as follows:

$$\begin{aligned} \boldsymbol{P}_{k+1|k} &= \mathbb{E}\left[\boldsymbol{e}_{k+1|k} \boldsymbol{e}^T{}_{k+1|k}\right] \\ &= \mathbb{E}[(\boldsymbol{A}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\right) + \boldsymbol{G}_k (\boldsymbol{\xi}_k - \boldsymbol{\mu}))(\boldsymbol{A}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\right) + \boldsymbol{G}_k (\boldsymbol{\xi}_k - \boldsymbol{\mu}))^T] \end{aligned} \tag{4-9}$$

Note that $\boldsymbol{e}_{k|k} = \boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}$ is the posterior estimation error at time interval k. Therefore, the prior covariance at time interval $k + 1$ is as follows:

$$\boldsymbol{P}_{k+1|k} =$$
$$\mathbb{E}[(\boldsymbol{A}_k \boldsymbol{e}_{k|k} + \boldsymbol{G}_k(\boldsymbol{\xi}_k - \boldsymbol{\mu}))(\boldsymbol{e}^T{}_{k|k}\boldsymbol{A}^T{}_k + (\boldsymbol{\xi}_k - \boldsymbol{\mu})^T \boldsymbol{G}_k{}^T)\,]$$
$$= \mathbb{E}[\boldsymbol{A}_k \boldsymbol{e}_{k|k}\boldsymbol{e}^T{}_{k|k}\boldsymbol{A}^T{}_k + \boldsymbol{G}_k(\boldsymbol{\xi}_k - \boldsymbol{\mu})\boldsymbol{e}^T{}_{k|k}\boldsymbol{A}^T{}_k + \boldsymbol{A}_k \boldsymbol{e}_{k|k}(\boldsymbol{\xi}_k - \boldsymbol{\mu})^T \boldsymbol{G}_k{}^T$$
$$+ \boldsymbol{G}_k(\boldsymbol{\xi}_k - \boldsymbol{\mu})(\boldsymbol{\xi}_k - \boldsymbol{\mu})^T \boldsymbol{G}_k{}^T] \quad (4\text{-}10)$$

Since $\mathbb{E}[\boldsymbol{\xi}_k - \boldsymbol{\mu}] = \mathbb{E}[\boldsymbol{\xi}_k] - \mathbb{E}[\boldsymbol{\mu}] = \boldsymbol{\mu} - \boldsymbol{\mu} = \boldsymbol{0}$, and $\mathbb{E}[(\boldsymbol{\xi}_k - \boldsymbol{\mu})(\boldsymbol{\xi}_k - \boldsymbol{\mu})^T] = \boldsymbol{Q}$, Equation (4-10) can be simplified as a function of Jacobian matrix, posterior estimation of the covariance matrix at time interval k, and the covariance of the process uncertainty distribution, as given by Equation (4-3).

4.1.2 Gaussian Mixture Model

Based on the developments presented in the section 4.1.1, an efficient estimation scheme can be developed for nonlinear dynamic systems involving non-Gaussian process uncertainty. A key idea in this work is to assume that the non-Gaussian distribution can be approximated by a set of Gaussian distributions. To pursue this goal, this work considers a Gaussian mixture as an approximation of the non-Gaussian process uncertainty PDF.

Gaussian mixture model is an attractive method that provides an accurate approximation for any arbitrary non-Gaussian PDF. The Gaussian mixture model framework has been widely used in the context of Bayesian estimation and monitoring for applications involving non-Gaussianity[160,161,162,163]. This model approximates the non-Gaussian density to a summation over ng Gaussian densities. In general, the number of required Gaussian components in the mixture (ng) increases as the level of non-Gaussianity of the original distribution increases. For a vector-valued non-Gaussian variables $\boldsymbol{\varpi}$, the Gaussian mixture density can be evaluated as follows:

$$p(\boldsymbol{\varpi}) = \sum_{i=1}^{ng}\left(\boldsymbol{\alpha}^i\, \mathcal{N}\left[\boldsymbol{\varpi};\, \boldsymbol{\eta}^i, \boldsymbol{M}^i\right]\right);$$
$$\boldsymbol{\alpha}^i, \boldsymbol{\varpi} \in \mathbb{R}^d; \boldsymbol{\eta}^i \in \mathbb{R}^d; \boldsymbol{M}^i \in \mathbb{R}^{d\times d} \quad (4\text{-}11)$$

$$\sum_{i=1}^{ng} \boldsymbol{\alpha}^i = \boldsymbol{1}; \quad \boldsymbol{\alpha}^i \geq \boldsymbol{0}; \quad (4\text{-}12)$$

where $\boldsymbol{\eta}^i, \boldsymbol{M}^i$ are the mean vector and covariance matrix for the i^{th} Gaussian component in the mixture. Moreover, symbol $\odot$ expresses the element-wise multiplication. The symbol d denotes the dimension of the Gaussian component distributions. Any random variable $\boldsymbol{\varpi}$ drawn from the original non-Gaussian distribution has a chance (weight) of $\boldsymbol{\alpha}^i$ to belong to the i^{th} Gaussian component in the Gaussian mixture model. Hence, $\boldsymbol{\alpha}^i$ is used in the Gaussian mixture model as the corresponding weights

for each Gaussian component in the model. These weights are normalized; thus, for each of the d variables, the summation of the weights over all the Gaussian components present in the mixture adds to the unity, as shown in constraint (4-12). In this study, the mean vector $\boldsymbol{\eta^{GM}} \in \mathbb{R}^d$ and the covariance matrix $\boldsymbol{M^{GM}} \in \mathbb{R}^{d\times d}$ of the Gaussian mixture models that approximate each of the non-Gaussian distributions of process uncertainties are considered in the proposed AGS-EKF framework. Note that the superscript $\boldsymbol{GM}$ stands for *Gaussian mixture*. Equation (4-13) presents the analytical evaluation of the mean vector and the covariance matrix of the Gaussian mixture model.

$$\begin{aligned}\boldsymbol{\eta^{GM}} &= \sum_{i=1}^{ng} \boldsymbol{\alpha}^i \odot \boldsymbol{\eta}^i \\ \boldsymbol{M^{GM}} &= \sum_{i=1}^{ng} \boldsymbol{\alpha}^i \odot \boldsymbol{M}^i + \sum_{i=1}^{ng} \boldsymbol{\alpha}^i \odot \left(\boldsymbol{\eta}^i - \boldsymbol{\eta^{GM}}\right)\left(\boldsymbol{\eta}^i - \boldsymbol{\eta^{GM}}\right)^{\boldsymbol{T}}\end{aligned} \tag{4-13}$$

For a known non-Gaussian distribution, $\boldsymbol{\eta}^i$, $\boldsymbol{M}^i$, and $\boldsymbol{\alpha}^i$ are parameters that should be specified accurately to ensure the accuracy of the Gaussian mixture model. This work considers the Expectation-Maximization (EM) algorithm[164] to estimate those parameters[164,165,166,167]. A brief description of the EM algorithm is presented in appendix A.

4.1.3 An Adopted Gaussian Sum Filter (GSF) for Non-Gaussian Process Uncertainties

All the components in the Gaussian mixture (i.e., Equations (4-11) and (4-12)) are non-zero mean Gaussian. Therefore, the modified EKF shown in Equation (4-4) can be used to estimate the prior states PDF with respect to each of the ng components in the Gaussian mixture model of the process uncertainties. As these components are independent, GSF runs a set of parallel EKFs considering the i^{th} component of the Gaussian mixture at every time interval, as shown in Equation (4-14). The posterior estimation for each component follows the standard EKF scheme, as in Equation (3-5). Based on the conventional formulation of GSF[93,96], the adopted GSF approach is as follows:

Prior estimation for the i^{th} component

$$\begin{aligned}\hat{\boldsymbol{x}}^i_{k+1|k} &= \boldsymbol{A}^i_k \hat{\boldsymbol{x}}^i_{k|k} + \boldsymbol{B}^i_k \boldsymbol{u}^i_k + \boldsymbol{\mu}^i \\ \boldsymbol{P}^i_{k+1|k} &= \boldsymbol{A}^i_k \boldsymbol{P}^i_{k|k} \boldsymbol{A}^T_k + \boldsymbol{Q}^i\end{aligned} \tag{4-14}$$

Posterior estimation for the i^{th} component

$$\begin{aligned}\boldsymbol{K}^i_{k+1} &= \boldsymbol{P}^i_{k+1|k} {\boldsymbol{H}^i_{k+1}}^{\boldsymbol{T}} / (\boldsymbol{H}^i_{k+1} \boldsymbol{P}^i_{k+1|k} {\boldsymbol{H}^i_{k+1}}^{\boldsymbol{T}} + \boldsymbol{R}) \\ \hat{\boldsymbol{x}}^i_{k+1|k+1} &= \\ &\quad \hat{\boldsymbol{x}}^i_{k+1|k} + \boldsymbol{K}^i_{k+1}\left(\boldsymbol{y}_{k+1} - h\left(\hat{\boldsymbol{x}}^i_{k+1|k}, \boldsymbol{u}^i_k\right)\right)\end{aligned} \tag{4-15}$$

$$\boldsymbol{P}^i_{k+1|k+1} = \left(\boldsymbol{I} - \boldsymbol{K}^i_{k+1}\boldsymbol{H}^i_{k+1}\right)\boldsymbol{P}^i_{k+1|k}$$

Point estimate for the system states:

$$\widehat{\boldsymbol{x}}^{\boldsymbol{GSF}}_{k+1|k+1} = \sum_{i=1}^{ngt} \beta^i \widehat{\boldsymbol{x}}^i_{k+1|k+1} \tag{4-16}$$

$$\boldsymbol{P}^{\boldsymbol{GSF}}_{k+1|k+1} = \sum_{i=1}^{ngt} \beta^i \left[\boldsymbol{P}^i_{k+1|k+1} + \left(\widehat{\boldsymbol{x}}^i_{k+1|k+1} - \widehat{\boldsymbol{x}}^{\boldsymbol{GSF}}_{k+1|k+1}\right)\left(\widehat{\boldsymbol{x}}^i_{k+1|k+1} - \widehat{\boldsymbol{x}}^{\boldsymbol{GSF}}_{k+1|k+1}\right)^T\right] \tag{4-17}$$

$$\sum_{i=1}^{ngt} \beta^i = 1; \quad \beta^i = \prod_{l=1}^{n_x} \alpha_l^{ip(l)}; \quad \alpha_l^{ip(l)}, \beta^i \geq 0; \; \forall ip = 1, \dots, ng(l) \tag{4-18}$$

Note that GSF performs a set of EKFs for all possible combinations of the Gaussian components in the Gaussian mixture models representing the process uncertainties. That is, for a system with n_x process uncertainty variables, GSF requires a total number of EKFs given by $ngt = \prod_{l=1}^{n_x} ng(l)$. Note that $ng(l)$ denotes the number of Gaussian components in the Gaussian mixture model corresponding to process uncertainty l. This means that GSF needs to consider all the possible combinations between the Gaussian components of each process uncertainty l. That is, GSF increases the required computational time by $\prod_{l=1}^{n_x} ng(l)$ when compared to the standard EKF. Moreover, β^i is the weight assigned to the i^{th} EKF in the set of GSF to represent the posterior estimation of the states at each time interval. In addition, $\alpha_l^{ip(l)}$ represents the weight corresponding to the ip^{th} component in the Gaussian mixture model of the l^{th} process uncertainty variable. $ip(l)$ denotes that the index of the Gaussian component for each process uncertainty can be different at each instance i. Note that for the applications where all the process uncertainties involved in the process follow the same PDF, the number of instances required to run EKFs would reduce to the number of Gaussian components in the representative Gaussian mixture model of this PDF, i.e., ng.

4.1.4 An Adaptive Gaussian Sum Filter (AGSF)

One limitation in GSF is that it runs individual EKFs based on a single Gaussian component in the Gaussian mixture. That is, the i^{th} EKF in the set evaluates the prior estimation $\widehat{\boldsymbol{x}}^i_{k+1|k}$ and $\boldsymbol{P}^i_{k+1|k}$ based on the posterior PDF that resulted from performing i^{th} EKF from the previous time interval k, i.e., $\widehat{\boldsymbol{x}}^i_{k|k}$ and $\boldsymbol{P}^i_{k|k}$. This may lead to an accumulative error and eventually to an inaccurate state estimation in GSF. In particular, when the process uncertainty distribution is highly nonlinear or multi-modal,

performing independent EKFs and assuming that the process uncertainties are obtained from one Gaussian density in the mixture might be a biased assumption and lead to inaccurate estimations. This issue can be prevented by updating the prior estimation for each i^{th} EKF, based on the available information of the overall posterior density of the states at time k, i.e., $\widehat{\boldsymbol{x}}_{k|k}^{\boldsymbol{GSF}}$ and $\boldsymbol{P}_{k|k}^{\boldsymbol{GSF}}$, given in Equations (4-16) and (4-17). This method is referred to as *Adaptive Gaussian Sum Filter* (AGSF). The prior estimation in AGSF is estimated as follows:

Prior estimation for the i^{th} component

$$\begin{aligned}\widehat{\boldsymbol{x}}_{k+1|k}^{i} &= \boldsymbol{A}_k^i \widehat{\boldsymbol{x}}_{k|k}^{\boldsymbol{GSF}} + \boldsymbol{B}_k^i \boldsymbol{u}_k^i + \boldsymbol{\mu}^i \\ \boldsymbol{P}_{k+1|k}^{i} &= \boldsymbol{A}_k^i \boldsymbol{P}_{k|k}^{\boldsymbol{GSF}} \boldsymbol{A}_k^T + \boldsymbol{Q}^i\end{aligned} \tag{4-19}$$

The "Posterior estimation" and the "Point estimate for states" are similar to the GSF algorithm presented in Equations (4-15)-(4-18).

4.1.5 Abridged Gaussian Sum Extended Kalman Filter (AGS-EKF)

Using the characteristics of the Gaussian mixture model presented in Equation (4-13), the proposed Abridged Gaussian Sum Extended Kalman Filter (i.e., AGS-EKF), avoids performing ng individual EKFs. To pursue this goal, AGS-EKF assumes that the random process uncertainty variable $\boldsymbol{\varpi}_k$ that follows an arbitrary non-Gaussian distribution can be approximated by a probability density that has the mean value and the covariance matrix as in a representative Gaussian mixture with model parameters defined as in Equation (4-13). That is, $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{GM}}, \boldsymbol{Q}^{\boldsymbol{GM}})$ for all k. Thus, given the modified EKF for non-zero mean Gaussian distributed process uncertainty (see section 4.1.1), the proposed AGS-EKF framework can be formulated as follows:

Prior estimation

$$\begin{aligned}\widehat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{\mu}^{\boldsymbol{GM}} \\ \boldsymbol{P}_{k+1|k} &= \boldsymbol{A}_k \boldsymbol{P}_{k|k} \boldsymbol{A}_k^T + \boldsymbol{Q}^{\boldsymbol{GM}}\end{aligned} \tag{4-20}$$

Posterior estimation

$$\begin{aligned}\boldsymbol{K}_{k+1} &= \boldsymbol{P}_{k+1|k} \boldsymbol{H}_{k+1}^T / (\boldsymbol{H}_{k+1} \boldsymbol{P}_{k+1|k} \boldsymbol{H}_{k+1}^T + \boldsymbol{R}) \\ \widehat{\boldsymbol{x}}_{k+1|k+1} &= \\ &\quad \widehat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1} \left(\boldsymbol{y}_{k+1} - h(\widehat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k) \right)\end{aligned} \tag{4-21}$$

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

where $\widehat{\boldsymbol{x}}_{k|k} \in \mathbb{R}^{n_x}$ is the posterior estimation of states at time interval k. The proposed approach offers the same number of calculations (i.e., required computational cost) as in standard EKF. As a result, AGS-EKF can be an efficient scheme to improve the EKF performance, in particular for practical applications in chemical engineering. The performance of the proposed approach is compared with GSF, AGSF and standard EKF performances in section 4.2.

4.2 Computational Experiments

The current study performs state estimation using the proposed AGS-EKF scheme introduced in section 4.1.5. In this section, the performance of the proposed method is compared with the estimations provided by standard EKF, GSF, and AGSF presented in sections 3.2, 4.1.3, and 4.1.4, respectively. The mean squared error (MSE) has been used as the main metric to perform such comparisons. The MSE for the m^{th} state using the n^{th} state estimation method is defined as follows:

$$MSE_{x_m}^{(n)} = \frac{1}{t_f}\sum_{k=0}^{t_f}\left(\hat{x}_{k,m}^{(n)} - x_{k,m}\right)^2 \tag{4-22}$$

where k is the time index; t_f is the final time interval considered in the experiments; the subscript m is the index of state in the vector of states $\boldsymbol{x}$, i.e., $m \in \{1, 2, \dots, n_x\}$; superscript n denotes the estimation method, i.e., $n \in \{\text{EKF}, \text{GSF}, \text{AGSF}, AGS-EKF\}$; $\hat{x}_{k,m}^{(n)}$ and $x_{k,m}$ are scalars that provide the estimated and true values for the m^{th} state at each time interval k using the estimation method n, respectively. The computational experiments 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 discretization of the nonlinear dynamic process models was performed using the backward method. The performance of the proposed AGS-EKF was tested for a motivating example, a gas-phase reactor, the Williams-Otto reactor, and a wastewater treatment plant (WTP). The results obtained for each of these case studies are presented next.

4.2.1 A Second-order Mathematical Example

The first case study considered in this work is a case study adopted from [106] that is subjected to a constraint on the process uncertainty. This case study was selected from the literature as it has been shown that the standard EKF could not provide an accurate estimation for the constrained linear

model[106]. Given that this case study is a linear model, the sensitivity matrices $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ remain constant throughout time. Consequently, the standard EKF formulation becomes equivalent to the standard KF. Likewise, GSF and AGS-EKF are performed assuming a KF. However, to keep consistency in this study, the discussion presented in this section uses the term EKF when making a reference to those estimator schemes. The case study is a dimensionless constrained linear system defined as follows:

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \begin{bmatrix} 0.99 & 0.2 \\ -0.1 & 0.3 \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varpi_k; \\ y_k &= [1 \quad -3]\boldsymbol{x}_k + v_k; \end{aligned} \tag{4-23}$$

$$\varpi_k \sim \mathcal{N}(0, Q);\ Q = 1;\ v_k \sim \mathcal{N}(0, R);\ R = 0.01;\ \boldsymbol{x}_0 = \begin{bmatrix} x_{1_0} \\ x_{2_0} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix};\ \boldsymbol{P}_0 = \boldsymbol{I}$$

$$\varpi_k \geq 0 \tag{4-24}$$

where ϖ and v are the process uncertainty and the measurement noise signal, respectively. Both states in the states vector $\boldsymbol{x}$, i.e., x_1 and x_2, need to be estimated. Moreover, y denotes the only measurement of this system. The covariance matrix for the initial states, i.e., $\boldsymbol{P}_0$, is an identity matrix. The time interval considered for this study is set to 1. As shown in Equation (4-24), the process uncertainties are non-negative, which makes the distribution of the process uncertainties Non-Gaussian with non-zero mean. The EM method presented in [164] was used to approximate the corresponding PDF of the process uncertainty. The Gaussian mixture provided by EM is as follows:

$$\begin{aligned} p(\varpi_k) = 0.21\, \mathcal{N}[\varpi_k;\ 0.18, 0.01] + 0.44\, \mathcal{N}[\varpi_k;\ 0.66, 0.08] \\ + 0.35\, \mathcal{N}[\varpi_k;\ 1.37, 0.36] \end{aligned} \tag{4-25}$$

Note that the number of Gaussian components in the Gaussian mixture model is arbitrary and should be chosen based on the non-Gaussianity of the density. Figure 4-1 shows that the PDF obtained by the Gaussian mixture presented in Equation (4-25) is an acceptable representation of the actual bounded distribution of the process uncertainty. Adding more Gaussian functions does not seem to improve the representation of the non-Gaussian function. The preliminary tests have been performed to attain an adequate number of Gaussian components such that the Gaussian model mixture used in the state estimation schemes can represent the non-Gaussian process uncertainties, accurately.

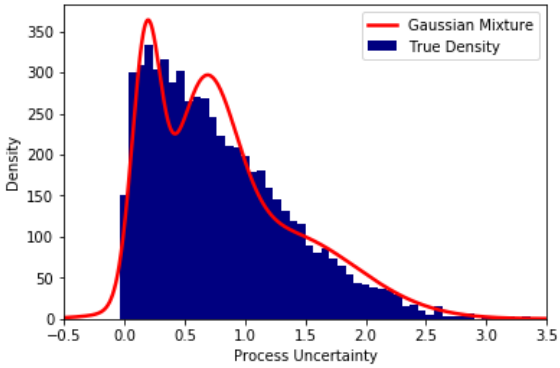


Figure 4-1: Histogram for the true non-Gaussian process uncertainty and the Gaussian mixture approximation

Figure 4-2 shows the state estimates provided by the estimation schemes presented in section 4.1 (and the standard EKF presented in section 3.2). As it shown in this figure, standard EKF was not be able to estimate the states accurately. On the other hand, the modified EKF formulation used in the GSF frameworks, i.e., GSF and AGSF, as well as the proposed AGS-EKF approach made a significant improvement in the accuracy of the estimation. In addition, the variability in the estimation of x_1 provided by AGS-EKF is smaller than that observed in AGSF. Table 4-1 presents the MSE in the estimation of x_1 and x_2 (i.e., $MSE_{x_1}^{(n)}$ and $MSE_{x_2}^{(n)}$) using the different estimation methods. As expected, standard EKF presents the largest errors in the estimation whereas the proposed approach (AGS-EKF) has the smallest MSE for x_1, i.e., both GSF and AGSF resulted in an MSE for x_1 that are 2% and 82% larger than that obtained for AGS-EKF, respectively. As for x_2, AGS-EKF exhibits a similar performance to that obtained by GSF; however, the most accurate estimation is provided by AGSF, as shown in Table 4-1. These results show that the proposed AGS-EKF schemes leads to a significant improvement in the performance of the EKF for an illustrative example involving bounded process uncertainty.

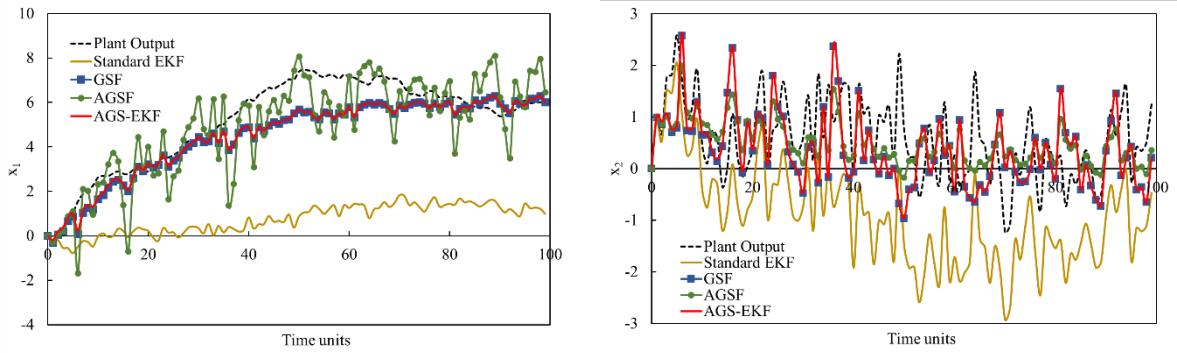


Figure 4-2: Estimation results based on the state estimation approach

Table 4-1: MSE for x_1 and x_2 using different estimation schemes

Estimation method (n)	$MSE_{x_1}^{(n)}$	$MSE_{x_2}^{(n)}$
Standard EKF	21.81	2.43
GSF	0.97	0.89
AGSF	1.78	0.56
AGS-EKF	0.95	0.89

4.2.2 Gas-phase Reactor

As the second case study, a gas-phase reactor was considered that is a typical case study in the field of state estimation to show the performance of Kalman filters in the presence of nonlinear distributions[3,84,97,98]. In this process, an irreversible reaction, i.e., $A \xrightarrow{k_r=16} B$, occurs in the Gas-phase reactor. The process is as follows:

$$\frac{dp_A}{dt} = -2k_r {p_A}^2$$
$$\frac{dp_B}{dt} = k_r {p_A}^2 \qquad (4\text{-}26)$$
$$y_k = [1 \quad 1]\left[p_{A,k}, p_{B,k}\right]^T + v_k;$$
$$v_k \sim \mathcal{N}(0, R);\ R = 0.01;\ \ p_{A_0} = 0.1;\ p_{B_0} = 4.5;\ \boldsymbol{P}_0 = \boldsymbol{I}$$

where states of the system, i.e., p_A and p_B, are the partial pressures of species *A* and *B*, respectively. Accordingly, both states can only take non-negative values. The time interval considered in this study is 0.1s. In this work, the initial conditions of the states, process uncertainties and the measurement noise signal are defined *a priori* based on the expected operating region for this process. The process uncertainty associated with the partial pressure of reactant *A* is assumed to follow a multi-modal distribution, as presented in Figure 4-3(a). Hence, a Gaussian mixture model that consists of three Gaussian components is used to approximate this multi-modal distribution. Moreover, the process

uncertainty variable associated with p_B is described with a Gamma distribution as shown in Figure 4-3(b). In order to approximate this non-Gaussian distribution with a proper Gaussian mixture model, the EM algorithm has been used to find the characteristics of each of the two Gaussian components considered for this mixture. Similar to the illustrative example, preliminary tests were performed to obtain an adequate number of Gaussian components that are required in the Gaussian mixture models of each of the process uncertainties.

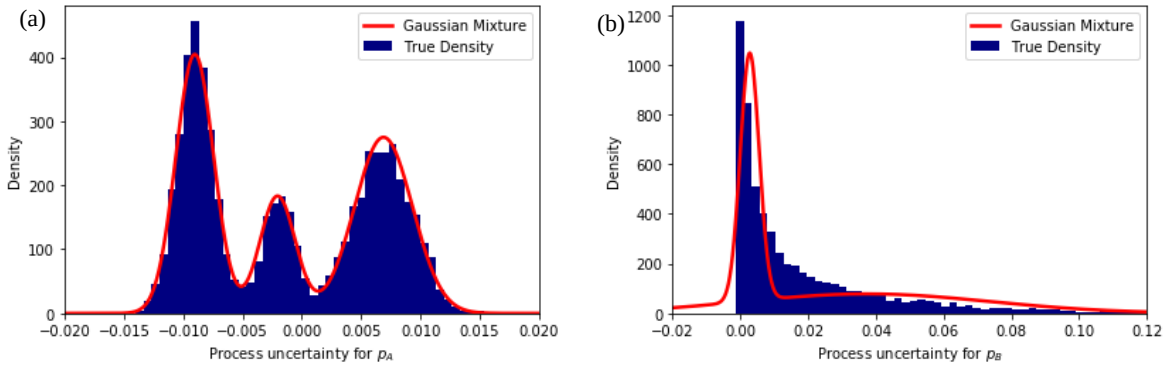


Figure 4-3: Histogram for the true non-Gaussian distribution and the Gaussian mixture approximation of the process uncertainty associated with (a) p_A; (b) p_B

For the process uncertainties shown in Figure 4-3, the characteristics of the Gaussian mixture model used in AGS-EKF algorithm are as follows:

$$\mu_A^{GM} = \sum_{i=1}^{3} \alpha_A^i \mu_A^i = -1.12e-3$$
$$\mu_B^{GM} = \sum_{i=1}^{2} \alpha_B^i \mu_B^i = 2.03e-2 \qquad (4\text{-}27)$$
$$Var_A^{GM} = \sum_{i=1}^{3} \alpha_A^i Var_A^i + \sum_{i=1}^{3} \alpha_A^i \left(\mu_A^i\right)^2 - \left(\sum_{i=1}^{3} \alpha_A^i \mu_A^i\right)^2 = 5.65e-5$$

$$Var_B^{GM} = \sum_{i=1}^{2} \alpha_B^i Var_B^i + \sum_{i=1}^{2} \alpha_B^i \left(\mu_B^i\right)^2 - \left(\sum_{i=1}^{2} \alpha_B^i \mu_B^i\right)^2 = 9.44e-4$$
$$\boldsymbol{\mu^{GM}} = [\mu_A^{GM}, \mu_B^{GM}]^T; \qquad \boldsymbol{Q^{GM}} = diag([Var_A^{GM}, Var_B^{GM}])$$

Note that ng is 3 and 2, for the uncertainty in the states p_A and p_B, respectively. On the other hand, GSF (and AGSF) algorithm runs six individual EKFs, i.e., there exists 6 different combinations of the Gaussian components in the Gaussian mixture models.

Figure 4-4 highlights the results for the gas-phase reactor. The inner graphs in Figures 4-4(a) and 4-4(b) show the actual scale of the states whereas the actual graphs present the same results for a limited scale of the states to highlight the differences in performance between the estimation schemes

considered in this work. As shown in this figure, the estimation provided by standard EKF deviates from the plant output. This inaccurate estimation was expected as standard EKF assumes that the process uncertainties follow Gaussian distribution, which cannot hold for the process uncertainties considered in this case study. Figure 4-4 also shows that the estimation provided by GSF is unstable. This is mostly because each EKF in the set of GSF only considers a small region of the actual non-Gaussian distribution of the process uncertainties. As a result, there are some EKFs in the set of GSF that evolve in the infeasible region of this process, i.e., provide negative valued estimates for the partial pressures for the state p_A. The AGSF presented in section 4.1.4 improves the estimation of the GSF significantly, as it corrects the initialization of each of the EKFs in the set of AGSF at every time interval, based on the point estimation of the states at the previous time interval. However, AGSF still needs to perform the EKF six times, which is computationally expensive when compared to standard EKF. On the other hand, the proposed AGS-EKF scheme reduces the estimation error for p_A and p_B by two and four orders of magnitude, respectively, when compared to the standard EKF (see Table 4-2), while offers the same required number of calculations as needed in standard EKF. Table 4-2 summarizes the estimation error for the various estimation schemes conducted to this study and further confirms that the proposed AGS-EKF provides acceptable estimates in short computational times.

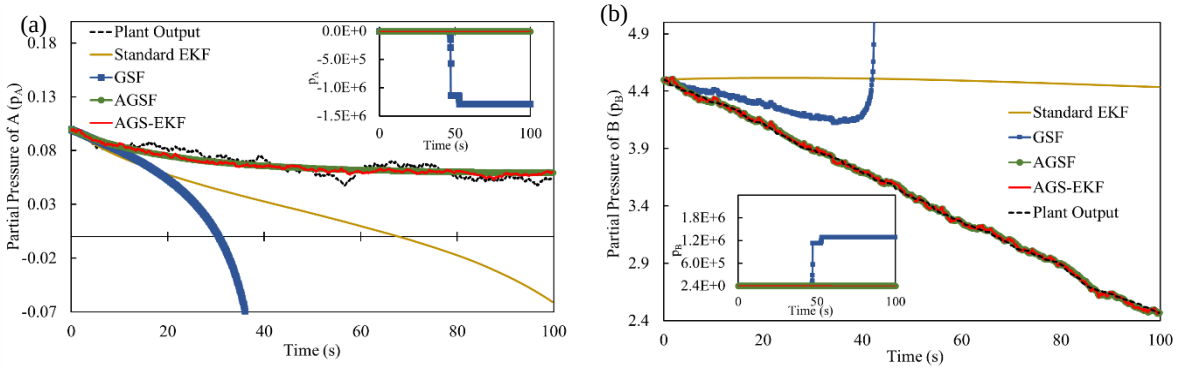


Figure 4-4: Estimation provided by various estimation schemes (a) p_A; (b) p_B

Table 4-2: MSE for p_A and p_B using different estimation schemes

Estimation method (n)	$MSE_{p_A}^{(n)}$	$MSE_{p_B}^{(n)}$
Standard EKF	1.09e-3	1.35
GSF	4.48e+11	8.48e+11
AGSF	3.14e-5	3.47e-4

AGS-EKF	2.74e-5	4.3e-4

The results presented above are based on the assumption that the initial guess in the estimation schemes is the same as the true value of the initial states. Note that the same observations obtained from Figure 4-4 and Table 4-2 (considering $\hat{\boldsymbol{x}}_0 = \boldsymbol{x}_0$) hold if the state estimation schemes are initialized at a value different than the true states ($\hat{\boldsymbol{x}}_0 \neq \boldsymbol{x}_0$). To confirm this observation, Figure 4-5 shows the state estimates provided using standard EKF, GSF, and AGS-EKF for the case that the initial estimate for each state is set to 95% of the true value of the corresponding initial state. As shown in this figure, the proposed AGS-EKF improves the estimation significantly when compared to standard EKF and GSF. Moreover, AGSF provides the same accurate estimates as given by AGS-EKF but at the expense of additional computational costs.

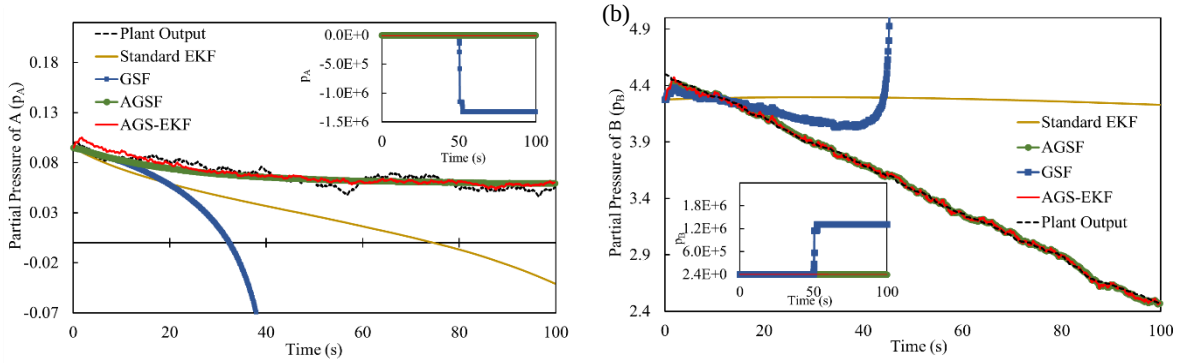


Figure 4-5: Estimation provided by various estimation schemes when initial estimates are at 95% of the true initial states (a) p_A; (b) p_B

4.2.3 Williams-Otto Reactor

A case study featuring the Williams-Otto reactor[156,157] is considered next. The Williams-Otto reactor is a highly nonlinear dynamic system that is widely used in control and state/parameter estimation studies[168,169,170,171]. Therefore, this model is a suitable candidate to test the performance of the proposed AGS-EKF framework. The following reactions take place inside the reactor:

$$A + B \xrightarrow{k_1} C \qquad k_1 = 1.6599 \times 10^6 e^{-6666.7/TR} \ \ s^{-1}$$

$$B + C \xrightarrow{k_2} P + E \qquad k_2 = 7.2117 \times 10^8 e^{-8333.3/TR} \ \ s^{-1}$$

$$C + P \xrightarrow{k_3} G \qquad k_3 = 2.6745 \times 10^{12} e^{-11111/TR} \ \ s^{-1}$$

The dynamic model for the Williams-Otto reactor is as follows:

$$W\frac{dX_A}{at} = F_A - (F_A + F_B)X_A - r_1$$

$$W\frac{dX_B}{at} = F_B - (F_A + F_B)X_B - r_1 - r_2$$

$$W\frac{dX_C}{at} = -(F_A + F_B)X_C + 2r_1 - 2r_2 - r_3$$

$$W\frac{dX_E}{at} = -(F_A + F_B)X_E + r_2 \qquad (4\text{-}28)$$

$$W\frac{dX_G}{at} = -(F_A + F_B)X_G + 1.5r_3$$

$$W\frac{dX_P}{at} = -(F_A + F_B)X_P + r_2 - 0.5r_3$$

$$r_1 = k_1X_AX_BW;\ \ r_2 = k_2X_BX_CW;\ \ r_3 = k_3X_CX_PW$$

where $W = 2104.7kg$ is the mass hold up of the reactor, $T_R = 366.05\ K$ is the temperature inside the reactor, and F_A and F_B are the mass flowrates of the reactants A and B, respectively. $F_B = 6.1\ kg/s$, and F_A is the main source of disturbances in the system and has a nominal value of $1.8\ kg/s$. Hence, a random disturbance with unit standard deviation is imposed to F_A. The state variables for the system are X_A, X_B, X_C, X_E, X_G, and X_P, which are the mass fractions of the corresponding chemical components. Moreover, online measurements are assumed to be available for X_A, X_B, and X_P, whereas X_C, X_E, and X_G are the states that need to be estimated. Note that the linear observability for the states X_A, X_B, and X_P was confirmed. That is, the Jacobian matrix for this system around the initial operating point considered in this study has been evaluated and correspondingly, the observability matrix at that initial condition was calculated. The resulting linear observability matrix is a full rank matrix (not shown for brevity). In addition, the measurement noise signals are zero-mean random variables; their corresponding standard deviations are set to 10% of the nominal steady-state values of the states. The nominal values of the states are reported in Table B-1 in the Appendix B. The Williams-Otto reactor considered in this study involves random independent process uncertainties that follow a uniform distribution. Since the states of the system are mass fractions bounded in the interval [0,1], this case study assumes that the process uncertainties associated with each of states follow a single uniform distribution. Moreover, the time interval used to discretize the nonlinear problem is set to 1 s. Although

the Williams-Otto reactor model is subjected to bounds on the states, i.e., non-negative mass fractions, these feasibility constraints are not strictly enforced in the present estimation since the main interest of this research is to focus on the non-Gaussian process uncertainties. Nevertheless, the initial conditions of the states, the additive uncertainties, disturbances, and measurement noise signals are selected to ensure feasible operation of the Williams-Otto reactor.

Figure 4-6 shows both the true and the approximated the process uncertainties associated with the state variables of the system. Part (a), (b), and (c) of this figure represent the Gaussian mixture model obtained by the EM when three, five, and ten Gaussian components were considered in the mixture, respectively. Preliminary tests have shown that a Gaussian mixture model featuring three Gaussian components is an adequate representation of the uniformed process uncertainties. Therefore, the Gaussian mixture model involving three components (i.e., Figure 4-6(a)) is used to accomplish this study. However, to make this study comprehensive, a brief discussion on the sensitivity of AGS-EKF to the other Gaussian mixture models shown in Figure 4-6 is presented at end of this section.

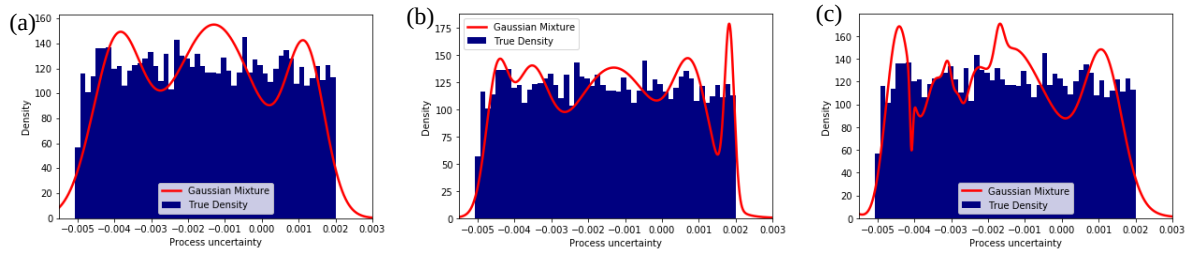


Figure 4-6: Uniform probability density of the process uncertainties in the Williams-Otto reactor compared to the approximated density by Gaussian mixture model involving (a) three Gaussian components; (b) five Gaussian components; (c) ten Gaussian components

Figures 4-7(a), (c), (e) compares the performance of AGS-EKF with standard EKF, GSF, AGSF, for estimating the states X_C, X_E, and X_G, respectively. Likewise, Figures 4-7(b), (d), (f) shows the estimates provided by each of the EKFs in the set of GSF and AGSF for X_C, X_E, and X_G, respectively. Note that GSF-G1, GSF-G2, and GSF-G3 represents the estimations provided by first, second, and third EKFs in the set of GSF, respectively. Likewise, AGSF-G1, AGSF-G2, and AGSF-G3 are the estimations resulting from performing the first, second, and third EKFs in the AGSF set, respectively. Moreover, Table 4-3 provides the corresponding MSE obtained for each estimation method.

According to the results shown in Figure 4-7(a) and Table 4-3, AGS-EKF accomplished the estimation of X_C at the smallest error compared to the rest of the estimation schemes. However, the error in the estimation of X_C provided by standard EKF is relatively small. The acceptable performance of the

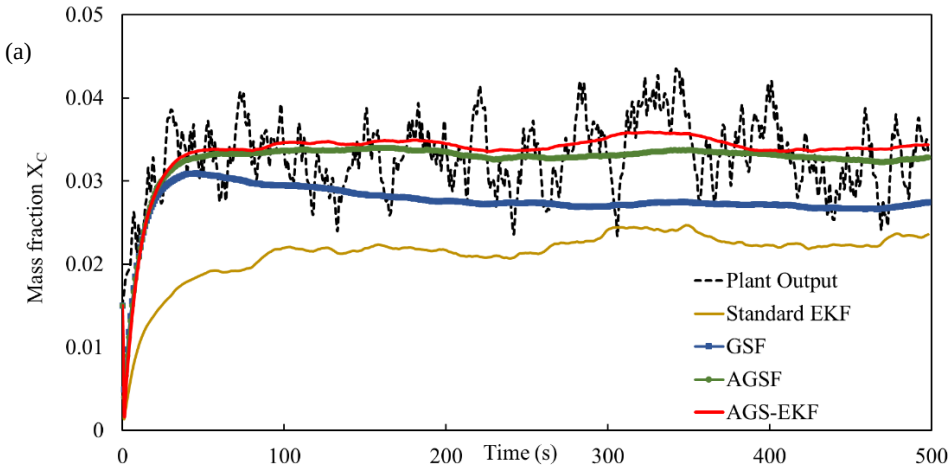
standard EKF for X_C is mostly due to the nature of the system since this state is highly related to the measurable states of the system. This can be verified by analyzing the value of the Kalman gain for this state X_C, which is in the order of 1e-4 and is relatively a large number when compared to the corresponding Kalman gain for the other state variables of the system (in the order of 1e-10). In general, the larger the Kalman gain, the more the estimation relies on the measurements and the less weight the method places on the process model (i.e., prior estimation). Thus, the standard EKF for X_C is not expected to lose performance due to the lack of a priori knowledge of the proper distribution of the uncertainties since the estimation of X_C highly depends on the available measurements in the system and not on the prior estimation.

The Kalman gain is in the order of 1e-8 to 1e-10 for the other unknown states, i.e., X_E and X_G, respectively; this implies that the estimation framework is very sensitive to the accuracy of the process model (i.e., prior estimation) and the provided Gaussian mixture for the process uncertainty variables. Therefore, the expectation is that under the uniform distribution of the process uncertainties (Figure 4-6(a)), the state estimation using the standard EKF may fail to provide adequate estimates to the unknown states in the Williams-Otto reactor. Thus, the modified EKF estimation schemes introduced in section 4.1 can be used to improve the state estimation (i.e., GSF, AGSF, AGS-EKF). As expected, Figures 4-7(c) and 4-7(e) indicate that the estimations provided by standard EKF significantly deviates from the true states. This outcome can be quantified in terms of the corresponding MSE evaluated for each estimation scheme, as shown in Table 4-3. For instance, the error in the estimation of X_E for the case of using the standard EKF as the estimator, i.e., $MSE_{X_E}^{(EKF)}$, is 8.2%, 8.3% and 8.4%higher than $MSE_{X_E}^{(GSF)}$, $MSE_{X_E}^{(AGSF)}$ and $MSE_{X_E}^{(AGS-EKF)}$, respectively. The calculated MSE for GSF, AGSF, and AGS-EKF are similar with a tolerance of 0.1%. This outcome suggests that these three approaches have the same performance in the estimation of X_E and X_G, while the required computational costs by the proposed approach (AGS-EKF) is three times smaller than that needed by GSF and AGSF. This is due to the fact that the proposed AGS-EKF needs to perform EKF only once whereas GSF and AGSF requires to perform EKF three times. As outlined above, the number of Gaussian components in the Gaussian mixture representing the process uncertainty density in this case study is set to 3; hence the corresponding reduction in the number of calculations performed by AGS-EKF.

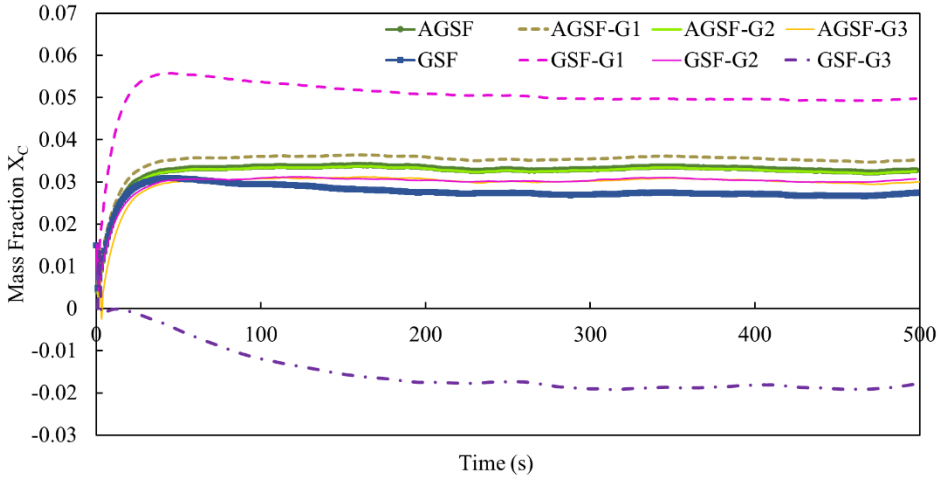
Note that for the mass fractions of components *C*, *E* and *G*, i.e., Figures 4-7(a), 4-7(c) and 4-7(e), AGSF improved the estimations of each of the EKFs in its set (when compared to GSF). The reason of this improvement can be explained by looking into the estimations provided by each of the EKFs in the set

of GSF and AGSF. As shown in Figures 4-7(b), 4-7(d), and 4-7(f), the third EKF in the GSF's set (denoted as GSF-G3) provides negative valued estimations for the mass fractions for the chemical compositions, which is an infeasible estimation. Moreover, GSF-G1 resulted in estimations that are greater than 1 in the estimation of the states X_E and X_G, which is an infeasible value for mass fraction. That is, the random uncertainties that are explicitly represented by the first and the third Gaussian component in the Gaussian mixture deviate the process to an infeasible region, hence the infeasible estimation resulted by GSF-G1 and GSF-G3. On the other hand, AGSF is able to reduce the effect of these infeasible estimates on the overall point estimates through a modification in the a priori information of the individual EKFs in the AGSF set. That is, in the AGSF framework, the infeasible estimations resulted by performing independent EKFs is corrected based on the available point estimation of the states at each time interval. Note that AGSF initializes the prior estimation step of all the EKFs in the set using the point estimates evaluated at the end of each time interval. This implies that AGSF offers useful features for the case where system operates far from the process constraints.

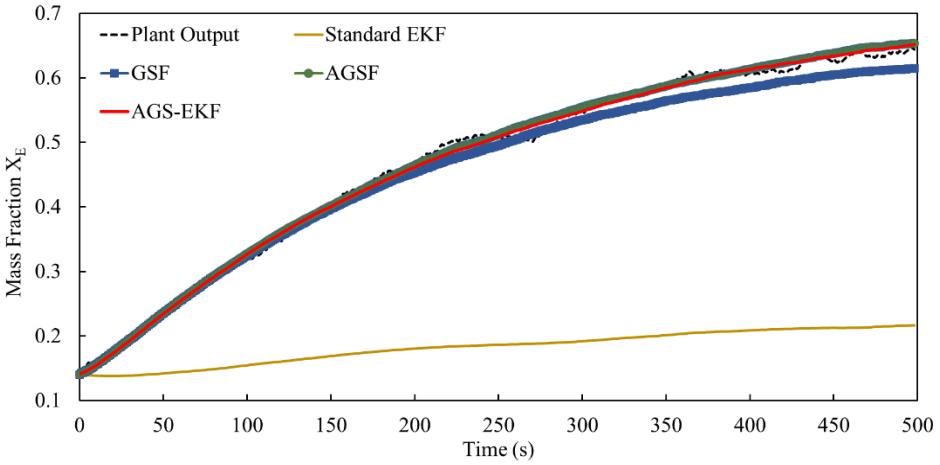
Moreover, the proposed AGS-EKF approach avoids performing the state estimation under only the first or the last Gaussian component in the Gaussian mixture model of the process uncertainties. The first Gaussian component suggests that the process uncertainties in the system are only negative, similarly, the last Gaussian component in the mixture represents only the positive valued process uncertainties presented in the actual non-Gaussian distribution (see Figure 4-6). However, AGS-EKF eliminate these biased estimations by considering that the process uncertainties can be drawn randomly from the overall Gaussian mixture model.



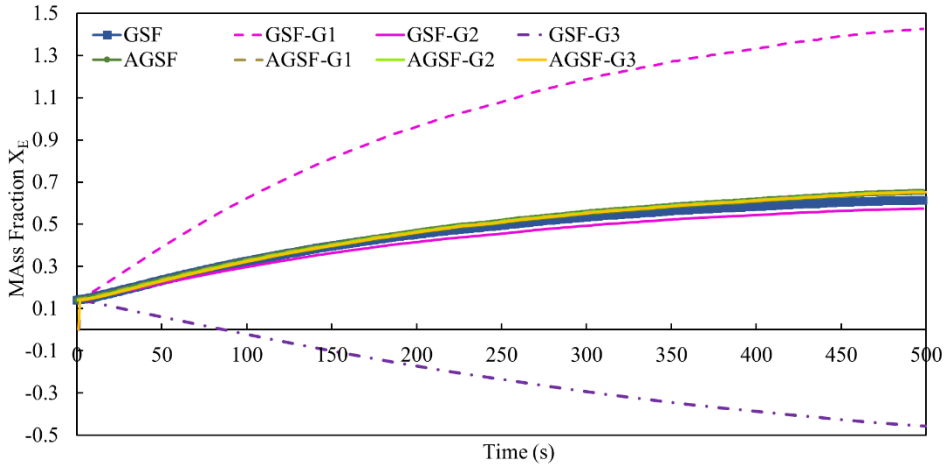
(b)


AGSF
AGSF-G1
AGSF-G2
AGSF-G3
GSF
GSF-G1
GSF-G2
GSF-G3
Mass Fraction X_C
Time (s)

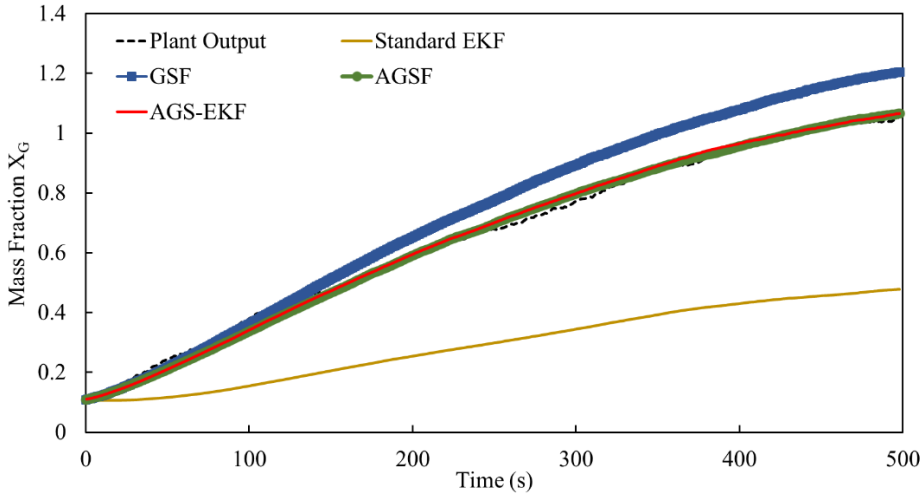
(c)


Plant Output
Standard EKF
GSF
AGSF
AGS-EKF
Mass Fraction X_E
Time (s)

(d)


GSF
GSF-G1
GSF-G2
GSF-G3
AGSF
AGSF-G1
AGSF-G2
AGSF-G3
MAss Fraction X_E
Time (s)

(e)



(f)

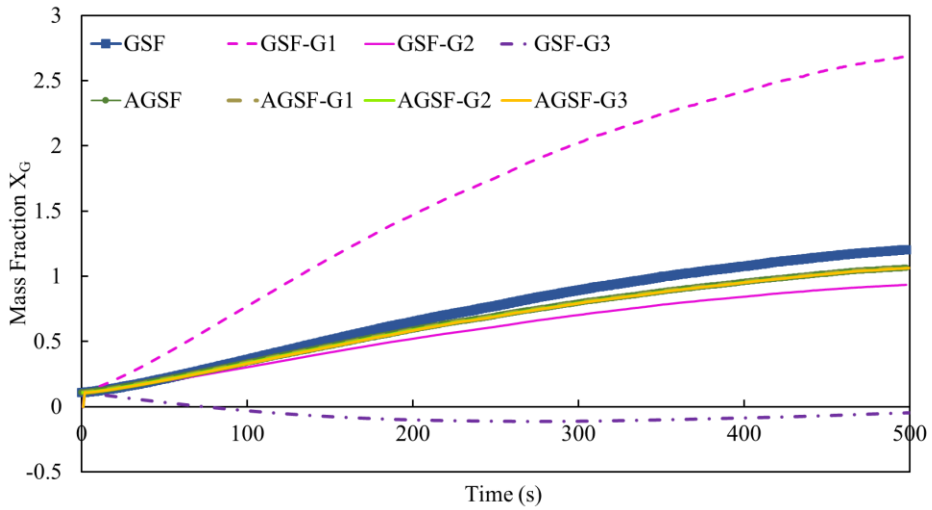


Figure 4-7: Estimation provided by various estimation schemes for mass fraction of the component (a), (b), X_C; (c), (d) X_E; (e), (f) X_G

Table 4-3: MSE for X_C, X_E, and X_G using different estimation schemes

Estimation method (n)	$MSE_{X_B}^{(n)}$	$MSE_{X_C}^{(n)}$	$MSE_{X_E}^{(n)}$	$MSE_{X_G}^{(n)}$
Standard EKF	2.3e-4	1.5e-4	9.8e-2	1.6e-1
GSF	1.3e-3	4.6e-5	3.3e-4	9.0e-3
AGSF	1.5e-4	1.7e-5	8.2e-5	3.8e-4
AGS-EKF	1.1e-4	1.7e-5	6.7e-5	3.7e-4

This observation suggests that for an unbounded state estimation similar to that considered in the current case study, GSF may not be a proper approach to estimate states that operate closely to their feasibility constraints. This issue can be addressed by considering a constrained estimation scheme for GSF, which is beyond the scope of the current work. Nevertheless, the proposed AGS-EKF approach provides accurate estimations to all states of the system. Therefore, this estimation scheme can perform closely to the actual process as they both are subjected to similar process uncertainties. In the current study, the required computational costs in GSF and AGSF are more than three times higher than that in AGS-EKF and standard EKF. Therefore, the proposed AGS-EKF scheme not only improves the accuracy of estimations of EKF, but also requires relatively lower computational costs (the same as needed in standard EKF) because it only performs EKF once.

4.2.3.1 Sensitivity Analysis

As mentioned earlier in this section, a sensitivity analysis has been conducted to investigate the required number of Gaussian components to obtain an adequate Gaussian mixture model as the approximation of the true non-Gaussian PDF of the process uncertainties. To perform this analysis, three Gaussian mixture models involving three, five and ten Gaussian components have been considered, as presented in Figures 4-6(a), 4-6(b), and 4-6(c), respectively. Figure 4-8 shows the performance of AGS-EKF under each of these assumptions. Here "$ng = 3$", "$ng = 5$", and "$ng = 10$" denotes the estimated states using AGS-EKF subjected to Gaussian mixture model contains three, five and ten Gaussian components, respectively. Based on the results shown in Figure 4-8, the estimation error remains the same by a tolerance of 0.001%. Therefore, three Gaussian components provide an accurate enough approximation of the non-Gaussian process uncertainties in this study. Hence, this selection to perform the current case study using "$ng = 3$" since additional Gaussian components increases computational costs without a significant improvement in the accuracy of the state estimation.

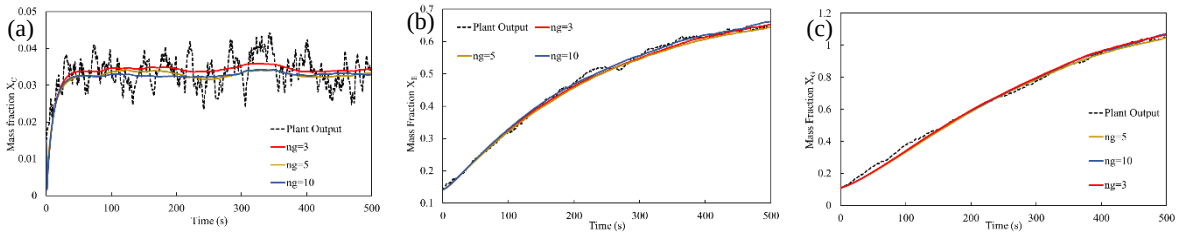


Figure 4-8: Estimation provided by various estimation schemes for mass fraction of the component (a) X_C; (b) X_E; (c) X_G

4.2.4 Wastewater Treatment Plant (WTP)

To further investigate the performance of the proposed AGS-EKF scheme, an industrial-scale wastewater treatment plant located in Manresa, Spain, is considered next. Current WTPs involve Nitrification/Denitrification processes[172,173], however, these processes have not been included in this case study for simplicity. Previous studies on WTP have shown that the standard EKF becomes unstable in the presence of the non-Gaussian process uncertainties[37]. Therefore, this system has been chosen to investigate the performance of AGS-EKF with respect to other estimation schemes presented in section 4.1 and the standard EKF (see section 3.2) under these conditions. To pursue this goal, the WTP considered in this study is subject to random process uncertainties with asymmetric bimodal densities.

Figure 4-9 illustrates a simplified flowsheet for the WTP considered in this study. In general, WTP aims to remove biodegradable pollutants (substrate) from the wastewater. This process is carried on in a bioreactor where substrate is converted to sludge and biomass. This process requires oxygen and biomass. Thus, an aeration turbine in the bioreactor is responsible for providing the required oxygen for this process, whereas a fresh biomass stream is entering the bioreactor to accomplish the reaction. The outlet stream of the bioreactor involving the activated sludge and biomass is sent to a decanter to settle down and remove the sludge from the treated water[158].

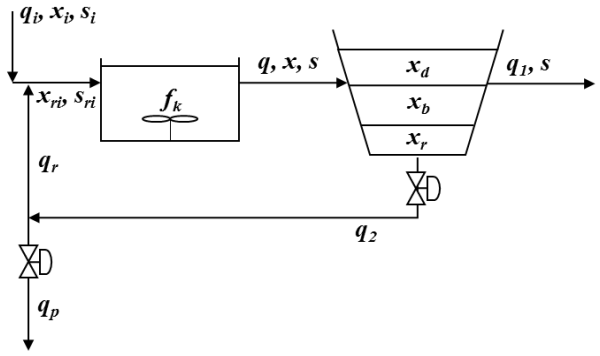


Figure 4-9: Wastewater treatment plant flowsheet

This process involves six states that are the biomass concentration (x_w), the organic substrate (s_w) and the dissolved oxygen concentration (c_w) inside the bioreactor, as well as the biomass concentration as the different layers in the decanter (i.e., x_d, x_b, x_r). Equation (4-29) represent the nonlinear model of WTP[158].

$$\frac{dx_w}{dt} = \mu_w y_w \frac{x_w s_w}{k_s + s_w} - k_d \frac{x_w}{s_w} - k_c x_w + \frac{q}{V_r}(x_{ir} - x_w) \tag{4-29}$$

$$\frac{ds_w}{dt} = -\mu_w \frac{x_w s_w}{k_s + s_w} + f_d k_d \frac{x_w}{s_w} + f_d k_c x_w + \frac{q}{V_r}(s_{ir} - s_w)$$

$$\frac{dx_d}{dt} = \frac{1}{A_d ld}\left(q_i - q_p\right)(x_b - x_d) - \frac{1}{ld}\text{vsd}$$

$$\frac{dx_b}{dt} = \frac{1}{A_d lb}\left(q_i + q_2 - q_p\right)(x_w - x_b) + \frac{1}{lb}(\text{vsd} - \text{vsb})$$

$$\frac{dx_r}{dt} = \frac{1}{A_d lr} q_2 (x_b - x_r)(x_b - x_d) + \frac{1}{lr}\text{vsb}$$

$$\frac{dc_w}{dt} = f_k k_{la}(c_s - c_w) - k_{01}\mu_w \frac{x_w s_w}{k_s + s_w} - \frac{q}{V_r} c_w$$

Model parameters for this process can be found elsewhere[158]. Online measurements are assumed to be available for s_w, c_w, and x_d. Note that the linear observability of system was confirmed, i.e., the observability matrix is full-rank (not shown for brevity). The Jacobian and sensitivity matrices required to form the observability matrix for WTP are presented in Appendix E. The random measurement noise signals follow zero-mean Gaussian distributions where the standard deviations are set to 10% of the nominal steady-state values of the states presented in Table B-2 in Appendix B. Moreover, a 1h sampling interval is considered for this process. Similar to the Williams-Otto reactor, WTP is subjected to feasibility constraints, i.e., non-negative concentrations, which has not been explicitly enforced during the present estimation.

The histograms for the asymmetric bimodal PDF of the process uncertainties in WTP are presented in Figure 4-10. Moreover, the red lines in Figure 4-10 illustrate the Gaussian mixture provided by the EM algorithm, which approximate the true non-Gaussian densities. Preliminary tests have been conducted to find the proper number of required Gaussian components in the mixture model to achieve both efficiency and accuracy. The current Gaussian mixture model for each of the process uncertainty variables involves two Gaussian components.

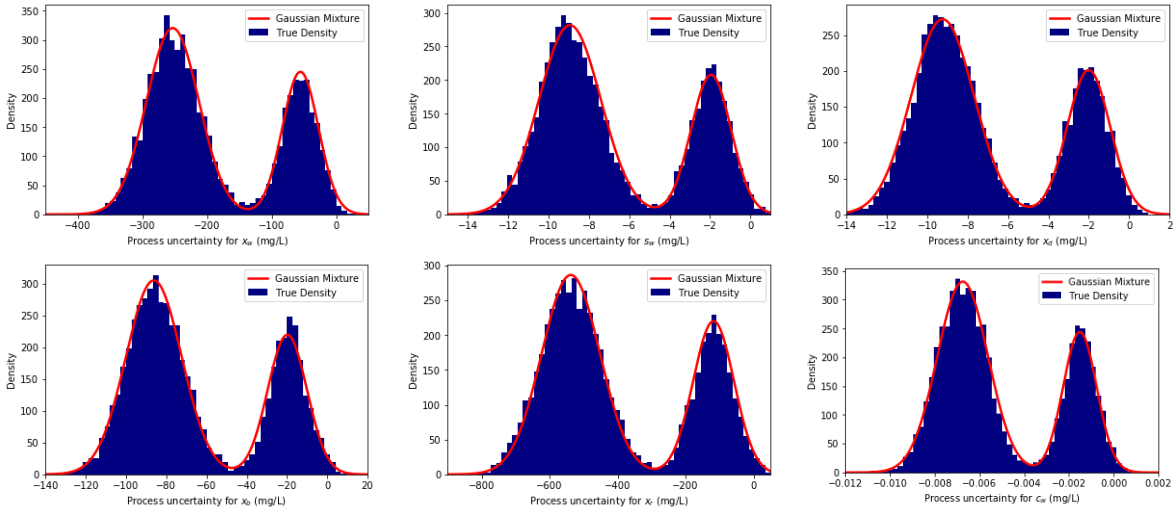
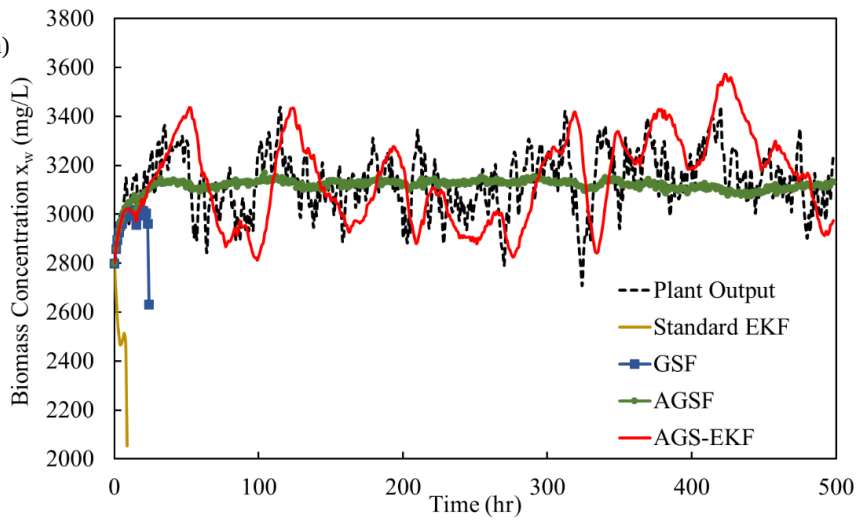


Figure 4-10: Asymmetric bimodal probability density of the process uncertainties in WTP compared to the approximated density by Gaussian mixture model

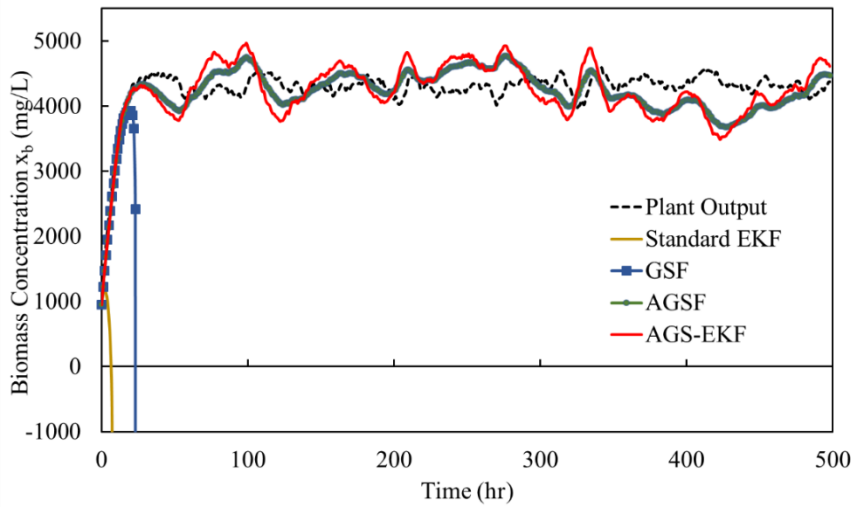
Based on the process uncertainties presented in Figure 4-10, GSF and AGSF needs to perform 64 EKFs (i.e., 2^6) at each time interval to provide the point estimates. Figure 4-11 highlights the estimation results for the unknown states in the WTP, i.e., x_w, x_b, and x_r. According to Figures 4-11(a), 4-11(b), and 4-11(c), standard EKF failed after at iteration 11 since this method considers zero-mean Gaussian PDFs for the process uncertainties while the actual process uncertainties in the system follow non-Gaussian densities, as shown in Figure 4-10. That is, poor estimates were used to update the Jacobian matrix, which eventually made the Jacobian matrix unstable, i.e., positive eigenvalues. Consequently, the EKF framework became unstable and was unable to compute the estimates. Both AGSF and AGS-EKF estimated the states accurately. As shown in Table 4-4, AGS-EKF improved the accuracy in the estimation of x_w, x_b, x_r, by several orders of magnitude when compared to standard EKF. Another feature that can be drawn from the results presented in Figure 4-11 and Table 4-4 is that GSF resulted into instability after 26 iterations. For instance, the evaluate of MSE in the case of GSF is in the order of 1e+9 and 1e+6 for x_b, and x_r, respectively. Similar to the discussion presented in section 3.3, the divergence of GSF occurs because this method performs individual EKFs, i.e., each EKF only uses one of combination of the Gaussian components in the process uncertainties mixture models. As a result, there are several EKFs in the set of GSF who perform state estimation using the process uncertainties drawn from the edge of the PDFs. These EKFs drive the system to the estimates into infeasibility region, which eventually leads to an infeasible point estimates and consequently a divergence in GSF. As in the case of the Williams-Otto reactor, AGSF uses the overall point estimate as a priori information to correct the biased estimations in GSF. Hence, AGSF results in acceptable estimates. Nevertheless, the

proposed AGS-EKF is not performing individual EKFs who operates under the extreme partition of the process uncertainty distribution, e.g., combination of the Gaussian components that only draws large negative random values from the non-Gaussian process uncertainties' PDF. That is, AGS-EKF uses the characteristics of the overall non-Gaussian distribution of the process uncertainties and run the modified EKF in which random process uncertainties are drawn from the general Gaussian mixture model. Thus, AGS-EKF is able to provide an accurate estimation. For instance, according to Table 4-4, the error in the estimation of states x_b and x_r provided by AGS-EKF is reduced by 4 and 2 orders of magnitudes when compared to the GSF estimation scheme, respectively.

(a)



(b)



(c)

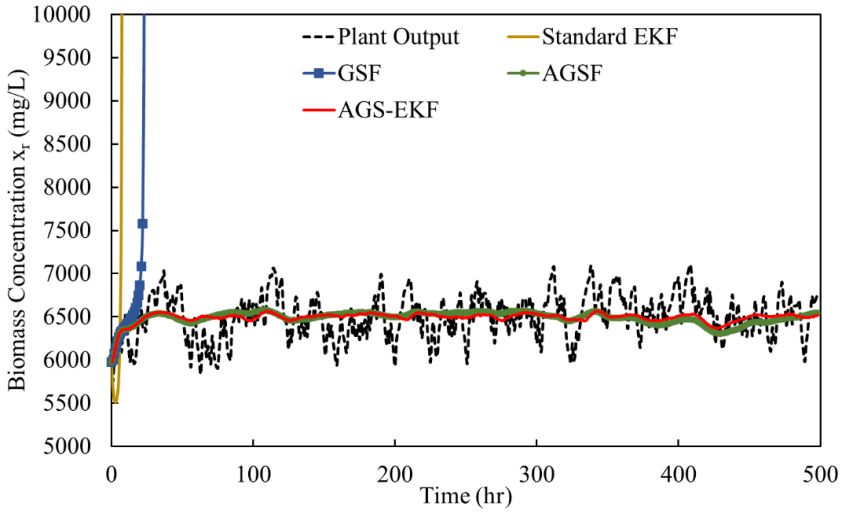


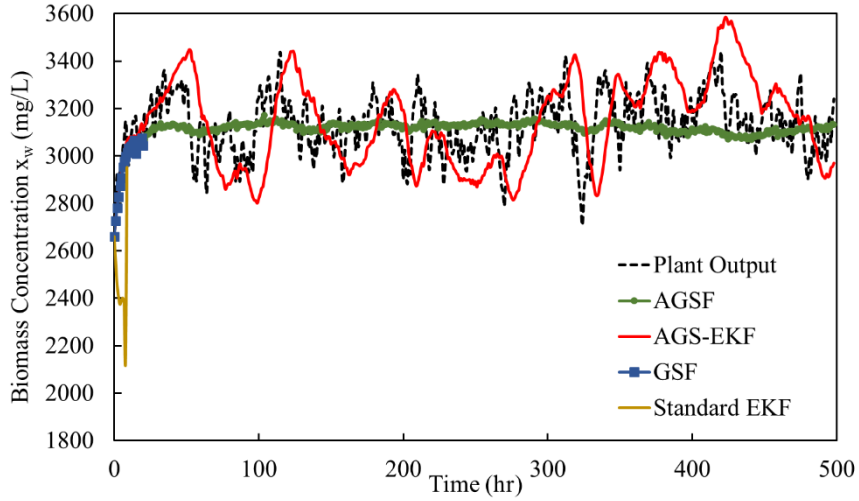
Figure 4-11: Estimation provided by various estimation schemes (a), (b), biomass concentration in the bioreactor x_w; (c), (d) biomass concentration in the second layer of the decantor x_b; (e), (f) biomass concentration in the bottom layer of the decantor x_r

Table 4-4: MSE for x_w, x_b, and x_r using different estimation schemes

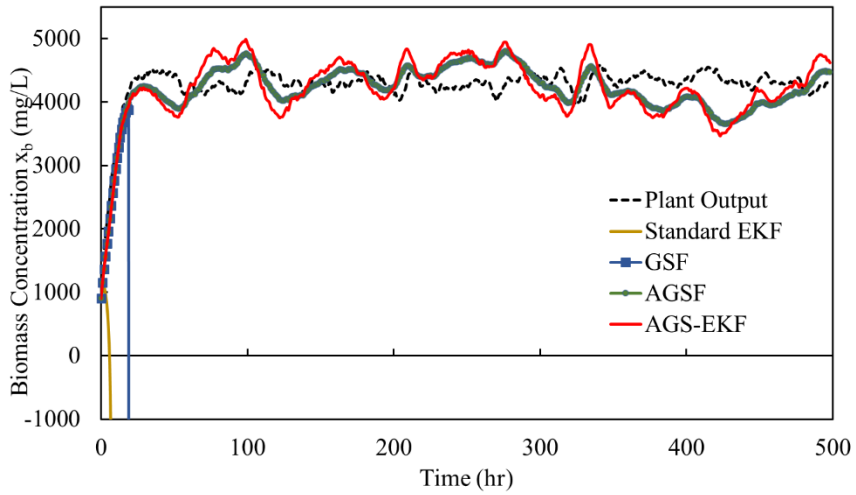
Estimation method (n)	$MSE_{x_w}^{(n)}$	$MSE_{x_b}^{(n)}$	$MSE_{x_r}^{(n)}$
Standard EKF	2.8e+5	5.7e+11	2.3e+12
GSF	2.1e+4	2.0e+9	4.0e+6
AGSF	1.4e+4	9.1e+4	6.4e+4
AGS-EKF	3.4e+4	1.5e+5	6.2e+4

The results presented above for the WTP are based on the assumption that the estimation schemes start from the true initial conditions. To further confirm the performance of the proposed estimation scheme, Figure 4-12 presents the state estimates provided by the estimation schemes when the initial guess provided to the estimators is 5% smaller than the true value of states at the initial time interval. As expected, the performance and outcomes observed for each state estimation scheme is similar to that obtained for the previous scenario shown in Figure 4-11, i.e., both the proposed AGS-EKF and AGSF improved significantly the estimation accuracy when compared to standard EKF and GSF. Note that AGSF is a computationally expensive scheme while AGS-EKF does not increase the computational costs thus making it a more attractive estimation method.

(a)


Biomass Concentration x_w (mg/L)
3600
3400
3200
3000
2800
2600
2400
2200
2000
1800
0
100
200
300
400
500
Time (hr)
Plant Output
AGSF
AGS-EKF
GSF
Standard EKF

(b)


Biomass Concentration x_b (mg/L)
5000
4000
3000
2000
1000
0
-1000
0
100
200
300
400
500
Time (hr)
Plant Output
Standard EKF
GSF
AGSF
AGS-EKF

(c)

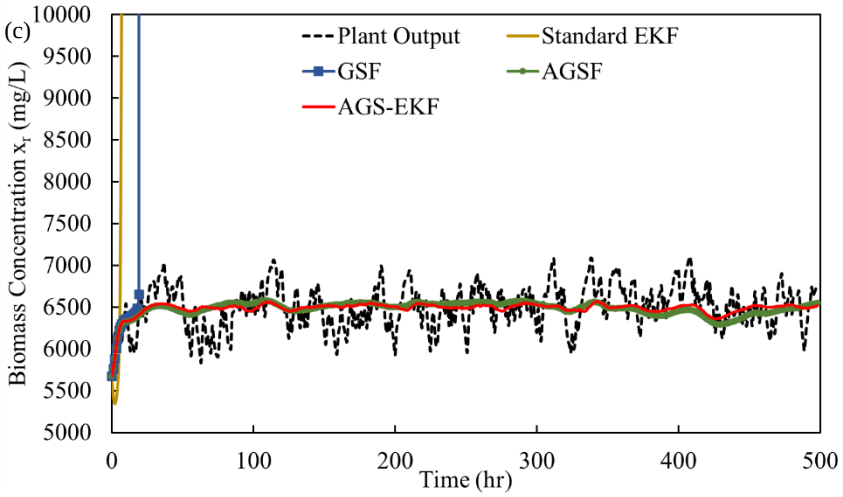

Biomass Concentration x_r (mg/L)
10000
9500
9000
8500
8000
7500
7000
6500
6000
5500
5000
0
100
200
300
400
500
Time (hr)
Plant Output
Standard EKF
GSF
AGSF
AGS-EKF

Figure 4-12: Estimation provided by various estimation schemes that are initilized by 95% of the true initial states (a), (b), biomass concentration in the bioreactor x_w; (c), (d) biomass concentration in the second layer of the decantor x_b; (e), (f) biomass concentration in the bottom layer of the decantor x_r.

The main advantage of the proposed AGS-EKF approach when compared to AGSF is that it only performs one EKF, regardless of how many Gaussian components are present in the Gaussian mixture models of the process uncertainties. On the other hand, the number of EKFs in the set of AGSF (and GSF) increases by both the number of Gaussian components in the mixture and the size of the process (number of states and consequently number of process uncertainties associated with the states). Table 4-5 presents the averaged CPU time reported for one iteration of the state estimation using AGS-EKF, AGSF (and GSF), and standard EKF. The values reported represents the actual and normalized CPU times for each approach considering a parallelization in the GSF (and AGSF) simulation, i.e., EKFs in the set are performed in parallel for each time interval. The normalization is performed based on the average CPU time reported for the case of performing standard EKF. This table shows that the CPU time for the standard EKF and AGS-EKF are the same. The largest CPU time is for the case of AGSF (and GSF), which is 5 times larger than that reported for AGS-EKF (and standard EKF).

Table 4-5: Averaged CPU time reported for different estimation schemes

Estimation method	Averaged (Normalized) CPU time
Standard EKF	0.0011 s (1)
GSF and AGSF	0.0055 s (5)
AGS-EKF	0.0011 s (1)

To generate the results presented in Table 4-5, the 8 cores of the computer used for this case study to parallelize the AGSF (and GSF) schemes. WTP involves six process uncertainty variables where each of these variables belong to a Gaussian mixture model that only requires two Gaussian components. Therefore, AGSF (and GSF) requires running 64 EKFs to estimate the states at each time interval. However, a larger set of EKFs would be required to perform AGSF (and GSF) for larger applications. For instance, if the Gaussian mixture models consists of two Gaussian components that describe the process uncertainties associated with the state of a large-scale model featuring n_x states variables, then GSF would need to perform a set of 2^{n_x} EKFs to provide the online estimation of the states. This may make the problem highly intensive, particularly when this approach is implemented online for closed-loop applications. On the other hand, AGS-EKF may provide the similar accuracy as in AGSF at the

same required CUP time as in standard EKF thus making it quite attractive for online estimation and control.

4.3 Summary

This chapter presented a novel EKF-based state estimation scheme (AGS-EKF) for nonlinear applications where process uncertainties are non-zero mean and do not follow a Gaussian distribution. The proposed framework is a modified version of EKF to deal with non-zero mean process uncertainties. The modified EKF does not add any computational burden to the standard EKF. In the AGS-EKF framework, the non-Gaussian density of the process uncertainty is approximated by a Gaussian mixture model similar to GSF. However, AGS-EKF uses the main characteristics of the Gaussian mixture of the process uncertainties to perform only one EKF calculation and consequently is computationally attractive. In addition to the efficiency offered by AGS-EKF, this estimation scheme does not suffer from operating at the edge of feasibility region as it avoids running individual EKFs featuring only an extreme range of the process uncertainties. As a result, AGS-EKF manages to stay far away from the edge of the process and avoids biased estimations. This feature makes the AGS-EKF an attractive method for large applications subject to non-Gaussian process uncertainties. In addition to AGS-EKF, an adaptation in the GSF framework was introduced that can help to correct the biased estimations provided by some of the EKFs in the set of AGSF. Hence, AGSF improved the performance of GSFs. The performance of AGS-EKF compared to standard EKF, GSF, and AGSF is assessed by performing state estimation on four case studies involving non-Gaussian process uncertainties. The results showed that AGS-EKF is able to provide accurate estimations for the unknown states in nonlinear chemical engineering systems without increasing the computational demands thus making it attractive for online estimation and control applications. The AGS-EKF introduced in this chapter only focuses on the problem of state estimation under non-Gaussian uncertainties. An AGS-EKF state estimation framework featuring constraints on the state variables is presented in the next chapter.

Chapter 5

Constrained Abridged Gaussian Sum Extended Kalman Filter

As described in sections 2.2 and 2.3, an efficient development of EKF to deal with constraints on the states is absent from the literature. Moreover, the performance of the GSF, i.e., that is known as the EKF-based framework to capture the non-Gaussianity in process, have not been investigated for a general constrained nonlinear application with non-zero mean non-Gaussian process uncertainties/ measurement noises. As discussed above, the available EKF methods, and particularly GSFs, face the following limitations: (I) high computational costs, (II) biased and infeasible estimations, (III) large covariance matrices and consequently, inappropriate estimations. The AGS-EKF presented in the previous chapter avoids limitations (I)-(III), improves the estimation accuracy, and requires no additional costs than that needed in the standard EKF. However, the AGS-EKF method presents two major limitations: it cannot deal with non-Gaussian measurement noises or/and feasibility bounds on the states.

This work fills a gap in the literature by presenting a generalized and efficient AGS-EKF-based state estimation scheme for nonlinear applications involving bounds on the states (i.e., states with non-Gaussian probability density functions) and non-Gaussian measurement noises and process uncertainties. The proposed approach is referred from heretofore as *constrained abridged Gaussian sum extended Kalman filter* (constrained AGS-EKF). To deal with non-Gaussian measurement noises, new modifications in the posterior estimation step of the standard EKF formulation is considered to adopt non-zero mean measurement noises in the estimation scheme. The modified posterior estimation step requires the same computational costs as needed in the standard EKF framework. The proposed constrained AGS-EKF includes an additional step in between the prior estimation step and the posterior estimation step in the EKF scheme to satisfy the constraints on the states. In this intermediate step, the constrained AGS-EKF uses the prior distribution of the states (resulted by the prior estimation step in EKF) as well as the knowledge of the present constraints/bounds to approximate the non-Gaussian distribution of the states by a Gaussian mixture model. Then, the main characteristics of the Gaussian mixture model, i.e., mean-vector of the states and covariance matrix, are used as the estimation of the constrained prior distributions of the states to continue performing the posterior estimation step in EKF. The multivariate expectation-maximization (EM) method is used to approximate the Gaussian mixture models of the non-Gaussian densities of the states. Note that the constrained AGS-EKF assumes that the distribution of the process uncertainties and measurement noises are known *a priori* from historical

data or process heuristics, which is also the common assumption made in standard EKF. The performance of these estimation schemes for the case of unknown distributions of noise/uncertainty is beyond the scope of this work.

The outline of this chapter is as follows: the proposed state estimation scheme for constrained AGS-EKF is introduced in section 5.1. Later, the results obtained from the proposed approach on three typical chemical processes (i.e., a gas-phase reactor presented in Equation (4-26), the Williams-Otto reactor presented in Equation (4-28), and the styrene polymerization process featuring six states) is presented. Section 5.3 summarizes this chapter.

5.1 Filters for constrained states and non-Gaussian uncertainties/ noises

This section presents the proposed constrained Abridged Gaussian sum extended Kalman filter (AGS-EKF). A modified version of EKF for the case that process uncertainty and measurement noise follow non-zero mean Gaussian distribution is presented first. The constrained AGS-EKF formulation is presented next. Both the EKF and GSF schemes are used in this work as benchmarks to investigate and compare the performance of the proposed constrained AGS-EKF method. Hence, section 5.1.3 provides a brief discussion on GSF adopted for applications with non-Gaussian states, noises, and uncertainties.

5.1.1 Modified EKF for non-zero mean Gaussian process uncertainty and measurement noise

General nonlinear systems with bounded states are the focus of this work. These systems can be modelled as follows:

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k, k) \tag{5-1}$$

$$\boldsymbol{y}_k = h(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{v}_k, k) \tag{5-2}$$

$$\boldsymbol{x}^{\boldsymbol{l}} \leq \boldsymbol{x}_k \leq \boldsymbol{x}^{\boldsymbol{u}} \tag{5-3}$$

$\boldsymbol{w} \in \mathbb{R}^{n_w}$, $\boldsymbol{v} \in \mathbb{R}^{n_v}$, $\boldsymbol{x}^{\boldsymbol{l}}, \boldsymbol{x}^{\boldsymbol{u}}, \boldsymbol{x} \in \mathbb{R}^{n_x}$, $\boldsymbol{y} \in \mathbb{R}^{n_y}$, $\boldsymbol{u} \in \mathbb{R}^{n_u}$,

f: $\mathbb{R}^{n_x \times n_u \times n_w} \rightarrow \mathbb{R}^{n_x}$, h: $\mathbb{R}^{n_x \times n_u \times n_v} \rightarrow \mathbb{R}^{n_y}$

where $\boldsymbol{x}^{\boldsymbol{l}}$ and $\boldsymbol{x}^{\boldsymbol{u}}$ denote the lower and upper bound on the state variables, respectively. Moreover, $\mathbb{E}\left[\boldsymbol{w}_{\varsigma}\boldsymbol{x}_0^T\right] = 0$ for all k where $\boldsymbol{x}_0$ is the vector of initial states. A common assumption is that the random process uncertainties ($\boldsymbol{w}$) and measurement noises ($\boldsymbol{v}$) are mutually uncorrelated, respectively. In this thesis, it is assumed that each measurement is corrupted with one unique random measurement noise

signal and each state of the system is associated with one random variable as its process uncertainty, which is a common assumption considered for chemical processes. Hence, $\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k$, $\boldsymbol{y}_k = h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{v}_k$, $n_v = n_y$ and $n_w = n_x$. As a reminder, Equations (3-4)-(3-5) describe the prior and posterior estimation steps in the standard formulation of EKF for an unconstrained nonlinear system such as that described by Equations (5-1)-(5-2) with zero-mean Gaussian random variables, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$.

In order to take into account the non-Gaussian process uncertainties and measurement noises, first, the standard EKF formulation needs to be modified accordingly for the case that these random variables follow non-zero mean Gaussian distributions. In chapter 4, a modification in the prior step of the standard EKF has been introduced to adopt non-zero mean Gaussian process uncertainties (section 4.1.1). Likewise, the current work considers a new modification in the posterior step of the standard EKF to capture non-zero mean Gaussian distribution of the measurement noises. That is, consider the unconstrained nonlinear system presented in Equations (5-1)-(5-2) in which $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{Q})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}, \boldsymbol{R})$. The modified EKF for applications involving non-zero mean Gaussian noises/uncertainties as follows:

$$\begin{aligned}
&\widehat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{\mu} \\
&\boldsymbol{P}_{k+1|k} = \boldsymbol{A}_k \boldsymbol{P}_{k|k} \boldsymbol{A}_k^{\boldsymbol{T}} + \boldsymbol{Q}
\end{aligned} \tag{5-4}$$

$$\begin{aligned}
&\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1|k} \boldsymbol{H}_{k+1}^T / (\boldsymbol{H}_{k+1} \boldsymbol{P}_{k+1|k} \boldsymbol{H}_{k+1}^T + \boldsymbol{R}) \\
&\widehat{\boldsymbol{x}}_{k+1|k+1} = \\
&\qquad \widehat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1} \big(\boldsymbol{y}_{k+1} - h(\widehat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k) - \boldsymbol{\tau} \big) \\
&\boldsymbol{P}_{k+1|k+1} = (\boldsymbol{I} - \boldsymbol{K}_{k+1} \boldsymbol{H}_{k+1}) \boldsymbol{P}_{k+1|k}
\end{aligned} \tag{5-5}$$

$$\begin{aligned}
&\big[(\boldsymbol{w}_\varsigma - \boldsymbol{\mu})(\boldsymbol{w}_\varsigma - \boldsymbol{\mu})^T \big] = \begin{cases} \boldsymbol{Q}, & \varsigma = k \\ \boldsymbol{0}, & \varsigma \neq k \end{cases} \; ; \; \boldsymbol{\mu} \in \mathbb{R}^{n_x}, \boldsymbol{Q} \in \mathbb{R}^{n_x \times n_x} \\
&\big[(\boldsymbol{v}_\varsigma - \boldsymbol{\tau})(\boldsymbol{v}_\varsigma - \boldsymbol{\tau})^T \big] = \begin{cases} \boldsymbol{R}, & \varsigma = k \\ \boldsymbol{0}, & \varsigma \neq k \end{cases} \; ; \; \boldsymbol{\tau} \in \mathbb{R}^{n_y}, \boldsymbol{R} \in \mathbb{R}^{n_y \times n_y}
\end{aligned} \tag{5-6}$$

Note that the modified EKF presented in Equations (5-4)-(5-6) is a general case of standard EKF. That is, for the special case where $\boldsymbol{\tau} = \boldsymbol{0}$ and $\boldsymbol{\mu} = \boldsymbol{0}$, the modified EKF represents the same formulation as the standard EKF (Equations (3-4)-(3-5)). Moreover, the derivation of the prior estimation step, i.e.,

Equation (5-4), can be found in section 4.1.1. The derivation of the posterior estimation, i.e., Equation (5-5), is presented next.

Derivation of the posterior estimation: this derivation is a mathematical proof that the posterior estimation step in the constrained AGS-EKF framework (i.e., Equation (5-5)) should be used for applications where the measurement noises follow non-zero mean distributions.

Consider a nonlinear dynamic model presented in Equations (5-1)-(5-2) associated with non-zero mean Gaussian measurement noises as described by Equation (5-6). For this system, the estimation of the measurements $\widehat{\boldsymbol{y}}_k$ is determined as follows:

$$\begin{aligned}\widehat{\boldsymbol{y}}_{k+1} &= \mathbb{E}\left[h\left(\widehat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k, \boldsymbol{v}_{k+1}\right)\middle|\boldsymbol{Y}_{k+1}\right] \\ &= \boldsymbol{H}_{k+1}\mathbb{E}[\widehat{\boldsymbol{x}}_{k+1|k}|\boldsymbol{Y}_{k+1}] + \boldsymbol{C}_{k+1}\mathbb{E}[\boldsymbol{u}_k|\boldsymbol{Y}_{k+1}] + \boldsymbol{D}_{k+1}\mathbb{E}[\boldsymbol{v}_{k+1}|\boldsymbol{Y}_{k+1}]\end{aligned} \tag{5-7}$$

where $\boldsymbol{Y}_{k+1} = [\boldsymbol{y}_0, \boldsymbol{y}_1, \dots, \boldsymbol{y}_{k+1}]$ is the set of measurements from the initial sampling time to the current time interval $k+1$. Moreover, $\boldsymbol{C}_k = \frac{\partial h(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})}{\partial \boldsymbol{u}}|\boldsymbol{x} = \widehat{\boldsymbol{x}}_{k|k}, \boldsymbol{u} = \boldsymbol{u}_k, \boldsymbol{v} = \boldsymbol{v}_k$ and $\boldsymbol{D}_k = \frac{\partial h(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})}{\partial \boldsymbol{v}}|\boldsymbol{x} = \widehat{\boldsymbol{x}}_{k|k}, \boldsymbol{u} = \boldsymbol{u}_k, \boldsymbol{v} = \boldsymbol{v}_k$. The sensitivity matrices $\boldsymbol{H}$, $\boldsymbol{C}$ and $\boldsymbol{D}$, as well as the measurement noises and inputs of the system, are independent from the historical measurements. The estimation of the measurements can be expressed as below:

$$\widehat{\boldsymbol{y}}_{k+1} = \boldsymbol{H}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{C}_{k+1}\boldsymbol{u}_k + \boldsymbol{D}_{k+1}\boldsymbol{\tau} \tag{5-8}$$

The innovation, i.e., the difference between actual measurements ($\boldsymbol{y}_k$) and the estimated measurements ($\widehat{\boldsymbol{y}}_k$), can be evaluated based on Equations (5-2) and (5-8), that is:

$$\begin{aligned}\boldsymbol{y}_{k+1} - \widehat{\boldsymbol{y}}_{k+1} &= \boldsymbol{H}_{k+1}\boldsymbol{x}_{k+1} + \boldsymbol{C}_{k+1}\boldsymbol{u}_k + \boldsymbol{D}_{k+1}\boldsymbol{v}_{k+1} \\ &\quad - \left(\boldsymbol{H}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{C}_{k+1}\boldsymbol{u}_k + \boldsymbol{D}_{k+1}\boldsymbol{\tau}\right) \\ &= \boldsymbol{H}_{k+1}\left(\boldsymbol{x}_{k+1} - \widehat{\boldsymbol{x}}_{k+1|k}\right) + \boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})\end{aligned} \tag{5-9}$$

Given $\boldsymbol{e}_{k+1|k} = \boldsymbol{x}_{k+1} - \widehat{\boldsymbol{x}}_{k+1|k}$ is known as the prior state estimation error, Equation (5-9) can be re-written as follows:

$$\boldsymbol{y}_{k+1} - \widehat{\boldsymbol{y}}_{k+1} = \boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k} + \boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau}) \tag{5-10}$$

The estimation of the measurement's covariance ($\boldsymbol{S}$) can be determined based on the innovation obtained by Equation (5-9). That is:

$$\begin{aligned}
&\boldsymbol{S}_{k+1} = \mathbb{E}[(\boldsymbol{y}_{k+1} - \widehat{\boldsymbol{y}}_{k+1})(\boldsymbol{y}_{k+1} - \widehat{\boldsymbol{y}}_{k+1})^T] \\
&= \mathbb{E}\left[\left(\boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k} + \boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})\right)\left(\boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k} + \boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})\right)^T\right] \\
&= \mathbb{E}\left[\left(\boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k} + \boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})\right)\left(\boldsymbol{e}_{k+1|k}{}^T\boldsymbol{H}_{k+1}{}^T + (\boldsymbol{v}_{k+1} - \boldsymbol{\tau})^T\boldsymbol{D}_{k+1}{}^T\right)\right] \\
&= \mathbb{E}\left[\boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k}\boldsymbol{e}_{k+1|k}{}^T\boldsymbol{H}_{k+1}{}^T\right] + \mathbb{E}\left[\boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})^T\boldsymbol{D}_{k+1}{}^T\right] + \\
&\mathbb{E}\left[\boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})\boldsymbol{e}_{k+1|k}{}^T\boldsymbol{H}_{k+1}{}^T\right] + \mathbb{E}\left[\boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})^T\boldsymbol{D}_{k+1}{}^T\right]
\end{aligned} \tag{5-11}$$

where $\boldsymbol{P}_{k+1|k} = \mathbb{E}\left[\boldsymbol{e}_{k+1|k}\boldsymbol{e}_{k+1|k}{}^T\right]$ is the prior covariance matrix of the estimated states. According to Equation (5-6), $\boldsymbol{R} = \mathbb{E}[(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})^T]$ is the covariance matrix of the measurement noises. Moreover, the estimation errors are uncorrelated with the measurement noises, i.e., the second and third terms in the left-hand side of Equation (5-11) are equal to zero. Equation (5-11) can be simplified to Equation (5-12).

$$\boldsymbol{S}_{k+1} = \boldsymbol{D}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{H}_{k+1}{}^T + \boldsymbol{D}_{k+1}\boldsymbol{R}\boldsymbol{D}_{k+1}{}^T \tag{5-12}$$

To determine the Kalman gain matrix, the cross-covariance matrix ($\boldsymbol{Cov}_k^{\boldsymbol{S-M}}$) between the state estimation error and the innovation needs to be evaluated first. Equation (5-13) express this cross-covariance matrix, as below:

$$\begin{aligned}
&\boldsymbol{Cov}_k^{\boldsymbol{S-M}} = \mathbb{E}\left[\left(\boldsymbol{e}_{k+1|k}\right)(\boldsymbol{y}_{k+1} - \widehat{\boldsymbol{y}}_{k+1})^T\right] \\
&= \mathbb{E}\left[\left(\boldsymbol{e}_{k+1|k}\right)\left(\boldsymbol{H}_{k+1}\boldsymbol{e}_{k+1|k} + \boldsymbol{D}_{k+1}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})\right)^T\right] \\
&= \mathbb{E}\left[\boldsymbol{e}_{k+1|k}\left(\boldsymbol{e}_{k+1|k}{}^T\boldsymbol{H}_{k+1}{}^T + (\boldsymbol{v}_{k+1} - \boldsymbol{\tau})^T\boldsymbol{D}_{k+1}{}^T\right)\right] \\
&= \mathbb{E}\left[\boldsymbol{e}_{k+1|k}\boldsymbol{e}_{k+1|k}{}^T\boldsymbol{H}_{k+1}{}^T + \boldsymbol{e}_{k+1|k}(\boldsymbol{v}_{k+1} - \boldsymbol{\tau})^T\boldsymbol{D}_{k+1}{}^T\right]
\end{aligned} \tag{5-13}$$

$$\boldsymbol{Cov}_k^{\boldsymbol{S-M}} = \boldsymbol{P}_{k+1|k}\boldsymbol{H}_{k+1}{}^T \tag{5-14}$$

In the end, Kalman gain of the modified EKF is determined based in Equations (5-12) and (5-14), that is:

$$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1|k}\boldsymbol{H}_{k+1}{}^{T} / \boldsymbol{D}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{H}_{k+1}{}^{T} + \boldsymbol{D}_{k+1}\boldsymbol{R}\boldsymbol{D}_{k+1}{}^{T} \quad (5\text{-}15)$$

Note that the sensitivity matrix $\boldsymbol{D}$ is assumed to be an identity matrix with the proper dimension, i.e., $\boldsymbol{D} \in \mathbb{R}^{n_y \times n_y}$. Therefore, Equation (5-15) is equivalent to the Kalman gain of standard EKF presented in Equation (3-5); the posterior estimation of the states can be evaluated based on this Kalman gain and the innovation presented in Equation (5-10), as shown in Equation (5-5).

Note that derivations presented above and in section 4.1.1 represent the mathematical expressions that proves the need of Equations (5-5) and (5-4) for cases involving non-zero mean random noises/uncertainties.

5.1.2 Constrained Abridged Gaussian Sum Extended Kalman Filter (Constrained AGS-EKF)

The proposed constrained abridged Gaussian sum extended Kalman filter (constrained AGS-EKF) is developed based on the modified EKF scheme presented in Equations (5-4)-(5-6) plus an additional intermediate step. Constrained AGS-EKF uses the mean-value and the covariance matrix of the overall Gaussian mixture models to approximate the non-Gaussian states (constrained/bounded states) as well as non-Gaussian measurement noises and process uncertainties. That is, the mean value and the covariance matrix of the states ($\widehat{\boldsymbol{x}}_{k+1|k} \sim \mathcal{N}(\widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}}, \boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}})$), the process uncertainties ($\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}^{\boldsymbol{GM}}, \boldsymbol{R}^{\boldsymbol{GM}})$), and the measurement noises ($\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{GM}}, \boldsymbol{Q}^{\boldsymbol{GM}})$) obtained from the Gaussian mixture models are used to represent the non-Gaussian densities of the states, process uncertainties and measurement noises in the constrained AGS-EKF framework. For clarity, Equations (5-16)-(5-18) are provided to describe the multivariate Gaussian mixture model of the states, as well as the univariate Gaussian mixture models of the process uncertainties and the measurement noises, respectively. Note that a system involving n_x states (and process uncertainty variables) and n_y measurements (and measurement noise variables) is considered in this study.

$$p(\widehat{\boldsymbol{x}}_{k+1|k}) = \sum_{is=1}^{ngs} \left(\alpha s^{is}\, \mathcal{N}\left[\widehat{\boldsymbol{x}}_{k+1|k};\, \widehat{\boldsymbol{x}}_{k+1|k}^{is}, \boldsymbol{P}_{k+1|k}^{is}\right]\right); \quad \forall k \quad (5\text{-}16)$$

$$p(w_{l_k}) = \sum_{ip=1}^{ngp(l)} \left(\alpha p_l^{ip}\, \mathcal{N}\left[w_{l_k};\, \mu_l^{ip}, Q_l^{ip}\right]\right); \qquad \forall k;\ \forall l = 1, \dots n_x \quad (5\text{-}17)$$

$$p(v_{c_k}) = \sum_{im=1}^{ngm(c)} \left(\alpha m_c^{im}\, \mathcal{N}\left[v_{c_k};\, \tau_c^{im}, R_c^{im}\right]\right);\ \forall k;\ \forall c = 1, \dots n_y \quad (5\text{-}18)$$

$$\sum_{is=1}^{ngs} \alpha s^{is} = 1; \quad \alpha s^{is} \geq 0 \qquad \forall k;$$

$$\sum_{ip=1}^{ngp(l)} \alpha p_l^{ip} = 1; \quad \alpha p_l^{ip} \geq 0 \qquad \forall k; \forall l = 1, \dots n_x$$

$$\sum_{im=1}^{ngm(c)} \alpha m_c^{im} = 1; \quad \alpha m_c^{im} \geq 0 \quad \forall k; \forall c = 1, \dots n_y$$

$$\mu_l \in \boldsymbol{\mu}, \tau_c \in \boldsymbol{\tau}, Q_l \in \boldsymbol{Q}, R_c \in \boldsymbol{R}\,; \qquad \forall k;\ \forall c = 1, \dots n_x;\ \forall c = 1, \dots n_y$$

where is, ip, and im denote the indexes for the Gaussian component in the Gaussian mixture model of states, process uncertainties, and measurement noises, respectively. Moreover, the scalar w_{l_k} refers to the process uncertainty associated with the l^{th} state variable at time interval k. The number of Gaussian components for the process uncertainties associated with l^{th} state is $ngp(l)$. Likewise, scalar v_{c_k} denotes to the c^{th} element of vector $\boldsymbol{v}_k$, which represents the measurement noise associated with c^{th} measurement at time interval k; $ngm(c)$ denotes the corresponding number of Gaussian components in the mixture. In Equation (5-16), $\widehat{\boldsymbol{x}}_{k+1|k}$ is the vector of states obtained from the prior estimation step in constrained AGS-EKF, and ngs represents the number of Gaussian components that form the multivariate Gaussian mixture model of the states. Moreover, αs^{is}, αm_c^{im}, and αp_l^{ip} are the weights assigned to the is^{th} component in the Gaussian mixture model of states, im^{th} component in the Gaussian mixture model of c^{th} measurement noise, ip^{th} Gaussian component in the Gaussian mixture model of the l^{th} process uncertainty variable, respectively. Typically, the distribution of the process uncertainties and measurement noises are assumed to remain unchanged throughout the process. In this case, the EM algorithm can be performed offline to evaluate the corresponding parameters of the Gaussian mixture models of these random variables. This means that Equations (5-17)-(5-18) are used only once (and prior to the state estimation) to approximate the Gaussian mixture model parameters for these random variables. That is, Equation (5-17) provides the process uncertainties' specifications required in the priori estimation step, i.e., $\boldsymbol{\mu}^{\boldsymbol{GM}}$ and $\boldsymbol{Q}^{\boldsymbol{GM}}$ (Equation (5-19)); likewise, Equation (5-18) provides $\boldsymbol{\tau}^{\boldsymbol{GM}}$ and $\boldsymbol{R}^{\boldsymbol{GM}}$ to the posterior estimation step (Equation (5-21)). Note that $\boldsymbol{\mu}^{\boldsymbol{GM}}$, $\boldsymbol{Q}^{\boldsymbol{GM}}$, $\boldsymbol{\tau}^{\boldsymbol{GM}}$, and $\boldsymbol{R}^{\boldsymbol{GM}}$ are assumed to remain constant during operation. However, the density of the constrained/ bounded states changes at every time interval throughout the operation. With this in mind, it is essential to update the parameters of the Gaussian mixture model (Equation (5-16)) that approximates the constrained/bounded distribution of the states once the prior estimation of the states is available. As a reminder, the prior estimation step assumes that the states follow unconstrained/unbounded Gaussian distributions, as shown in Equation (5-19). Thereby, Equation (5-20) applies the constraints on the prior distributions of states and consequently, the distributions are no longer Gaussian. That is, $p(\widehat{\boldsymbol{x}}_{k+1|k})$ in

Equation (5-20) represents the non-Gaussian probability density function of states. Hence, the multivariate EM algorithm is implemented to approximate the constrained/bounded prior distribution of the states at each time interval, as shown in Equation (5-20). This intermediate step does not propagate the prior distribution of states (obtained by the prior estimation step) through time. The effect of bounds and constraints on the state variables of the system is taken into account by Equation (5-20), which provides the mean value and covariance of the constrained prior estimation of the states' distribution, i.e., $\widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}}$ and $\boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}}$. Then, the constrained prior estimation of the states ($\widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}}$ and $\boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}}$) is used in the posterior step (Equation (5-21)) to estimate the posterior estimation of the states.

Prior estimation step

$$\begin{aligned}\widehat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{A}_k\widehat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{\mu}^{\boldsymbol{GM}} \\ \boldsymbol{P}_{k+1|k} &= \boldsymbol{A}_k\boldsymbol{P}_{k|k}\boldsymbol{A}_k^{\boldsymbol{T}} + \boldsymbol{Q}^{\boldsymbol{GM}}\end{aligned} \tag{5-19}$$

Intermediate step

$$\begin{aligned}&\boldsymbol{x}^{\boldsymbol{l}} \le \widehat{\boldsymbol{x}}_{k+1|k} \le \boldsymbol{x}^{\boldsymbol{u}} \\ &(\widehat{\boldsymbol{x}}_{k+1|k}^{is}, \boldsymbol{P}_{k+1|k}^{is}) = EM(p(\widehat{\boldsymbol{x}}_{k+1|k}), k) \\ &\widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}} = \textstyle\sum_{is=1}^{ngs} \alpha s^{is}\widehat{\boldsymbol{x}}_{k+1|k}^{is} \\ &\boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}} = \textstyle\sum_{is=1}^{ngs} \alpha s^{is}\boldsymbol{P}_{k+1|k}^{is} + \\ &\qquad \textstyle\sum_{is=1}^{ngs} \alpha s^{is}\left(\widehat{\boldsymbol{x}}_{k+1|k}^{is} - \widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}}\right)\left(\widehat{\boldsymbol{x}}_{k+1|k}^{is} - \widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}}\right)^{\boldsymbol{T}}\end{aligned} \tag{5-20}$$

Posterior estimation step

$$\begin{aligned}&\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}}\boldsymbol{H}_{k+1}^{T}/(\boldsymbol{H}_{k+1}\boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}}\boldsymbol{H}_{k+1}^{T} + \boldsymbol{R}^{\boldsymbol{GM}}) \\ &\widehat{\boldsymbol{x}}_{k+1|k+1} = \\ &\qquad \widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}} + \boldsymbol{K}_{k+1}\left(\boldsymbol{y}_{k+1} - h\left(\widehat{\boldsymbol{x}}_{k+1|k}^{\boldsymbol{GM}}, \boldsymbol{u}_k\right) - \boldsymbol{\tau}^{\boldsymbol{GM}}\right) \\ &\boldsymbol{P}_{k+1|k+1} = (\boldsymbol{I} - \boldsymbol{K}_{k+1}\boldsymbol{H}_{k+1})\boldsymbol{P}_{k+1|k}^{\boldsymbol{GM}}\end{aligned} \tag{5-21}$$

where the EM function in Equation (5-20) is the multivariate EM algorithm; the specification of noises and uncertainties used in that EM algorithm are as follows:

$$\left[(\boldsymbol{w}_{\varsigma}-\boldsymbol{\mu}^{\boldsymbol{GM}})(\boldsymbol{w}_{\varsigma}-\boldsymbol{\mu}^{\boldsymbol{GM}})^{T}\right]=\begin{cases}\boldsymbol{Q}^{\boldsymbol{GM}}, & \varsigma=k\\ \boldsymbol{0}, & \varsigma\neq k\end{cases};\ \boldsymbol{\mu}^{\boldsymbol{GM}}\in\mathbb{R}^{n_x},\boldsymbol{Q}^{\boldsymbol{GM}}\in\mathbb{R}^{n_x\times n_x}$$
$$\left[(\boldsymbol{v}_{\varsigma}-\boldsymbol{\tau}^{\boldsymbol{GM}})(\boldsymbol{v}_{\varsigma}-\boldsymbol{\tau}^{\boldsymbol{GM}})^{T}\right]=\begin{cases}\boldsymbol{R}^{\boldsymbol{GM}}, & \varsigma=k\\ \boldsymbol{0}, & \varsigma\neq k\end{cases};\boldsymbol{\tau}^{\boldsymbol{GM}}\in\mathbb{R}^{n_y},\boldsymbol{R}^{\boldsymbol{GM}}\in\mathbb{R}^{n_y\times n_y} \quad (5\text{-}22)$$

As shown in Equation (5-20), the proposed constrained AGS-EKF scheme involves performing the EM algorithm to complete the point estimation at each time interval k, which adds computational costs when compared to the standard EKF. Further details on the effect of this intermediate step on the CPU time of the constrained AGS-EKF scheme is presented at the end of section 5.2. Moreover, the proposed estimation scheme may not perform well for constrained multivariate systems in which the covariance matrix is very small. In those cases, the estimation of the multivariate Gaussian mixture models (intermediate step in Equation (5-20)) may become challenging as such a small covariance may cause numerical issues while evaluating the likelihood in the EM algorithm.

5.1.3 Gaussian Sum Filter for (GSF) for general non-zero mean non-Gaussian applications

In current chapter, the conventional GSF is used as a benchmark to compare the performance of the proposed constrained AGS-EKF method. Hence, this section provides a brief description on the Gaussian sum filter (GSF). In GSFs, Equations (5-16)-(5-18) are evaluated only once and at the beginning of the process, i.e., at $k=0$. That is, the mean value and covariance of Gaussian components in the Gaussian mixture models of the constrained initial states as well as non-Gaussian process uncertainties and measurement noises are evaluated offline and only at $k=0$. Equations (5-23)-(4-27) represents the formulation of GSF adopted to this study, i.e., nonlinear systems involving bounded states with non-Gaussian distributed process uncertainties and measurement noises.

Prior estimation for the i^{th} component

$$\begin{aligned}\widehat{\boldsymbol{x}}^{i}_{k+1|k}&=\boldsymbol{A}^{i}_{k}\widehat{\boldsymbol{x}}^{i}_{k|k}+\boldsymbol{B}^{i}_{k}\boldsymbol{u}^{i}_{k}+\boldsymbol{\mu}^{i}\\ \boldsymbol{P}^{i}_{k+1|k}&=\boldsymbol{A}^{i}_{k}\boldsymbol{P}^{i}_{k|k}\boldsymbol{A}^{T}_{k}+\boldsymbol{Q}^{i}\end{aligned} \quad (5\text{-}23)$$

Posterior estimation for the i^{th} component

$$\boldsymbol{K}^{i}_{k+1}=\boldsymbol{P}^{i}_{k+1|k}{\boldsymbol{H}^{i}_{k+1}}^{\boldsymbol{T}}/(\boldsymbol{H}^{i}_{k+1}\boldsymbol{P}^{i}_{k+1|k}{\boldsymbol{H}^{i}_{k+1}}^{\boldsymbol{T}}+\boldsymbol{R}^{i}) \quad (5\text{-}24)$$
$$\widehat{\boldsymbol{x}}^{i}_{k+1|k+1}=$$

$$\widehat{\boldsymbol{x}}_{k+1|k}^{i} + \boldsymbol{K}_{k+1}^{i}\left(\boldsymbol{y}_{k+1} - h\left(\widehat{\boldsymbol{x}}_{k+1|k}^{i}, \boldsymbol{u}_{k}^{i}\right) - \boldsymbol{\tau}^{i}\right)$$

$$\boldsymbol{P}_{k+1|k+1}^{i} = \left(\boldsymbol{I} - \boldsymbol{K}_{k+1}^{i}\boldsymbol{H}_{k+1}^{i}\right)\boldsymbol{P}_{k+1|k}^{i}$$

Point estimate for the system states:

$$\widehat{\boldsymbol{x}}_{k+1|k+1}^{\boldsymbol{GSF}} = \sum_{i=1}^{ngt} \beta^{i}\widehat{\boldsymbol{x}}_{k+1|k+1}^{i} \tag{5-25}$$

$$\boldsymbol{P}_{k+1|k+1}^{\boldsymbol{GSF}} = \sum_{i=1}^{ngt} \beta^{i}\left[\boldsymbol{P}_{k+1|k+1}^{i} + \left(\widehat{\boldsymbol{x}}_{k+1|k+1}^{i} - \widehat{\boldsymbol{x}}_{k+1|k+1}^{\boldsymbol{GSF}}\right)\left(\widehat{\boldsymbol{x}}_{k+1|k+1}^{i} - \widehat{\boldsymbol{x}}_{k+1|k+1}^{\boldsymbol{GSF}}\right)^{T}\right] \tag{5-26}$$

$$\beta^{i} = \alpha s^{is}\prod_{l=1}^{n_x}\prod_{c=1}^{n_y}\alpha m_{c}^{im(c)}\,\alpha p_{l}^{ip(l)}\,\Lambda \tag{5-27}$$

where:

$$\Lambda = \frac{1}{(2\pi)^{n_y/2}\left(det\left(\boldsymbol{R}^{i}\right)\right)^{1/2}} \times exp\left[-\frac{1}{2}\left(\boldsymbol{y}_{k+1} - h\left(\widehat{\boldsymbol{x}}_{k+1|k}^{i}, \boldsymbol{u}_{k}^{i}\right) - \boldsymbol{\tau}^{i}\right)^{T}\boldsymbol{R}^{i^{-1}}\left(\boldsymbol{y}_{k+1} - h\left(\widehat{\boldsymbol{x}}_{k+1|k}^{i}, \boldsymbol{u}_{k}^{i}\right) - \boldsymbol{\tau}^{i}\right)\right]$$

$$\sum_{i=1}^{ngt}\beta^{i} = 1; \quad \forall is = 1, \dots, ngs; \forall im = 1, \dots, ngm(c); \ \forall ip = 1, \dots, ngp(l);$$
$$\alpha s^{is}, \alpha m_{c}^{im(c)}, \alpha p_{l}^{ip(l)}, \beta^{i} \geq 0;$$

where $im(c)$ in Equation (5-27) represents the im^{th} Gaussian component in the Gaussian mixture model of the c^{th} measurement noise variable; likewise, $ip(l)$ denotes the ip^{th} Gaussian component in the corresponding Gaussian mixture model of the l^{th} process uncertainty variable. That is, the i^{th} EKF in the set of GSF was performed based on the mentioned components of each process uncertainty and measurement noise. Moreover, the term Λ in Equation (5-27) aims to correct the weights assigned to each EKFs based on the current available measurements. GSF performs individual EKFs for each possible combination of the Gaussian components that are present in the Gaussian mixture models of states, process uncertainties, and measurement noises. As a reminder, a multivariate Gaussian mixture model and two univariate Gaussian mixture models are required to approximate the constrained/bounded states, as well as non-Gaussian process uncertainties and measurement noises, respectively. The total number of unique combinations of the Gaussian components in these three mixtures is $ngt = ngs\prod_{l=1}^{nx}\prod_{c=1}^{ny} ngm(c)\,ngp(l)$. That is, GSF requires to run ngt parallel EKFs to provide the point estimate of the state variables of the system. In addition, β^{i} is a scalar number that

represents the weight assigned to the i^{th} EKF in the set of GSF. This weight is determined based on $\alpha s^{is}, \alpha m_c^{im(c)}, \alpha p_l^{ip(l)}$.

5.2 Computational Experiments

The performance of the proposed constrained AGS-EKF scheme has been tested on three chemical engineering applications, which are presented in this section. In this work, the mean squared error (MSE) defined in Equation (4-22) is used to compare the proposed approach with the conventional GSF and standard EKF, where $n \in \{\text{EKF}, \text{GSF}, \text{constrained } AGS-EKF\}$. Similar to the previous chapters, in the computational experiments presented in this section, the output of the mechanistic process model complemented with additive process uncertainties represent the true value of the states (denoted as "Plant Output"). Note that the "Plant output" does not include the measurement noises. Pyomo 5.6 (Python 3.7) has been used to implement the computational experiments. All the simulations have been performed 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. In current work, the backward method has been used to discretize the nonlinear dynamic process models.

The proposed AGS-EKF framework is a deterministic-based estimation strategy; hence, comparing the performance of this method with stochastic (sampling)-based methods such as PF may not result in a direct and fair comparison since these methods rely on different mathematical principles. Instead, the present approach was compared with other deterministic-based estimation methods that can handle non-Gaussian distributions in the estimation such as GSF. To test the performance of the proposed constrained AGS-EKF three case studies have been considered, i.e., a gas-phase reactor, the Williams-Otto reactor, and a styrene polymerization process (SPP). To show the impact of active constraints on the performance of the estimation schemes considered in this work, the gas-phase reactor has been used as an illustrative small-scale case study considering active/inactive constraints on the states combined with relatively small process uncertainties where both process uncertainties and measurement noises follow zero-mean Gaussian distributions. Note that larger process uncertainties may negatively impact the performance in the estimation. Also, two chemical engineering case studies featuring Williams-Otto reactor and a styrene polymerization process are used in this work to show the benefits offered by the proposed estimation scheme. For the Williams-Otto reactor, constraints are imposed on the state variables and subject to Gaussian process uncertainties and non-Gaussian measurement noises. Likewise, the polymerization process is a highly nonlinear system that was simulated under the assumptions of constrained states, non-Gaussian process uncertainties and measurement noises.

Therefore, each case study provides insights regarding the performance of the proposed estimation framework under different operating conditions often found in chemical plants. Note that the distributions of the process uncertainties and measurement noises considered in the present case studies were defined based on the expected values of the plant states around their nominal operating conditions. The distributions selected for each of the case studies are expected to capture the real operation in an industrial application. Each of these case studies and their results are presented next.

5.2.1 Gas-phase Reactor

The gas-phase reactor introduced in Equation (4-26) consisting of two states (the partial pressures of species A and B) and one measurement (the total partial pressure) is considered to show the performance of the constrained AGS-EKF. This system is a widely used example in the context of state estimation[89,84,174]. In addition to the process model described in Equation (4-26), a set of constraints are imposed on the states of the gas-phase reactor, as follows:

$$p^l \leq (p_A, p_B) \leq p^u \tag{5-28}$$

where p^l and p^u are the lower and upper bounds on the state variables (p_A and p_B), respectively. The additive process uncertainties ($\boldsymbol{w}_k$) and the single measurement noise signal (v_k) in this system are assumed to follow zero-mean Gaussian distributions, i.e.,

$$\begin{aligned} &v_k \sim \mathcal{N}(0, R);\ R = 0.01; \\ &\boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{Q});\ \boldsymbol{Q} = diag([1 \times 10^{-6}, 1 \times 10^{-6}]); \end{aligned} \tag{5-29}$$

The time interval considered on this case study is 0.1 s. After conducting preliminary tests, three Gaussian components are considered to form the Gaussian mixture model that approximates the original constrained distribution of the states with an appropriate accuracy. Two scenarios have been considered for this case study. The first scenario (*Scenario I*) aims to assess the performance of all three estimation schemes (standard EKF, GSF, and constrained AGS-EKF) for the case that the constraints on the states are inactive. That is, this scenario considers that the process operates far from the infeasibility region, i.e., near the bounds on the states shown in Equation (5-28). Moreover, *Scenario II* is considered to evaluate the performance of the proposed constrained AGS-EKF compared to standard EKF and GSF for the case in which the constraints on the states are active. In both scenarios the true values of the states have been used to initialize the estimation schemes. The results obtained for each scenario are discussed next.

Scenario I: inactive constraints

This scenario considers that $p^l = 0$ and $p^u = 10$. Figure 5-1 shows the estimations obtained with the standard EKF, conventional GSF, and the constrained AGS-EKF methods. As expected, the three estimation schemes performed well since all the constraints present in the system are inactive and consequently, an unconstrained approach such as the standard EKF can estimate the states accurately. Ideally, both GSF and constrained AGS-EKF are expected to have high performance in the presence of active/inactive constraints on the states. As shown in Figure 5-1, three of the approaches provide similar and accurate estimates after almost 60 s (600 iterations). Note that GSF begins with individual Gaussian components in the Gaussian mixture model of the initial states, which drive the system away in the first iteration. Then, due to the weight correction step in GSF, the estimator can recover from the initial deviation within the first 60 s of the operation. Nevertheless, the GSF performance improved over the time and converged.

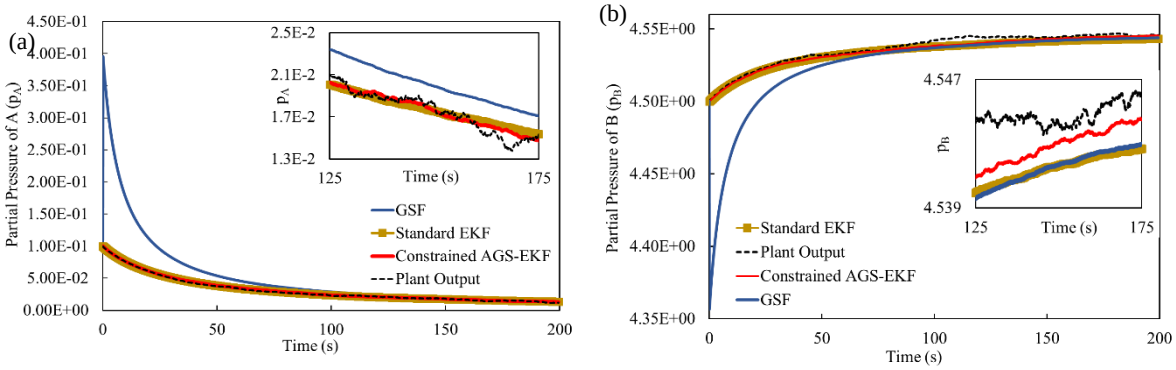


Figure 5-1: Estimation provided by EKF, GSF, and constrained AGS-EKF for (a) p_A; (b) p_B given *Scenario I*

Scenario II: active constraints

In the present scenario, as an illustrate proof-of-concept of the present estimation framework, a tight bound on the state variables was considered, i.e., $p^l = 0.05$ and $p^u = 10$. These constraints were added to evaluate the performance of the proposed constrained AGS-EKF, GSF, and standard EKF when the process operates near the feasibility limits. Note that the low partial pressure constraint on reactant A could represent the hypothetical case that an input in the system operates at a saturation limit thus producing this hard constraint on p_A. Note that the mechanistic model for this process shown in Equations (4-28) and (5-28) does not include inputs to the system.

Figure 5-2 compares the estimations provided by the estimation schemes considered in this study. As expected, standard EKF was not able to take into account the constraints on the states and the estimates provided by EKF for p_A violates the lower bound (Figure 5-2(a)). Consequently, the estimation provided for p_A by standard EKF is not accurate either, as shown in Figure 5-2(b). Likewise, the estimations provided by GSF were inaccurate and violated from the lower bound on p_A. This is because GSF performs individual EKFs based in each Gaussian component in the Gaussian mixture model of the initial states. The Gaussian mixture approximates the distribution of the constrained states at the initial time. In this way, GSF guarantees that the distribution of the initial states is bounded properly. Then, GSF propagates each of the Gaussian components in the Gaussian mixture model of the initial states by performing individual EKFs. However, GSF cannot ensure that the states remain within their feasible limits because it does not explicitly take into account these constraints in their framework during the operation of the process. As shown in Figure 5-2(a), GSF is not able to maintain p_A within their operational limits. Moreover, running individual EKFs may lead to biased/infeasible estimates; in particular for those EKFs that perform the estimation based on the Gaussian components describing the edges of the overall non-Gaussian distribution of the states, i.e., Gaussian components located at the extreme left or right side of the distribution. Note that Equation (5-27) in the GSF scheme is used to re-evaluate the weights on each Gaussian component (each EKF in the set) to lower the effect of the biased estimations in the overall point estimates. For this case study though, this re-evaluation step set all the weights on the EKFs to zero, except for the one EKF with the weight equal to 1. As a result, GSF provides the point estimates based on the estimations provided by only one the EKFs, which makes the estimation biased as it does not consider other partitions of the distribution of the states.

The proposed constrained AGS-EKF takes into account the constraints on the states since it approximates the constrained distribution of the states not only at the initial time, but also at each sampling interval *k*. As shown in Figure 5-2(a) the estimates resulting from the constrained AGS-EKF complies with the bounds on the states thus providing an accurate estimation for p_A. Moreover, the constrained AGS-EKF considers the overall distribution of states at any time interval and avoids both losing partial information and biased estimations observed in the case of using GSF. According to Figure 5-2(b), the constrained AGS-EKF can accurately estimate p_B. The estimation errors evaluated using different estimation schemes confirmed that the estimation accuracy improved significantly when the constrained AGS-EKF was used as the estimator. For instance, $MSE_{p_B}^{EKF}$ (5.1e-4) and $MSE_{p_B}^{GSF}$ (4.69e-4) are two orders of magnitude larger than $MSE_{p_B}^{AGS-EKF}$ (3.19e-6).

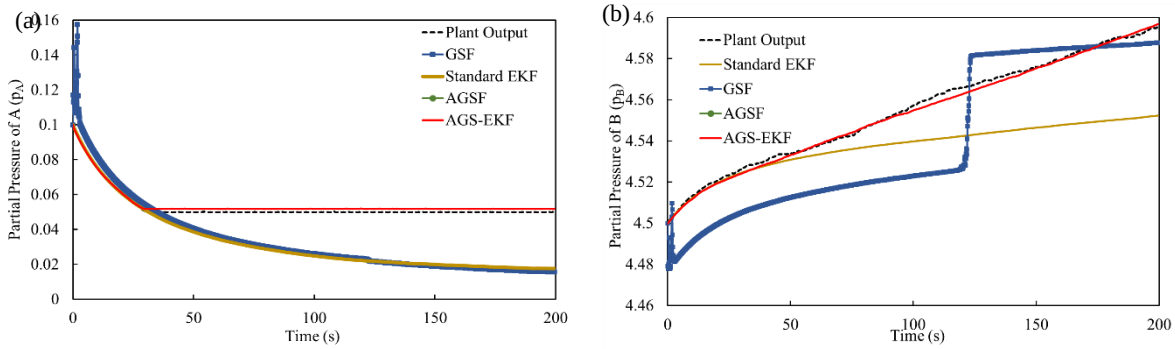


Figure 5-2: Estimation provided by EKF, GSF, and constrained AGS-EKF for (a) p_A; (b) p_B given *Scenario II*

5.2.2 Williams-Otto Reactor

To further investigate the performance of the proposed approach, the Williams-Otto reactor presented in Equation (4-28) has been considered. Note that the Williams-Otto reactor is regarded as a highly nonlinear dynamic system that is widely used in studies involving online control and state estimation[171,170,169]. According to section 4.2.3, the implementation of AGS-EKF for this process showed that this estimation scheme offers a promising performance for the unconstrained systems, with zero-mean Gaussian measurement noises, and in the presence of non-Gaussian process uncertainties[35]. The current work aims to examine the performance of the constrained AGS-EKF for the constrained Williams-Otto reactor in the presence of non-zero mean non-Gaussian measurement noises while the process uncertainties are assumed to follow zero-mean Gaussian distributions. For this goal, Equation (5-30) represents the bounds on the state variables, i.e., all the mass fractions have a lower bound of 0.1 and an upper bound of 1.

$$0.1 \leq X_A, X_B, X_C, X_E, X_G, X_P \leq 1 \qquad (5\text{-}30)$$

These bounds are considered to investigate the performance of the proposed constrained AGS-EKF, GSF, and standard EKF for the cases in which the process operates near the feasibility region. The random variables present in the system, i.e., process uncertainties and measurement noises, are represented by mutually independent probability distributions. The random process uncertainty associated with each state follow a zero-mean Gaussian distribution with 1% of the steady-state value of the corresponding state as its standard deviation. The nominal values for the initial states are provided in Table 5-1.

Table 5-1: Base case values of states for the Williams-Otto reactor

Process Variables	Base case value
X_A	0.100
X_B	0.399
X_C	0.150
X_E	0.141
X_G	0.110
X_P	0.105

As mentioned in section 4.2.3, online measurements are assumed to be available for the mass fractions of the reactants A and B as well as the product P, i.e., X_A, X_B and X_P. The unknown states of this system are X_C, X_E, and X_G. To make this study more realistic, the available measurements are assumed to be corrupted with non-zero mean non-Gaussian measurement noises, as shown in Figure 5-3. The procedure to identify the required number of Gaussian components in the Gaussian mixture model consists of a one-to-one increase in the number Gaussian components until a Gaussian mixture model that captures the main features of the actual non-Gaussian density is detected. Note that this test is performed offline (i.e., prior to the implementation of the proposed framework). The key metrics used to determine the suitability of the number of Gaussian components are quality in the representation of the non-Gaussian distribution, quality in the state estimation and the overall computational effort. In general, adding Gaussian components improves the representation of the Gaussian mixture model thereby improving the overall state estimation at the expense of additional computational costs needed to evaluate the point estimates in the estimation scheme. Note that performing these tests are problem specific. However, this test is not intensive since the non-Gaussian density of the noises/uncertainties is assumed to be known a priori. 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. According to the author's own experience, three Gaussian components may be an appropriate choice to describe bounded Gaussian distributions. A detailed analysis on this topic was presented in section 4.2.3.1.

Based on above, five, two, and three Gaussian components are needed to form Gaussian mixture models representing the distributions of measurement noises associated with X_A, X_B, X_P, respectively. The results of the tests performed to make this set of assumption has not been presented here for simplicity.

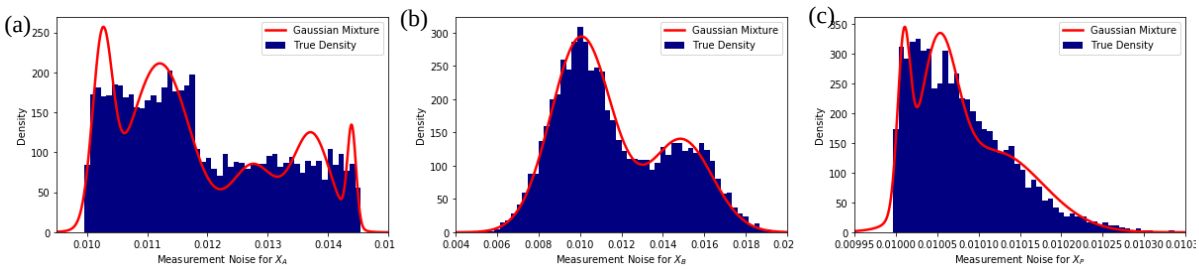


Figure 5-3: Histogram of the non-Gaussian measurement noises present in the Williams-Otto reactor and the approximated density by Gaussian mixture model for (a) X_A; (b) X_B; (c) X_P

The initial condition used to perform state estimations was assumed to be 5% larger than the true values of corresponding initial state. The performance of each estimation scheme is presented in Figure 5-4. This figure highlights the results for the unknown states, i.e., X_C, X_E, X_G, as well as the measurable state X_P. Moreover, Table 5-2 reports the MSE in the state estimates provided by the estimation schemes considered in this work. As shown in Figure 5-4(a), the standard EKF failed to estimate X_C accurately because standard EKF does not consider the non-Gaussianity assumption for the measurement noises and the lower bound on this state. For the same reasons, the estimation error for the other two unknown states is relatively large in the case of using standard EKF, as shown in Figure 5-4(b)-(c). The estimations provided by GSF for the three unknown states are not accurate either. Similar to the discussion presented for the first case study (see section 5.2.1-*Scenario II*), GSF is not able to comply with the constraints on the states for the Williams-Otto reactor because it does not explicitly take into account the bounds on the states at each time interval k. In addition, biased estimations are expected in GSF as it performs individual EKFs that led to constraint violations in some of the EKFs. Although the re-evaluation of the weights assigned to each EKF in the GSF framework helps to overcome the latter issue, GSF scheme discarded all the EKFs in the set from the estimation scheme except for one, which makes the estimation biased. Figures 5-4(a)-(c) illustrates that the proposed constrained AGS-EKF estimation scheme exhibits an acceptable performance. This is because the constrained AGS-EKF takes into account the constraints on the states, explicitly, to modify the prior estimation of the states' distribution based on Equation (5-20). By approximating the constrained priori distribution of the states to a multi-variate Gaussian mixture model at each time interval, the constrained AGS-EKF is able to satisfy the bounds on the states thereby improving the estimation of the states significantly. The MSE reported in Table 5-2 supports these observations. For instance, the error in the estimation for X_C offered by the proposed constrained AGS-EKF is three orders of magnitude smaller than that obtained by performing standard EKF and GSF, as shown in

Table 5-2. Moreover, both Figure 5-4(d) and Table 5-2 show that GSF and standard EKF provided smaller errors in the estimation of the measurable state X_P, when compared to the constrained AGS-EKF. This is due to the non-Gaussian measurement noises considered in the system. As the noises are small, the estimation scheme often tends to follow the online measurements rather than the plant outputs. As shown in Figure 5-4(d), the estimation provided by the constrained AGS-EKF closely follows the measurements whereas the standard EKF and GSF rely on the process model and tend to follow the plant outputs. This is because the constrained AGS-EKF uses the overall univariate Gaussian mixture models that approximate the corresponding non-Gaussian distributions of the measurement noises. Hence, the proposed approach avoids the Gaussian assumption as in EKF, and it does not eliminate any partition of the Gaussian mixture models of the noises as it is performed in the GSF scheme. Thus, constrained AGS-EKF has access to the reliable measurements with accurate information on the noises' distributions, and as a result, relies on the available measurements for X_P, and use that information to estimate the unknown states. Furthermore, Table 5-2 also shows that the estimations obtained by the different estimation schemes for the other two measurable states (X_A, and X_B) are similar, i.e., there is an approximately $\sim 0.009\%$ and $\sim 0.4\%$ difference in between the estimation errors obtained using various estimation schemes for X_A and X_B, respectively.

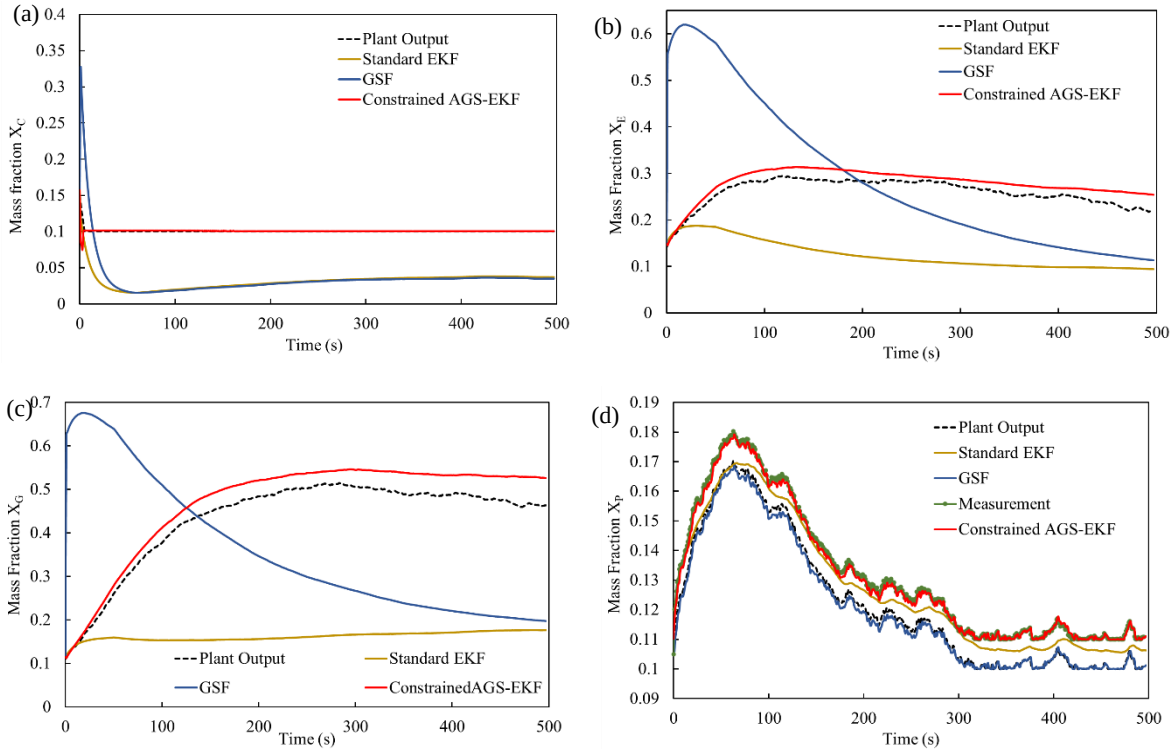


Figure 5-4: Estimation provided by EKF, GSF, and constrained AGS-EKF for (a) X_C; (b) X_E; (c) X_G; (d) X_P

Table 5-2: MSE for X_A, X_B, X_C, X_E, X_G, and $X_P$using different estimation schemes

Estimation method (n)	$MSE_{X_A}^{(n)}$	$MSE_{X_B}^{(n)}$	$MSE_{X_C}^{(n)}$	$MSE_{X_E}^{(n)}$	$MSE_{X_G}^{(n)}$	$MSE_{X_P}^{(n)}$
Standard EKF	8.62e-5	4.04e-2	4.73e-3	1.97e-2	8.11e-2	3.58e-5
GSF	1.26e-5	3.48e-2	5.15e-3	2.87e-2	6.42e-2	2.26e-5
Constrained AGS-EKF	9.65e-5	3.86e-2	1.41e-6	4.05e-4	1.59e-3	1.13e-4

5.2.3 Styrene Polymerization Process

Styrene polymerization process is a constrained complex nonlinear dynamic process featuring a continuous manufacturing of styrene. Due to its inherent complexity and nonlinearity, this polymerization reactor makes an attractive case study for the current work. Note that the styrene polymerization system has been used to show the performance of state estimation approaches presented in the literature including GSF-based estimation schemes[98,86,175,176]. The dynamic model for this process is as follows:

$$\frac{dC_i}{at} = -\left(\frac{F_t}{V} + k_i\right) C_i + \frac{F_i C_{i_i}}{V}$$

$$\frac{dC_s}{at} = -\frac{F_t C_s}{V} + \frac{F_i C_{s_i} + F_m C_{s_m}}{V}$$

$$\frac{dC_m}{at} = -k_p C_m P + \frac{F_m C_{m_m} - F_t C_m}{V} \tag{5-31}$$

$$\frac{d\lambda_1}{at} = -\frac{F_t \lambda_1}{V} + \left(\left(k_{fm} C_m + k_{t_d} P + k_{fs} C_s\right)(2\alpha - 2\alpha^2) + k_{t_c} P\right) \frac{P M_n}{(1-\alpha)}$$

$$\frac{dM_n}{at} = \left\{\left(\left(k_{fm} C_m + k_{t_d} P + k_{fs} C_s\right)(2\alpha - 2\alpha^2) + k_{t_c} P\right) M_n - \left(\left(k_{fm} C_m + k_{t_d} P + k_{fs} C_s\right)\alpha + 0.5 k_{t_c} P\right) M_n (1-\alpha)\right\} \frac{P M_n}{\lambda_1 (1-\alpha)}$$

$$\frac{dM_w}{at} = \left\{\left(\left(k_{fm}C_m + k_{t_d}P + k_{fs}C_s\right)(\alpha^3 - 3\alpha^2 + 4\alpha) + +k_{t_c}P(\alpha + 2)\right)M_n - \left(\left(k_{fm}C_m + k_{t_d}P + k_{fs}C_s\right)(2\alpha - 2\alpha^2) + k_{t_c}P\right)M_n(1 - \alpha)\right\}\frac{PM_n}{\lambda_1(1-\alpha)^2}$$

where,

$P = \sqrt{2fC_ik_i/k_t}$; $F_t = F_i + F_m$;

$k_i = 0.693/\left(60 \times 10^{\left((A_i/T)+B_i\right)}\right)$; $k_p = A_p exp\left(\frac{-E_p}{RT}\right)$; $k_{fm} = A_{fm}exp\left(\frac{-E_{fm}}{RT}\right)$;

$k_{fs} = A_{fs}exp\left(\frac{-E_{fs}}{RT}\right)$; $k_t = A_t exp\left(\frac{-E_t}{RT}\right)$; $k_{t_d} = 0.15k_t$; $k_{t_c} = 0.85k_t$

Note that the description and the nominal value for all model parameters are presented elsewhere[86,175]. The styrene polymerization process presented in Equation (5-31) has six state variables, namely the initiator C_i, solvent C_s and monomer C_m concentrations inside the polymerization reactor, the first moment of dead polymer molecular weight λ_1, the number average molecular weight M_n, and the weight average molecular weight M_w. The online measurements assumed to be available for process are C_i, M_n, and M_w. The remaining three states need to be estimated. The linear observability matrix for this setting is full-rank; thus, the system is observable around the initial operating condition considered for this case study. The sampling interval for this process is set to 1 s. The process uncertainties and measurement noises are randomly chosen from their corresponding non-Gaussian distributions. The process uncertainties (and the measurement noises) are mutually independent. The non-Gaussian distribution of the process uncertainty/measurement noise associated with each state is presented in Figure 5-5. Each non-Gaussian distribution has been approximated to a Gaussian mixture model (red solid line in Figure 5-5) using an adequate number of Gaussian components. Moreover, the feasibility bounds on the states of the system as well as the nominal values of the initial states are presented in Table 5-3. Note that the state bounds listed in Table 5-3 are considered in the current work with the sole purpose to assess and illustrate the performance and benefits of the proposed constrained AGS-EKF on a highly complex and nonlinear system. These bounds are not expected to represent the actual operation of this process. Similar to the previous case studies, GSF has considered three Gaussian components in multivariate Gaussian mixture model to represent an adequate approximation of the bounded distribution of the initial state variables. All of the estimation schemes have been initialized

by considering 95% of the true initial states (95% of the values reported in “base-case value” column heading in Table 5-3).

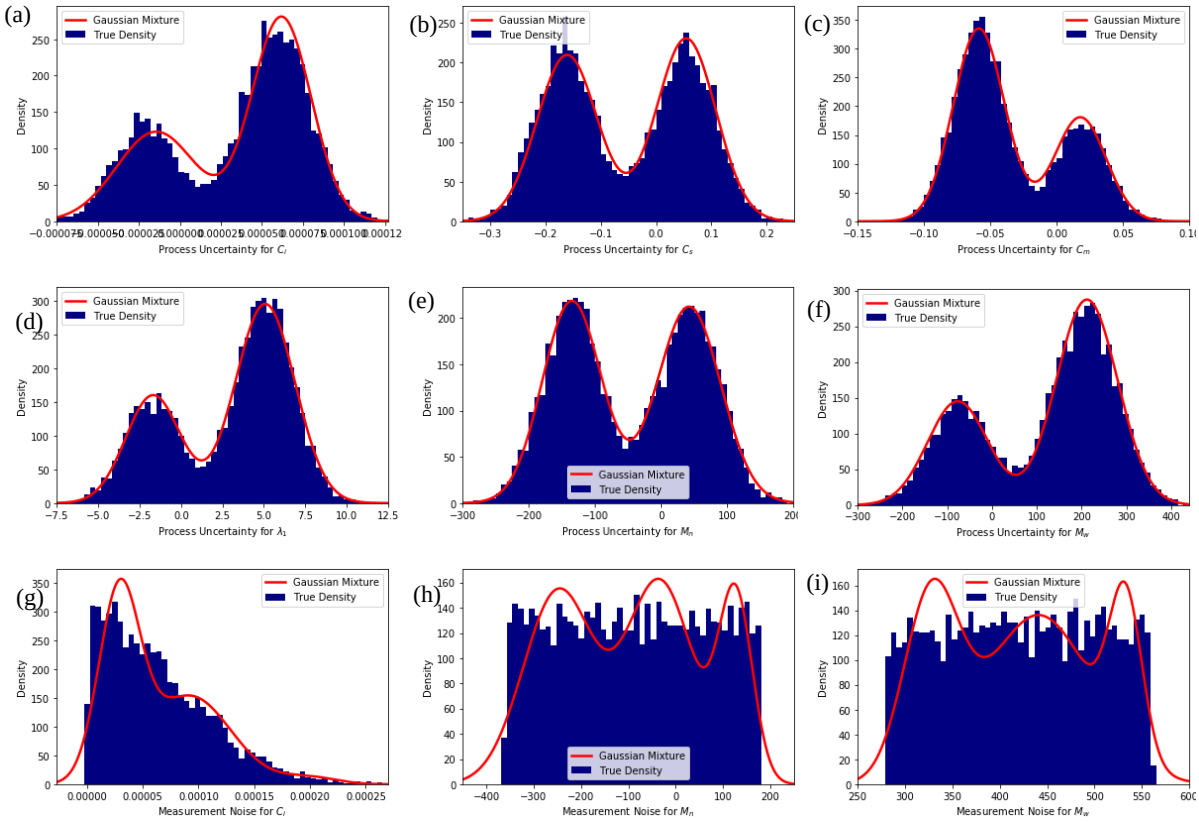


Figure 5-5: Histogram for the original non-Gaussian distribution and the corresponding Gaussian mixture model of the process uncertainties and the measurement noises in styrene polymerization process

Table 5-3: Initial condition and bounds on the states

Process Variable	Base case value[98]	Lower bound	Upper bound
C_i (kmol/m^3)	1.9854e-3	0.001	1
C_s (kmol/m^3)	5.475	0	15
C_m (kmol/m^3)	1.9408	0	3
λ_1(kg/m^3)	1.7e+2	0.01	500
M_n (kmol/m^3)	4.5e+3	100	25000
M_w (kg^2/kmol.m^3)	7.0e+3	100	40000

Figure 5-6(a)-(c) shows the results obtained for C_i and C_m, and λ_1, respectively, which are the states with active constraints during operation. Figure 5-6(d) presents the estimation obtained for an unknown state (C_s) that is far from its upper bound limit. Moreover, Table 5-4 compares the estimation errors obtained from each estimation scheme. According to Figure 5-6(a), the estimation provided by GSF violated from the lower bound on C_i; this is because GSF does not consider the bounds on the states at every time interval k. The standard EKF satisfies this constraint as the online measurement is available for this state and the estimator relies on this measurement. However, the error in the estimation obtained for C_i using standard EKF is larger than that resulted by performing the constrained AGS-EKF, as shown in Table 5-4. Both Figure 5-6(a) and Table 5-4 confirm that the constrained AGS-EKF offers the highest estimation accuracy for C_i among the estimation schemes considered in this study. That is, $MSE_{C_i}^{(AGS-EKF)}/MSE_{C_i}^{(EKF)}$ and $MSE_{C_i}^{(AGS-EKF)}/MSE_{C_i}^{(GSF)}$ are approximately 0.09 and 0.49, respectively, which shows the improvement achieved by the proposed estimation scheme. The estimates for the state C_i satisfied the lower bound on this state as the constrained AGS-EKF explicitly takes into account the constraints on the states at each sampling interval by re-approximating the constrained prior estimation of the states' distribution. As for the states C_m and λ_1, the online measurements were not available; hence, the estimates provided by standard EKF for C_m and λ_1 were not able to follow the plant outputs, as shown in Figures 5-6(b) and 5-6(c). Moreover, the first estimation provided by GSF is associated with a large error for C_m and λ_1. Similar to the discussion presented for *Scenario I* in section 5.2.1, this is because GSF starts from a Gaussian mixture model of the bounded states at the initial time interval. The first point estimates are determined by propagating these Gaussian components of the initial states for one time interval, which does not check for the active constraints on the states. The weight correction step in the GSF framework was not able to resolve this issue and made all the weights equals to zero expect for only one of the EKFs in the set. Following this event, the initial estimation error coupled with the fact that GSF does not consider constraints in the formulation led to inaccurate estimations for C_m and λ_1, which are biased and do not comply with the bounds on these states. Note that the similar observations have been made in the estimations provided by GSF in section 5.2.1 (*Scenario II*) and section 5.2.2 for applications featuring active constraints on the states. Nevertheless, the constrained AGS-EKF scheme satisfied the bounds on C_m and λ_1 and provided appropriate estimations for these states, as shown in Figures 5-6(b) and 5-6(c), respectively. According to Table 5-4, the constrained AGS-EKF decreased the estimation error of C_m by two orders of magnitude when compared to the error resulted by standard EKF and GSF. Likewise, performing

the constrained AGS-EKF reduced the estimation error by two and one orders of magnitude than that resulted by performing GSF and the standard EKF, respectively. As for the unmeasurable state C_s, Figure 5-6(d) shows that the standard EKF was not able to provide proper estimates for this state due to the absence of information about the non-Gaussian process uncertainties in the standard EKF framework, whereas both GSF and the constrained AGS-EKF provided appropriate estimations for C_s. As shown in Figure 5-6(d), GSF can provide proper estimations for states that are far from the constraints, which is similar to the performance of GSF observed in *Scenario I* in section 5.2.1. Table 5-4 also shows that the constrained AGS-EKF reduced the estimation error for all the other states by at least one order of magnitude in comparison to the MSE evaluated for standard EKF and GSF. Note that the estimates provided by the estimation schemes for the unmeasurable states are smooth (Figures 5-6(b)-(d)). In all the estimation schemes, the prior estimation is performed based on the plant model that considers no noise/uncertainties (in both EKF and the modified EKF formulation). That is, the estimates to the unknown states often seem to be smooth and free of noise. The stronger correlation between the states and the measurements makes the state estimates noisier. For instance, the estimates provided by the schemes for the measurable state C_i (see Figure 5-6(a)) are associated with noises as the estimation is strongly affected by the noisy measurements present in the posterior estimation step.

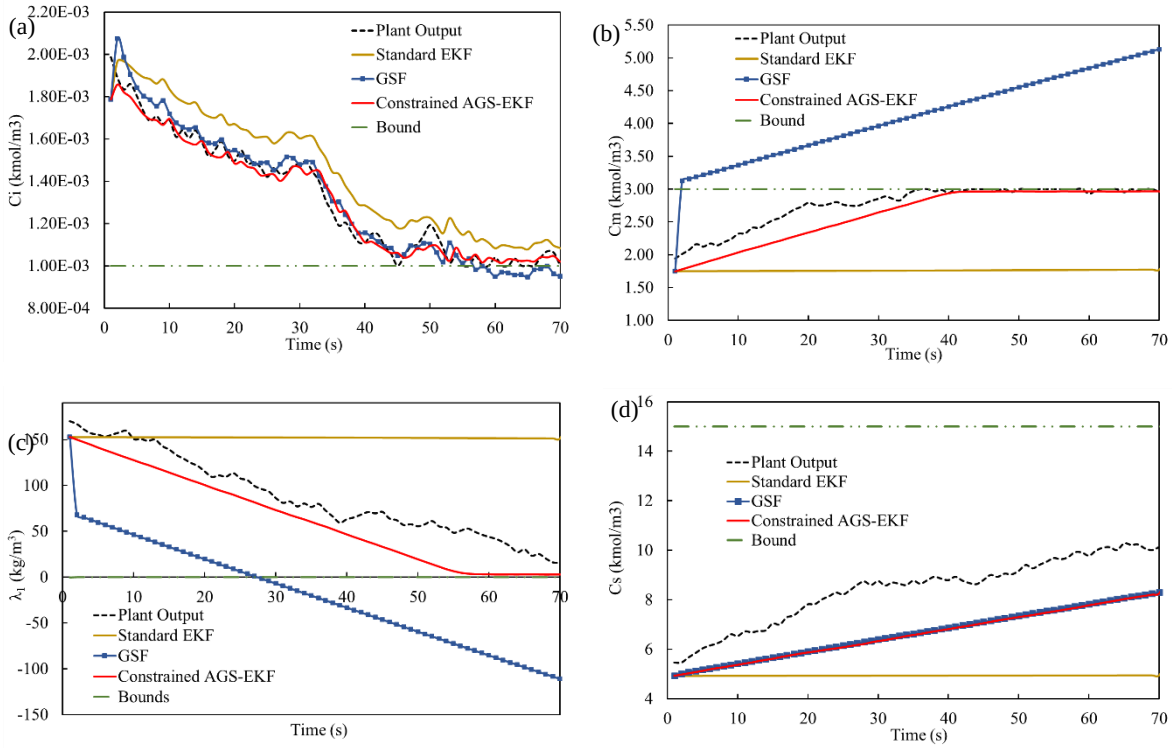


Figure 5-6: Estimation provided by EKF, GSF, and constrained AGS-EKF for (a) C_i; (b) C_m; (c) λ_1; (b) C_s

Table 5-4: MSE for C_A, C_s, C_m, λ_1, M_n, and $M_w$using different estimation schemes

Estimation method (n)	$MSE_{C_i}^{(n)}$	$MSE_{C_s}^{(n)}$	$MSE_{C_m}^{(n)}$	$MSE_{\lambda_1}^{(n)}$	$MSE_{M_n}^{(n)}$	$MSE_{M_w}^{(n)}$
Standard EKF	1.98e-8	1.39e+1	1.11	5.56e+3	9.530e+4	2.52e+5
GSF	3.93e-9	3.33	1.95	1.14e+4	4.45e+6	7.82e+4
Constrained AGS-EKF	1.94e-9	3.54	4.16e-2	7.29e+2	1.97e+4	1.37e+3

Note that the results presented in this section are based on the bounds considered in the current work, which are not the bounds reported on the actual styrene polymerization process. In fact, the bounds on the process are significantly wider than those considered in this scenario and consequently, the actual process may operate differently. Initially, this study performed the estimation using the original setting of the process and observed that all three estimation schemes perform adequately when the constraints on the states are not active (which is the case in the actual setting of this process). Hence, the outcomes of this study can be used as proof-of-concept of the present estimation framework when active constraints are present in complex nonlinear systems subject to non-Gaussian distributions. Note that the number of Gaussian components in the multivariate Gaussian mixture of the states is three. Each univariate Gaussian mixture model for the corresponding process uncertainty consists of two Gaussian components. Three Gaussian components are present in the Gaussian mixture model of each measurement noise signal. Given this set of assumptions and based on the discussion presented in section 5.1.3, the number of EKFs in the set of GSF is $5{,}184$ (i.e., $3 \times (2^6 \times 3^3)$). Thus, it is expected to observe a significant increase in the CPU time for the case of performing GSF in comparison to standard EKF. On the other hand, the proposed AGS-EKF performs EKF only once at each time interval and does not update the Gaussian mixture model of the process uncertainties and measurement noises. However, the intermediate step in the constrained AGS-EKF framework involves running the EM algorithm at each time interval k (Equation (5-20)), which leads to an increase in the required CPU time when compared to standard EKF. Nevertheless, the additional CPU time required in the constrained AGS-EKF is not expected to be significant as the EM algorithm is a computationally efficient approach, i.e., 1.74 s on average per sampling interval. The modifications conducted to the prior and posterior estimation steps in the constrained AGS-EKF does not add any computational costs than that required in the standard EKF. This means that the difference between the CPU time reported by performing the standard EKF and the constrained AGS-EKF schemes represents the CPU time

required to perform the EM algorithm in the intermediate estimation step. Table 5-5 reports the averaged CPU time for point estimate using various estimation schemes. As expected, performing 5,184 parallel EKFs in the set of GSF leads to a significantly larger CPU time than that required for the proposed constrained AGS-EKF and standard EKF. According to Table 5-5, the constrained AGS-EKF increases the average CPU time by one order of magnitude in comparison with standard EKF. Nevertheless, the proposed estimation scheme decreased the required CPU time by one order of magnitude than that required by GSF for this polymerization process.

Table 5-5: Computational costs reported for different estimation schemes

Estimation method	Averaged CPU time per sampling interval
Standard EKF	0.14 s
GSF	12.52 s
AGS-EKF	2.72 s

5.3 Summary

This chapter introduced a novel EKF-based estimation scheme referred to as Constrained Abridged Gaussian sum extended Kalman filter (constrained AGS-EKF). This framework aims to improve the state estimation for general applications involving bounds on the states, non-zero non-Gaussian process uncertainties and measurement noises. Note that the distribution of a constrained state is more likely non-Gaussian and may follow arbitrary probability distribution functions. In the constrained AGS-EKF framework, Gaussian mixture models are used to represent an adequate approximation of the original non-Gaussian distributions of states/uncertainties/noises. The proposed framework performs a modified version of EKF based on the main characteristics of the overall Gaussian mixture models to capture the non-Gaussianity present in the process. These adaptations involve changes in the prior and posterior estimation steps to capture the non-Gaussianity of the uncertainties and noises, as well as considering an additional intermediate step in the EKF framework to take into account the constrains on the states. The latter requires additional computational costs than that in the standard EKF as this intermediate step perform the EM algorithm to approximate the Gaussian mixture model of the constrained estimates states online. Note that the additional averaged CPU time required to perform the point estimates using the constrained AGS-EKF is not significant when compared to the case of using conventional GSFs. Moreover, the proposed approach avoids biased estimations as observed in the

GSFs. The results obtained from the constrained AGS-EKF, GSF and the standard EKF show that the constrained AGS-EKF method improved the estimation accuracy significantly. This improvement holds under different scenarios, i.e., active/inactive constraints on the states, for the chemical engineering case studies considered in this work. Another standard state estimation scheme that can successfully take into account the process constraints is moving horizon estimation (MHE). The next chapters of the thesis focus on addressing the knowledge gaps mentioned in chapter 2 in the context of the MHE framework.

Chapter 6

Assessing the Impact of EKF as the Arrival Cost estimator in Moving Horizon Estimation under Nonlinear Model Predictive Control

The aim of this chapter is to assess the impact of arrival cost (AC) on an NMPC-MHE closed-loop framework for industrial and complex applications under scenarios that may occur during operation. In particular, this study assesses the capability of EKF (when compared to Constrained Particle Filter (C-PF)) as the AC estimator engaged with MHE under the cases that EKF may fail as an individual state estimator. Previous studies have shown that EKF may fail to provide an accurate estimation of the states for the constrained nonlinear systems. Considering this fact, this study aims to point out the potential of EKF as AC estimator coupled with MHE and NMPC to perform online estimation and control of large and challenging processes. To address this point, an error analysis on the performance of EKF as AC estimator is presented to gain insight on the convergence of this method as AC estimator in the MHE-NMPC closed-loop scheme. In addition, this work investigates the effect of key operational factors such as plant sizes and non-symmetric bounded probability distributions on the process uncertainty. Multiple aspects such as computational costs, accuracy in the estimation and closed-loop performance under a wide variety of scenarios have been used to gain new insights on the effect of EKF as an AC estimator. To pursue these objectives, an industrial Wastewater Treatment Plant (WTP) and a High Impact Polystyrene (HIPS) process, which has been identified as a challenging (open-loop unstable) industrial process, have been considered. Although multiple studies have individually assessed the controllability and estimation of WTP[177,178,179,180], studies focusing on the operation of WTP using an MHE-NMPC closed-loop framework are lacking. Moreover, multiple studies have proposed optimal open-loop control schemes for HIPS[181,182,183]; to the author's knowledge, this is the first study that presents an MHE-NMPC closed-loop framework for this challenging process.

Note that this research was performed prior to the AGS-EKF and constrained AGS-EKF. Hence, the standard EKF has been considered as the AC-estimator in the MHE framework. The structure of this chapter is organized as follows: section 6.1 presents the closed-loop framework to assess the impact of AC approximation. Section 6.2 presents the computational experiments preformed in this research for two case studies, i.e., WTP and HIPS, under different scenarios. Moreover, an error analysis section

that provides insights on the performance of EKF as AC estimator is presented at the end of this section. Chapter summary is presented at the end.

6.1 Closed-loop framework

An adequate state estimation technique can provide a good approximation to the arrival cost (AC) and therefore improve the estimation in the MHE framework. This can also lead to a reduction in the length of the horizon N, and therefore a reduction in CPU costs, which are attractive features for online control of large-scale applications when engaged with an NMPC framework. Providing an accurate initial condition is key to attain good closed-loop performance. Thus, an acceptable AC approximation can help MHE to provide fast and accurate estimations to NMPC. Figure 6-1 illustrates the feedback control system considered in this study. At every time step k, the plant updates the state output vector $\boldsymbol{x}_k$ (subject to process uncertainties $\boldsymbol{w}_k$) using the past optimal control actions obtained from the NMPC. The sensors in the plant update the measurement output vector $\boldsymbol{y}_k$, which may include measurement noise $\boldsymbol{v}_k$. Then, MHE is engaged and used to provide the estimations of unmeasurable states $\hat{\boldsymbol{x}}_k$, which are then used as the initial condition in NMPC to compute the control actions for the next time interval. An estimator such as EKF or C-PF needs to be coupled with the MHE framework to provide the approximation of the AC parameters, i.e., expected value ($\overline{\boldsymbol{x}}_{k-N}$) and covariance matrix($\boldsymbol{P}_{k-N}$) of the approximated AC distribution for the states. $k-N$ is the initial time interval in the finite estimation horizon window (N) considered in the MHE framework. As shown in Figure 6-1, the AC estimator uses the outcome of MHE at $(k-N-1)^{th}$time interval (denoted by $\hat{\boldsymbol{x}}_{k-N-1}$) as the prior estimation of AC to estimate $\overline{\boldsymbol{x}}_{k-N}$. The AC approximation method estimates the covariance of the AC distribution $\boldsymbol{P}_{k-N}$ in a recursive form, i.e., $\boldsymbol{P}_{k-N}$ is estimated from the AC calculations performed in the previous time interval (i.e., $k-N-1$). For instance, the posterior covariance computed by EKF is used in MHE as the approximation of AC parameter $\boldsymbol{P}_{k-N}$. As the time updates to $k+1$, approximation of the AC covariance $\boldsymbol{P}_{k-N+1}$will be performed using $\boldsymbol{P}_{k-N}$ obtained by the AC estimator at the $K-N$ time interval. Note that $\hat{\boldsymbol{x}}_{k-N-1}$ represents the MHE estimation, which includes the measured and unknown states. Once the NMPC receives the initial condition $\hat{\boldsymbol{x}}_k$, it uses this information to find the next optimal control actions that minimizes the offset between the controlled variables and their reference set points $\boldsymbol{y}_{\boldsymbol{ref}}$ over a prediction horizon (L). The new control actions $\boldsymbol{u}_k$ are implemented in the plant to complete the feedback control loop sequence. This algorithm is implemented in the same fashion at subsequent time intervals. In addition to the closed-loop framework shown in Figure 6-1, NMPC coupled with EKF

as the state estimation method (i.e., with no MHE involved) is often used in the literature[53,54]. Hence, this approach will also be considered in this work for comparison purposes (see sections 6.2.1.2 and 6.2.2.1).

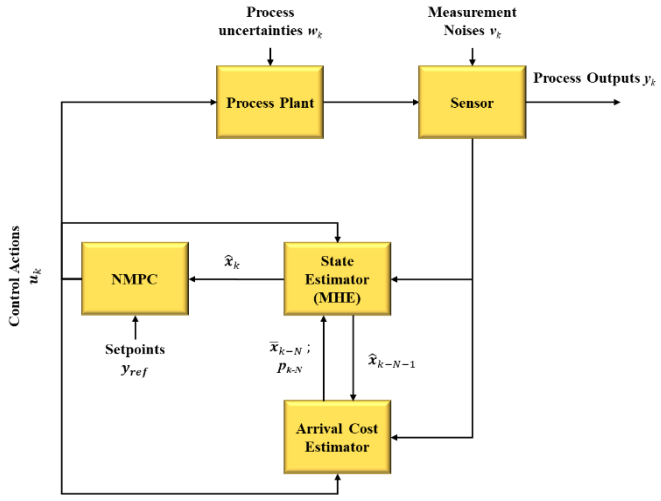


Figure 6-1: Block diagram of closed-loop feedback control

In the current study, both EKF and C-PF are considered first to evaluate an approximation of AC distribution. EKF is known as the most common AC approximation method due to its simplicity and reliability. However, for large-scale or complex nonlinear applications featuring bounded disturbances and constraints on the control and manipulate variables, a standard estimator such as EKF may not be expected to provide good approximations since the assumption of a normal distribution for the system's states may no longer hold. In the latter cases, C-PF may be expected to provide a better approximation of the AC parameters than the EKF, and therefore improve NMPC performance. In general, C-PF is the most popular approach in the literature to deal with non-Gaussian distributions that appears in the presence of bounds on the states and process uncertainties in the system. In particular, the literature often recommends to use C-PF for the case of non-symmetric bounded distribution uncertainties[50,84,126,174]. Based on the above, both EKF and C-PF are initially considered to assess the impact of AC estimation in the closed-loop framework of a large-scale system. The detailed reviews on C-PF algorithms can be found elsewhere[27,184,185,186]. The MHE and NMPC formulations considered in this work are presented next.

6.1.1 Moving Horizon Estimation (MHE)

MHE aims to find estimates for the unknown states at the current time interval k by minimizing the summation of the ℓ^2-norm of process uncertainty and measurement noises over a finite time horizon N in the presence of process constraints, e.g., bounds on the states, inputs and disturbances entering the plant. The general formulation of the MHE is as follows[29,187]:

$$\min_{\{\boldsymbol{x}_j,\boldsymbol{w}_j\}_{j=k-N}^{k-1}} \sum_{j=k-N}^{k-1} \left\|\boldsymbol{w}_j\right\|_{\boldsymbol{Q}^{-1}}^2 + \sum_{j=k-N+1}^{k} \left\|\boldsymbol{v}_j\right\|_{\boldsymbol{R}^{-1}}^2 + \varphi_{k-N}$$

s.t.

$$\begin{aligned}
&\boldsymbol{x}_{j+1} = f(\boldsymbol{x}_j, \boldsymbol{u}_j) + \boldsymbol{w}_j\,; && \forall j = k-N, \dots k-1 \\
&\boldsymbol{y}_j = h(\boldsymbol{x}_j, \boldsymbol{u}_j) + \boldsymbol{v}_j\,; && \forall j = k-N+1, \dots k \\
&g(\boldsymbol{x}_j, \boldsymbol{u}_j, \boldsymbol{w}_j, \boldsymbol{y}_j) \le 0; && \forall j = k-N, \dots k \\
&\boldsymbol{x}^{\boldsymbol{l}} \le \boldsymbol{x}_j \le \boldsymbol{x}^{\boldsymbol{u}}\,; && \forall j = k-N, \dots k
\end{aligned} \tag{6-1}$$

where:

$\boldsymbol{w} \in \mathbb{R}^{n_x}, \boldsymbol{v} \in \mathbb{R}^{n_y}, \boldsymbol{x}^{\boldsymbol{l}}, \boldsymbol{x}^{\boldsymbol{u}}, \boldsymbol{x} \in \mathbb{R}^{n_x}, \boldsymbol{y} \in \mathbb{R}^{n_y}, \boldsymbol{u} \in \mathbb{R}^{n_u}, \boldsymbol{Q} \in \mathbb{R}^{n_x \times n_x}, \boldsymbol{R} \in \mathbb{R}^{n_y \times n_y}$,

$f\colon \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_x}$, $h\colon \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_y}$, $g\colon \mathbb{R}^{n_x \times n_u \times n_x \times n_y} \to \mathbb{R}^{n_g}$

As shown in problem (6-1), k denotes the current time interval, N is the length of horizon, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are the covariance matrices of process uncertainties and measurement noises, respectively; $\boldsymbol{w}$, $\boldsymbol{v}$, $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{u}$ represent vectors that describe the process uncertainty, measurement noise, states, outputs, and inputs in the system, respectively. The index j is the time interval within the time horizon N. Note that k is the current time interval of the plant and is different from the time interval j, which denote past time intervals in the MHE framework. The function f is the process model whereas h is a model that describes the dependency of the online measurements with respect to the state variables. Moreover, g represents the inequality constraints in the system, excluding the bounds on the states and the inputs. The lower and upper bounds on the states and inputs are denoted by the vectors $\boldsymbol{x}^{\boldsymbol{l}}$, $\boldsymbol{u}^{\boldsymbol{l}}$, $\boldsymbol{x}^{\boldsymbol{u}}$, and $\boldsymbol{u}^{\boldsymbol{u}}$, respectively. The penalty term φ_{k-N} in problem (6-1) provides the AC estimated at time $k-N$. This term considers the past information of the system that was not considered in the current horizon of the MHE problem. This term is defined as follows:

$$\varphi_{k-N} = \|\boldsymbol{x}_{k-N} - \overline{\boldsymbol{x}}_{k-N}\|^2_{\boldsymbol{P}^{-1}_{k-N}} \quad (6\text{-}2)$$

where $\overline{\boldsymbol{x}}_{k-N}$ and $\boldsymbol{P}_{k-N}$ denote the expected value and covariance matrix of the approximated posterior distribution of the states at the time interval $k-N$. Both $\overline{\boldsymbol{x}}_{k-N}$ and $\boldsymbol{P}_{k-N}$ are approximated using only the measurements available at $k-N$, i.e., $\boldsymbol{y}_{k-N}$, and the past information of the system available at time $k-N-1$ (i.e., $\widehat{\boldsymbol{x}}_{k-N-1}$ and $\boldsymbol{P}_{k-N-1}$). $\widehat{\boldsymbol{x}}_{k-N-1}$ is the filtered state estimate obtained from MHE whereas $\boldsymbol{P}_{k-N-1}$is the outcome of the AC approximation method estimated at the time interval $k-1$. Note that $k-N$ is the starting time point in the finite horizon window. The estimated states obtained from the MHE problem at the current time $\widehat{\boldsymbol{x}}_k$ represent the initial condition in the NMPC formulation.

6.1.2 Nonlinear Model Predictive Control (NMPC)

NMPC aims to find the optimal control actions at the current time interval $\boldsymbol{u}_k$. As shown in problem (6-3), the optimal control actions are obtained from the minimization of the least squares errors between the predicted output variables and their corresponding set points, together with the magnitude of the changes made in the manipulated variables[188]. The general MPC framework is as follows[189]:

$$\min_{\substack{\{\boldsymbol{x'}_z\}_{z=k+1}^{k+L} \\ ,\{\boldsymbol{u}_z\}_{z=k}^{k+C-1}}} \sum_{z=k+1}^{k+L} \left(\boldsymbol{y'}_z - \boldsymbol{y}_{\boldsymbol{ref}}\right)^T \boldsymbol{Q}_{\boldsymbol{out}} \left(\boldsymbol{y'}_z - \boldsymbol{y}_{\boldsymbol{ref}}\right) + \sum_{z=k}^{k+C-1} \Delta \boldsymbol{u}_z{}^T \boldsymbol{Q}_{\boldsymbol{in}} \Delta \boldsymbol{u}_z$$

s.t.

$$\begin{aligned}
&\boldsymbol{x'}_{z+1} = f(\boldsymbol{x'}_z, \boldsymbol{u}_z) && \forall z = k, \dots k+L-1 \\
&\boldsymbol{y'}_z = h(\boldsymbol{x'}_z, \boldsymbol{u}_z) && \forall z = k, \dots k+L \\
&g(\boldsymbol{x'}_z, \boldsymbol{u}_z, \boldsymbol{y'}_z) \le 0, && \forall z = k, \dots k+L \\
&\boldsymbol{x^l} \le \boldsymbol{x'}_z \le \boldsymbol{x^u}, && \forall z = k, \dots k+L \\
&\boldsymbol{u^l} \le \boldsymbol{u}_z \le \boldsymbol{u^u}, && \forall z = k, \dots k+L \\
&\boldsymbol{u}_z = \boldsymbol{u}_{z-1}, && \forall z = k+C, \dots k+L \\
&\Delta \boldsymbol{u}_z = \boldsymbol{u}_z - \boldsymbol{u}_{z-1}, && \forall z = k, \dots k+L \\
&\boldsymbol{x'}_z = \widehat{\boldsymbol{x}}_k, && \forall z = k
\end{aligned} \quad (6\text{-}3)$$

where:

$\boldsymbol{x'} \in \mathbb{R}^{n_x}, \boldsymbol{y'} \in \mathbb{R}^{n_y}, \boldsymbol{u} \in \mathbb{R}^{n_u}, \boldsymbol{Q_{in}} \in \mathbb{R}^{n_u \times n_u}, \boldsymbol{Q_{out}} \in \mathbb{R}^{n_y \times n_y}, \boldsymbol{x^l}, \boldsymbol{x^u} \in \mathbb{R}^{n_x}, \boldsymbol{u^l}, \boldsymbol{u^p} \in \mathbb{R}^{n_u}$,
$f: \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_x}$, $h: \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_y}$, $g: \mathbb{R}^{n_x \times n_u \times n_y} \to \mathbb{R}^{n_g}$

As shown in problem (6-2), $\boldsymbol{x'}$ and $\boldsymbol{y'}$ denote the vectors of predicted states and outputs of the system at the z^{th} time interval; $\boldsymbol{Q_{out}}$ and $\boldsymbol{Q_{in}}$ are the matrices of the weights on the controlled and manipulated variables, respectively. The scalars L and C are the prediction and control horizon, respectively; $\boldsymbol{x'}_z = \widehat{\boldsymbol{x}}_k$ provides the initial condition of the NMPC problem, which describes the connection between the MHE ($\widehat{\boldsymbol{x}}_k$) and NMPC ($\boldsymbol{x'}_z$). Note that index z in Equation (6-3) is the time interval in the prediction

horizon, i.e., starting from current time interval k to the future time $k + L$; on the other hand, the index j in Equation (6-1) denotes the time interval within the estimation horizon, i.e., starting from historical time interval $k - N$ to the current time interval k. Moreover, for the case when the prediction horizon L is longer than the control horizon C, the constraint $\boldsymbol{u}_z = \boldsymbol{u}_{z-1}$ in Equation (6-3) ensures that the manipulated variables (MV) remain constant and equal to the last value in the control horizon until the end of the prediction horizon. In addition, $\boldsymbol{u}_{k-1}$ is the known control action evaluated by NMPC at the previous time interval, i.e., $k - 1$.

6.2 Computational Experiments

This section presents the results obtained for implementing the closed-loop framework shown in Figure 6-1 for each of the case studies considered in this chapter. The computational experiments were conducted using Pyomo 5.2 (and Python 3.6) on a computer running Microsoft Windows Server 2016 standard. The computer was equipped with 96 GB RAM and Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz 2.10 GHz (2 processors). The direct transcription approach was used to solve both the NMPC and MHE formulations at each time step k. This method discretizes the nonlinear differential equations into a set of nonlinear algebraic equations thus transforming the dynamic optimization problem into a large-scale nonlinear optimization problem (NLP)[190]. Both the NMPC and MHE formulations were solved using the interior-point method, which is an efficient method used to solve large-scale NLP[191]. The backward method was used to discretize the nonlinear dynamic process model. A normalized sum of squared error (SSE) of MHE has been used in this work to compare the performance of the different AC estimation methods considered in this work. The SSE for the m^{th} state using the n^{th} AC estimation method is as follows:

$$SSE_m^{(n)} = \sum_{k=0}^{t_f} \left(\hat{x}_{k,m}^{(n)} - x_{k,m} \right)^2 \tag{6-4}$$

As a reminder, k is the current time index and t_f is the final time horizon considered in the experiments; m denotes the index of state in the vector of states $\boldsymbol{x}$, i.e., $m \in \{1, 2, \dots, n_x\}$, where n_x is the number of state variables; $\hat{x}_{k,m}^{(n)}$and $x_{k,m}$ are scalars that provide the estimated and true values for the m^{th} state at each instant k, respectively; $\hat{x}_{k,m}^{(n)}$ is obtained from MHE using the n^{th} AC approximation method. The normalized SSE for state m is defined as follows:

$$\widehat{SSE}_m^{(n)} = \frac{SSE_m^{(n)}}{SSE_{m_{max}}} \; ; \quad SSE_{m_{max}} = \max SSE_m^{(n)} \tag{6-5}$$

where $SSE_{m_{max}}$ represents the largest SSE for m^{th} state, which is obtained from the different AC estimation methods considered in this work. The results of this study for WTP and HIPS are presented in sections 6.2.1 and 6.2.2, respectively. Section 6.2.3 presents an error analysis conducted in this work to analyze the results obtained from the case studies.

6.2.1 Wastewater Treatment Plant (WTP)

The WTP model presented in section 4.2.4. (see Equation (4-29)) is considered as the first case study to test the closed-loop framework presented in section 6.1. As a reminder, this plant considers four input variables (q_i, q_p, f_k, and q_2) and six state variables (x_w, s_w, x_d, x_b, x_r, and c_w). Table B-2 in Appendix B lists the nominal steady-state value of the states and model parameters for this process. The key states to control in this process are the biomass concentration inside the bioreactor (x_w), the organic substrate (s_w) and the dissolved oxygen concentration in the bioreactor (c_w). In the current study, the WTP presented in Equation (4-29) is subjected to the following process constraints:

$$s_w(t) \leq 100 \tag{6-6}$$

$$0.01 \leq \frac{q_p(t)}{q_2(t)} \leq 0.2 \tag{6-7}$$

Constraint (6-6) ensures that the water disposed to the river is not beyond the maximum allowed limit set by the municipality whereas constraint (6-7) ensures that the recycle/purge ratio is maintained within acceptable limits to maintain a profitable operation.

In general, WTP require considerable amounts of energy to maintain appropriate recycle ratios and supply sufficient oxygen to the bioreactor. Also, WTP are often subject to external perturbations (e.g., changes in the influent) that will impact the operability of the plant (e.g., the amount of pollutants disposed to natural effluents) and therefore the process economics[172,192]. Moreover, it has been widely recognized that key operating variables in WTP cannot be accurately measured online[180] for various reasons (e.g., significant measurement error) thereby adding an additional layer of complexity to the operational tasks required for this process. In addition, WTPs are difficult to obtain[192] and more likely may not completely capture the dynamic operation of this process; hence, WTPs are subject to process uncertainties (i.e., plant-model mismatch) that will eventually diminish the performance of these plants in closed-loop. To the author's knowledge, optimal model-based control and state estimation strategies

that consider bounded process uncertainty like that presented in this work has not been reported for this process, even though they can be effective to maintain the feasible operation of this plant within acceptable economic targets.

The goal of this case study is to investigate the impact of some of the most widely known AC approximation methods, i.e., EKF and C-PF, during closed-loop operation of an industrial-scale system such as the WTP. As mentioned in chapter 4, this case study considers that the online measurements available are the concentration of biomass at upper layer of decanter (x_d), substrate (s_w) and oxygen (c_w). The controlled variables (CV) considered in the NMPC are x_w and s_w and c_w whereas the manipulated variables (MV) are q_p, f_k, and q_2. The rest of the assumption are presented in Appendix C.

All state estimators rely on both plant measurements and the plant model; out of these two sets of information, the focus of the estimator will be to follow more reliable sources of information. The larger the covariance of a distribution with respect to other distributions, the lower the reliability of that particular distribution. Hence, this study considers that the process measurements are assumed to follow a normal distribution with a standard deviation (σ_{meas}) set to 10% of the nominal steady-state values of the states (see Table B.2). Moreover, σ_{unc} denotes the standard deviation of the process uncertainty associated with each of the plant states, which has been set to 5% of the nominal steady-state values of the states (see Table B-2 in Appendix B). Under this set of assumptions, MHE and AC estimators tend to rely more on the plant model rather than the measurements; hence, the effect of bounded process uncertainties are expected to have a more significant impact on the state estimation. Note that estimates for σ_{meas} and σ_{unc} can be obtained from a real setting using process heuristics or historical plant data.

Multiple scenarios under different conditions have been considered to perform this analysis. Table 6-1 summarizes the scenarios considered in this work. *Scenario I* and *Scenario II* explore the effects of different bounded distribution on the process uncertainties. As shown in Figure 6-2, both symmetric and non-symmetric bounded distributions are considered in the plant model for the individual process uncertainties associated with each of the WTP's state variables. *Scenario I* assumes a symmetric bounded distribution in the range of $(-2\sigma_{unc}, 2\,\sigma_{unc})$ whereas *Scenario II* considers a non-symmetric bounded distribution in the range of $(-2\sigma_{unc}, 0.6\sigma_{unc})$. *Scenario III* investigates the role of AC estimation using different plant designs. The response time of the system is highly related to the plant's design. For instance, larger vessel sizes may lead to higher settling times. Hence, the response time is a key parameter to adjust both the estimation (N) and prediction (L) horizons in a closed-loop

framework. As the time constant of the process increases, longer estimation and prediction horizons in the MHE and NMPC frameworks may be required to perform a better estimation and control for this type of processes, which would also affect the online computational costs. For instance, if MHE does not consider a long enough estimation horizon for this processes with large time constants, the true state dynamics may be interpreted as measurement noise and process uncertainties due to the slow changes in the system, e.g., a process that is slowly drifting away from a nominal operating point. Hence, this condition may eventually lead to an inaccurate estimation of the states due the inability of the MHE to differentiate between process/measurement noise and actual process dynamics. A longer estimation horizon provides MHE with a larger set of measurements and process constraints that helps MHE to accurately capture the evolution of the system and therefore provide accurate state estimations at the expense of higher computational costs. Alternatively, the AC approximation becomes more important for those cases since it directly affects the MHE performance in the closed-loop framework. Consequently, an accurate initialization of NMPC problem becomes more crucial as the response time of the process plant increases. To analyze this effect, *Scenario III* considers four instances featuring two different plant designs, each under bounded symmetric and non-symmetric distributions in the process uncertainty, respectively. As shown in Table 6-1, *Scenario III.A* and *Scenario III.B* assume a -20% and +20% change in the capacity of the WTP's decanter under a symmetric process uncertainty, respectively, whereas *Scenarios III.C* and *Scenario III.D* perform the same test under non-symmetric distributions for process uncertainties. The process uncertainty considered for *Scenario III*'s instances are the same used for *Scenarios I-II* (see Figure 6-2). Note that a test involving open-loop estimation under these scenarios was performed and is presented in Appendix C.

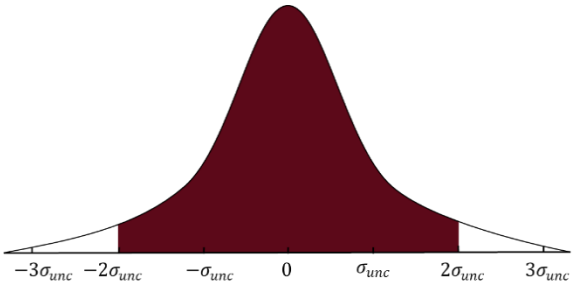
For each scenario, the closed-loop scheme was performed using six different approximations of the mean and covariance of the AC distribution $\mathcal{N}(\bar{\boldsymbol{x}}_{k-N}, \boldsymbol{P}_{k-N})$ featuring true states (TS), EKF and C-PF, i.e., $n \in \{\mathrm{TS}, 0.5\mathrm{TS}, \mathrm{EKF}, \mathrm{C-PF}, \mathrm{EKFexpc}, \mathrm{C-PFexpc}\}$. Table 6-2 shows the detailed information of each of these instances. In the case that $n = \{EKF\}$ or $n = \{C - PF\}$, the posterior distribution provided by their corresponding estimator at time $k - N$ has been used explicitly as the approximation of both AC parameters ($\bar{\boldsymbol{x}}_{k-N}$ and $\boldsymbol{P}_{k-N}$). Note that $\boldsymbol{P}_{k-N}$ is estimated using $\boldsymbol{P}_{k-N-1}$ obtained from the corresponding AC estimation method at the time interval $k - 1$ of the closed-loop framework. For the rest of the cases, i.e., $n \in \{\mathrm{TS}, 0.5\mathrm{TS}, \mathrm{EKFexpc}, \mathrm{C-PFexpc}\}$, a matrix $1 \times 10^{-10}\mathbf{I}$, where $\mathbf{I}$ is an identity matrix of proper dimensions, has been considered as the covariance matrix of AC distribution ($\boldsymbol{P}_{k-N}$). Statistically, a small covariance matrix is an indication of a narrow normal distribution; hence,

the expected values obtained from this distribution have a high probability to be the true states. Thus, based on Equations (6-1)-(6-2), a smaller $\boldsymbol{P}_{k-N}$ leads to a higher weight on AC in the objective function of the MHE framework, which is expected to drive the estimated states at time step $k-N$ ($\widehat{\boldsymbol{x}}_{k-N}$) to be closer to the expected state of the AC ($\overline{\boldsymbol{x}}_{k-N}$). In the case of EKFexpc (C − PFexpc), the expected value estimated by EKF (C-PF) has been assigned to $\overline{\boldsymbol{x}}_{k-N}$. When $n = \{\text{TS}\}$, the true states have been explicitly used as the mean value of AC distribution $\overline{\boldsymbol{x}}_{k-N}$; whereas in the case of $n = \{0.5\text{TS}\}$, a -50% of true state values is used as the expected states $\overline{\boldsymbol{x}}_{k-N}$. Both TS and 0.5TS have been used in this work as benchmark AC estimation methods, i.e., TS represents the ideal case (full access to the true states) whereas 0.5TS presents a case of a poor AC estimation of the states $\overline{\boldsymbol{x}}_{k-N}$. Each of the s*cenarios I-III* was repeated six times based on the n AC approximation methods considered in this work. The results reported in this study are focused on the estimation of biomass concentration (x_w) for WTP since this is the most critical state for monitoring and controlling this process.

Table 6-1: Characteristics of Scenarios considered in this study

Id	Plant Design			Bounded distribution of process uncertainties	
	Fixed	-20%	+20%	Symmetric	Non-symmetric
Scenario I	✓	×	×	✓	×
Scenario II	✓	×	×	×	✓
Scenario III.A	×	✓	×	✓	×
Scenario III.B	×	×	✓	✓	×
Scenario III.C	×	✓	×	×	✓
Scenario III.D	×	×	✓	×	✓

(a)



(b)

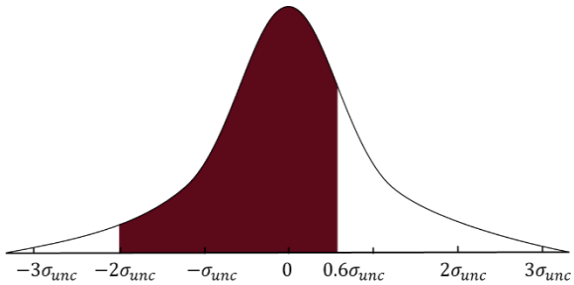


Figure 6-2: (a) Bounded symmetric distribution for process uncertainty: $(-2\sigma_{unc}, 2\ \sigma_{unc})$; (b) Bounded non-symmetric distribution for process uncertainty: $(-2\sigma_{unc}, 0.6\ \sigma_{unc})$

Table 6-2: Characteristic of AC estimation methods considered in this study

AC estimation method (n)	Approximation of AC distribution
EKF	$\mathcal{N}\left(\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{EKF}}, \boldsymbol{P}_{k-N}^{\boldsymbol{EKF}}\right)$
C-PF	$\mathcal{N}\left(\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{C-PF}}, \boldsymbol{P}_{k-N}^{\boldsymbol{C-PF}}\right)$
EKFexpc	$\mathcal{N}\left(\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{EKF}}, 1\times10^{-10}\mathbf{I}\right)$
C-PFexpc	$\mathcal{N}\left(\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{C-PF}}, 1\times10^{-10}\mathbf{I}\right)$
TS	$\mathcal{N}\left(\boldsymbol{x}_{k-N}^{\boldsymbol{TS}}, 1\times10^{-10}\mathbf{I}\right)$
0.5TS	$\mathcal{N}\left(\boldsymbol{x}_{k-N}^{\boldsymbol{0.5TS}}, 1\times10^{-10}\mathbf{I}\right)$

6.2.1.1 WTP: Closed-loop operation: Scenarios I-III

Figure 6-3 presents the performance of the NMPC, MHE and AC estimators during closed-loop operation for *Scenarios I* and *II* (see Table 6-1). As shown in this Figure, both scenarios showed poor performance in the absence of an acceptable AC estimation method in the MHE formulation (i.e., 0.5TS in Figure 6-3). For instance, in the case of 0.5TS as the AC approximation method, a 96% difference between $\widehat{SSE}_{x_w}^{(0.5TS)}$ and $\widehat{SSE}_{x_w}^{(TS)}$ was observed for *Scenario II*. On the other hand, the biomass concentration is close to its target values when a suitable EKF or C-PF is used to approximate the AC parameters. However, as shown in Figures 6-3(a) and 6-3(d), changing the AC approximation method did not seem to have a significant effect in closed-loop for *Scenarios I-II*, i.e., the difference between $\widehat{SSE}_{x_w}^{(EKF)}$ and $\widehat{SSE}_{x_w}^{(C-PF)}$ is less than 0.3% in *Scenario II.* This insensitivity of the control system to the AC approximation method can be explained by inspecting Figures 6-3(b) and 6-3(e), which illustrate the estimation of the biomass (obtained from MHE) in closed-loop for *Scenario I* and *II*, respectively. These figures show that, regardless of the type of bounded distributions, the MHE performance remains almost the same when C-PF or EKF is used as the AC estimation method. According to Figures 6-3(c) and 6-3(f), the AC estimation provided by EKF is slightly better than that obtained from C-PF (~0.2%); however, this can be attributed to either noises in the system (e.g., measurement noise) or round-off errors. Based on the above, both EKF and C-PF are able to provide

an accurate enough AC approximation. As the result, both MHE and NMPC show a similar performance using either EKF or C-PF as the AC approximation method.

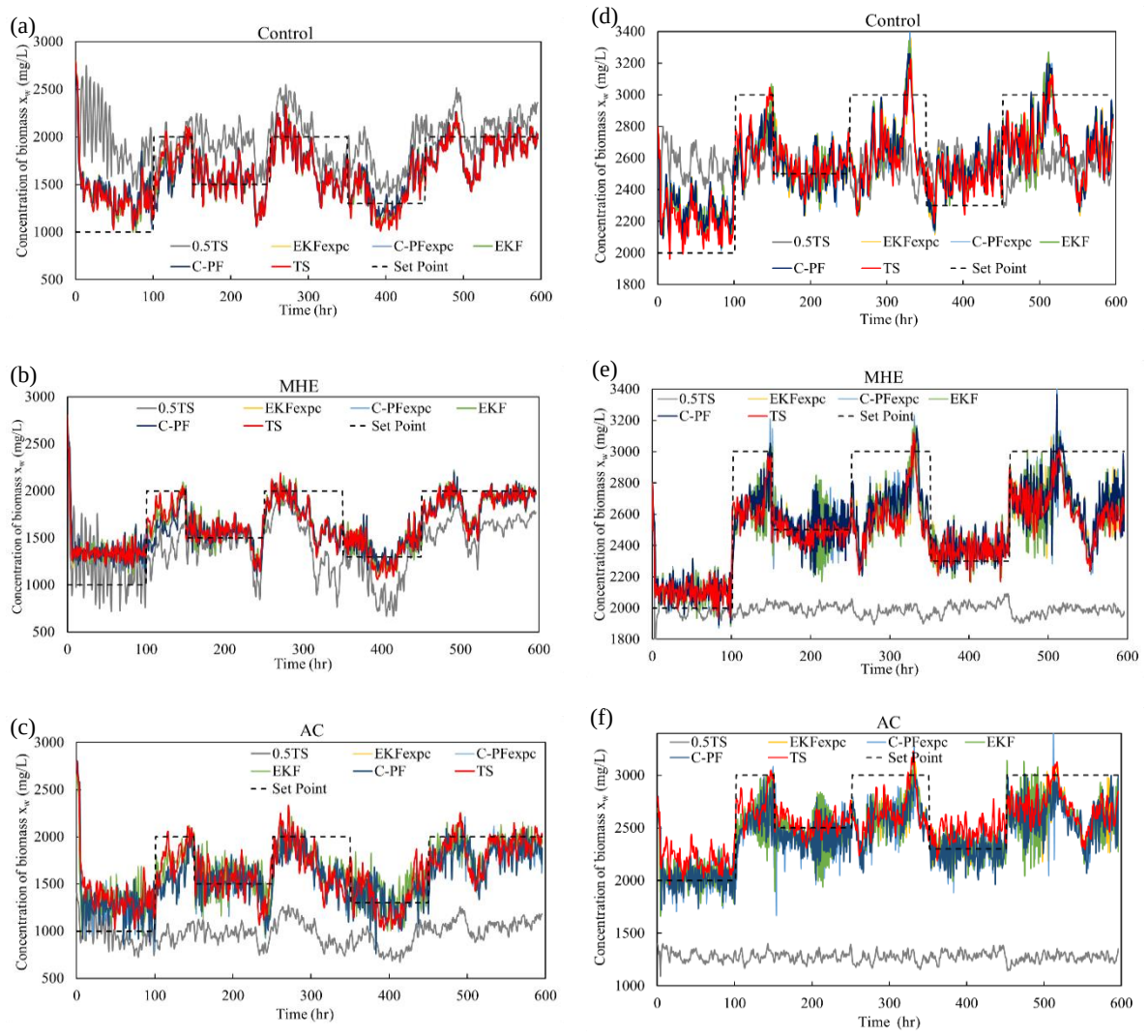


Figure 6-3: Impact of using different AC estimators on the control and estimation (MHE) of biomass in the closed-loop system (a), (b), (c) *Scenario I*; (d), (e), (f) *Scenario II*

Figure 6-4 shows the control actions implemented on the recycle flowrate q_2 under the different AC estimation methods. As shown in this figure, the control actions for q_2 are very similar when EKF, C-PF and TS are used as the AC estimation methods; hence, similar control actions are expected from these methods. Note that the control actions reported for 0.5TS are far from the benchmark, i.e., the true state estimation method (TS), which highlights the importance of using a good AC approximation method in the MHE-NMPC closed-loop framework. Moreover, the results shown in Figure 6-4 indicate

that the MVs are moving aggressively, particularly under *Scenario I*. In practice, a low-order filter can be used to filter out the noisy signal before passing the control actions to the process plant.

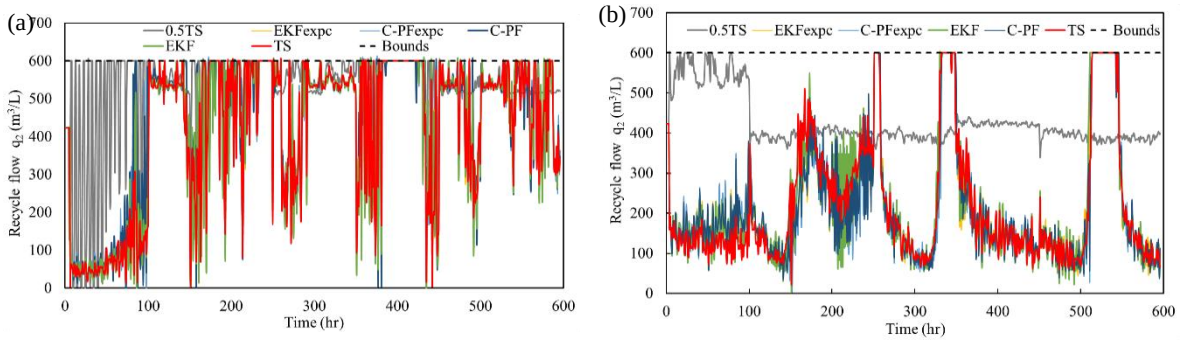


Figure 6-4: Original plant design: Recycle flow that leaves the decanter $q_2\ (m^3/L)$ (a) *scenario I*; (b) *scenario II*

Figure 6-5 highlights the results of closed-loop operation in the presence of symmetric bounded uncertainties for different plant designs (*Scenarios III.A* and *III.B*). Similar to the results obtained for *Scenario I* and *II* (closed-loop operation), the results obtained for *Scenarios III.A* and *III.B* confirm the importance of using an adequate AC estimation method in the closed-loop framework. As shown in Figures 6-5(a) and 6-5(d), NMPC failed to take a proper action when 0.5TS was used as the AC estimation method. Note that $\widehat{SSE}_{x_w}^{(0.5TS)}$ is higher than $\widehat{SSE}_{x_w}^{(TS)}$ by 44% and 83% for *Scenario III.A* and *Scenario III.B*, respectively. Similar to the open-loop scenarios, the closed-loop results suggest that a more reliable AC approximation method becomes critical when the size of the plant (and the response time) increases. Nevertheless, a somewhat similar $\widehat{SSE}_{x_w}^{(EKF)}$ and $\widehat{SSE}_{x_w}^{(C-PF)}$ (0.17) was recorded for *Scenario III.A* and *Scenario III.B*, hence, a similar closed-loop performance was observed from both AC approximation methods, as shown in Figures 6-5(a) and 6-5(d). This result was expected since the MHE performance for both instances remained the same, as shown in Figures 6-5(b) and 6-5(e). As depicted in Figures 6-5(c) and 6-5(f), C-PF did not improve the accuracy to the approximated distribution of AC.

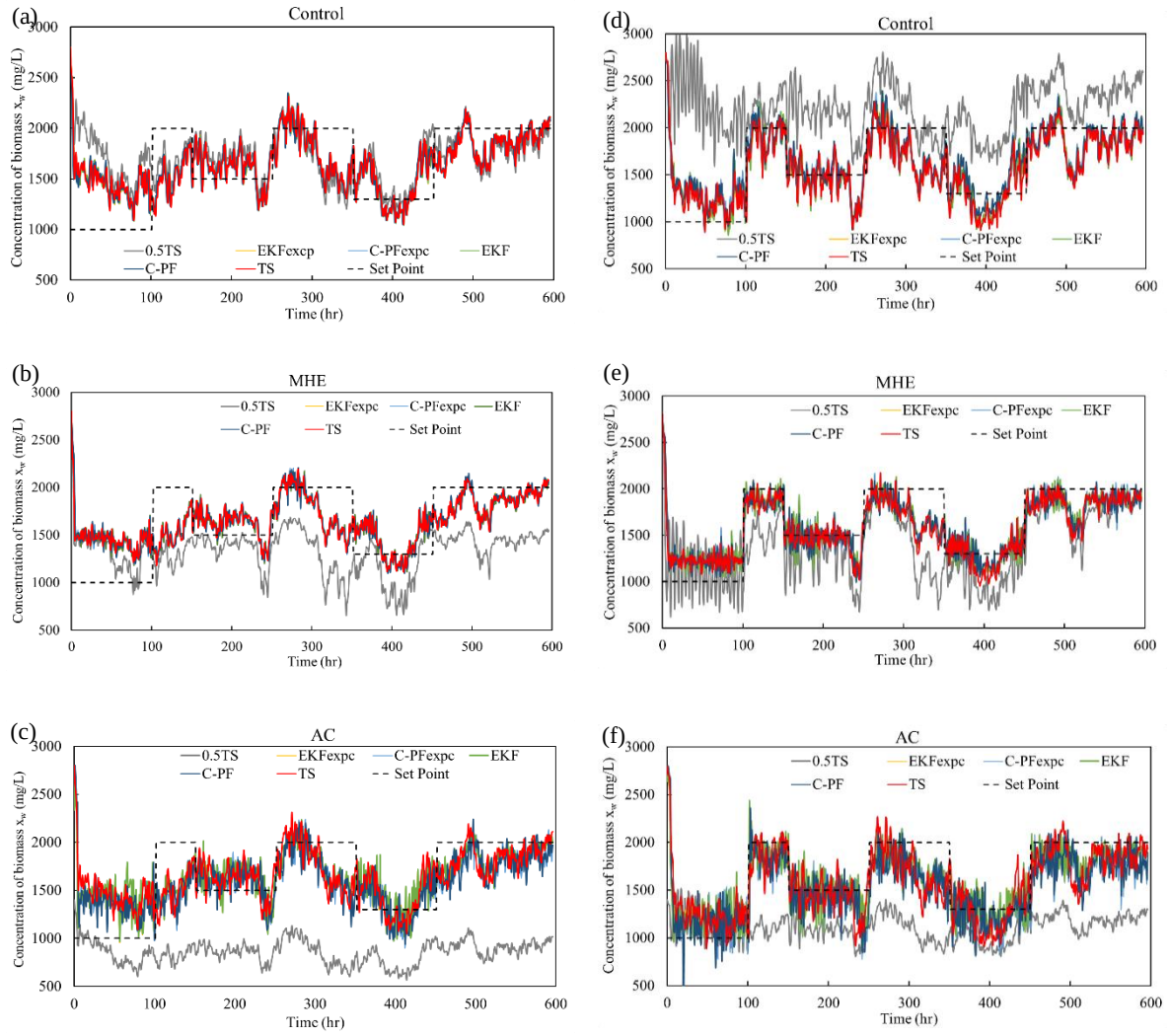


Figure 6-5: Impact of using different AC estimators on the control and estimation (MHE) of biomass in the closed-loop system (a), (b), (c) *Scenario III.A*; (d), (e), (f) *Scenario III.B*

Figure 6-6 shows the results for *Scenarios III.C* and *III.D* (see Table 6-1). As in *Scenarios III.A* and *III.B*. The $\widehat{SSE}_{x_w}^{(0.5TS)}$ obtained for *Scenario III.C* is 57% smaller than that reported for *Scenario III.D*; these suggests that larger equipment sizes may require more accurate AC estimation methods during closed-loop operation. Nevertheless, the results presented in Figure 6-6 show a slight improvement when C-PF is used as the AC estimation method. For instance, $\widehat{SSE}_{x_w}^{(C-PF)}$ is 1% and 19% smaller than the $\widehat{SSE}_{x_w}^{(EKF)}$ obtained for *Scenario III.C* and *III.D*, respectively. This suggests that, as the size of the plant increases, the closed-loop becomes more sensitive to the AC method and C-PF can handle

non-symmetric bounded uncertainties better compared to EKF. Although using C-PF in the AC estimation improves the overall performance of the control-loop system, this estimation method required an additional 12% and 17% in the averaged CPU costs for *Scenarios III.C* and *III.D*, respectively. On the other hand, EKF is able to provide reasonable estimates for the AC estimation in shorter CPU times, which is essential to ensure an economically feasible and safe operation of the WTP in closed-loop.

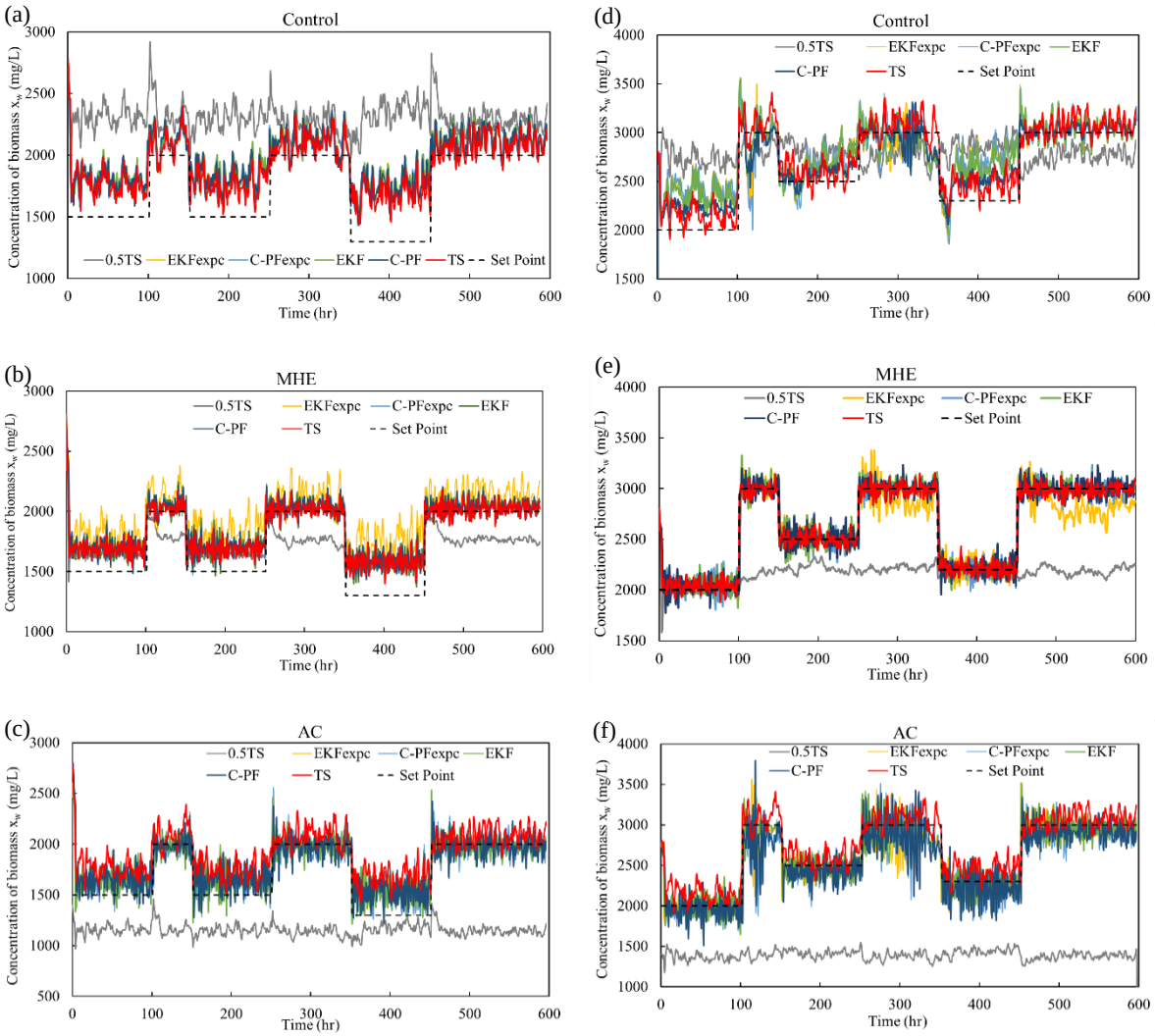


Figure 6-6: Impact of using different AC estimators on the control and estimation (MHE) of biomass in the closed-loop system (a), (b), (c) scenario III.C; (d), (e), (f) scenario III.D

The scenarios discussed above assume that the NMPC does not have access to the actual measurements. This is essential for the correct initialization of the NMPC framework. Thus, the control system relies

on the state estimator (i.e., MHE). As the result, there is no a direct connection between the plant and the NMPC, whereas there is an explicit connection between the MHE and plant (see Figure 6-1). Hence, the offset between the set point and the biomass estimation provided by MHE is smaller than the offset between the set point and the plant output for biomass. This issue can be seen in the scenarios presented above for the closed-loop operation. This problem can be resolved by designing a robust NMPC that takes process uncertainty into account in the NMPC formulation. In addition, adding integrating states to WTP may reduce this offset by inducing integral action in the NMPC. This may be achieved by considering nonstationary disturbances as augmented states to the WTP[55,56]. However, this is beyond the scope of this work. Note that the offset between estimation and set points is smaller compared to the offset between CVs and the set points.

6.2.1.2 WTP: EKF for state estimation vs EKF for AC estimation

The results presented above have shown that EKF can approximate the AC parameters even in the presence of bounded constraints, as shown in *Scenario I* and *Scenario II* above. Although EKF is expected to face difficulties when the Gaussian assumption approximation cannot hold, this is still an acceptable AC estimator as it can provide an accurate approximation of AC in short computational times. The closed-loop considered in section 6.2.1.1 follows the standard framework shown in Figure 6-1. This section considers two additional closed-loop strategies (*CL*) that are used here to provide insights regarding the efficiency of EKF as an AC estimator in the MHE-NMPC closed-loop framework involving nonlinear constrains such as bounds on the process uncertainties. To pursue this goal, two closed-loop strategies (*CL*) involving different features have been considered; these strategies are as follows:

CL1) As shown in Figure 6-7(a), this closed-loop strategy considers EKF as the only state estimation technique (i.e., without the MHE framework).

CL2) EKF is used as the AC estimator while EKF is working independent of MHE, i.e., the prior estimation of AC is provided by EKF itself and not by MHE framework, as shown in Figure 6-7(b). These two tests, i.e., *CL1* and *CL2*, have been compared to the base case control strategy (*BC*) used in this work described in section 6.1 and depicted in Figure 6-1. In the base case (*BC*) framework, EKF is used to estimate AC presented in MHE within a standard framework, where the prior estimation of the states in EKF is provided by MHE.

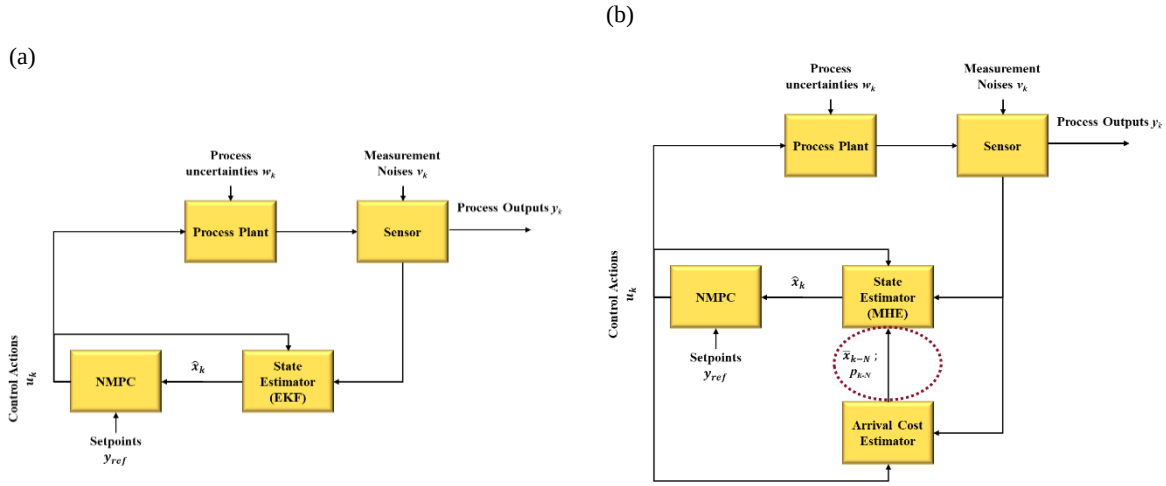


Figure 6-7: Block diagrams: (a) CL1, (b) CL2

The closed-loop frameworks described for *CL1* and *CL2* have been tested using *Scenario I* and *Scenario II* described above. As shown in Figures 6-8(a)-(b), *CL1* test does not lead to instability in the control or estimation in the presence of symmetric bounded process uncertainties (*Scenario I*). This result was expected as the constraint imposed on the probability distribution of the process uncertainties is symmetric and fairly wide, i.e., it closely approximates to a Gaussian distribution. Although the distributions of process uncertainties are not precisely Gaussian (see Figure 6-2(a)), EKF will not severely suffer from the absence of this bounds as the distributions of the process uncertainties are not far from a Gaussian distribution. Nevertheless, using MHE improves the accuracy of the state estimation and closed-loop performance as MHE considers the bounds and operational constraints for the WTP. As shown in Figure 6-8(b), the relative error in the AC approximation is reduced by approximately 30% in the base case (*BC*) closed-loop strategy compared to the case when EKF is used as an independent AC estimator (*CL2*). Figures 6-8(c)-(d) present a more realistic scenario in which process uncertainties may follow a non-symmetric probability distribution function. As shown in these figures, *CL1* (which explicitly relies on EKF estimations) may lead to instability in the presence of non-symmetric process uncertainties (*Scenario II*). As shown in Figure 6-8(d), the estimation provided by EKF in *CL1* becomes unstable as the estimation error is propagated over time, i.e., at time 24 hr (24th iteration) EKF failed to provide an estimation since the numerical solver diverged due to the large estimation error. An error analysis is presented in section 6.2.3 to further investigate the cause of this divergence. As shown in Figure 6-8(c), these unstable estimations caused by EKF in *CL1* eventually resulted in an unstable closed-loop operation. Moreover, the approximation of AC using only EKF

estimations (*CL2*) also becomes unstable (see Figure 6-8(d)) and eventually leads to closed-loop instability due to an inadequate initialization of NMPC, i.e., poor estimation of the states, as shown in Figure 6-8(c). On the other hand, the base case (*BC*) control strategy offers a stable and accurate estimation of both AC and states in closed-loop, which is consistent with the results presented in section 6.2.1.1 (*Scenarios I-III*). Based on the above, the estimation error caused by the absence of operational constraints and bounds in the EKF algorithm may propagate over time and may lead to inaccurate unstable estimations depending on the nonlinearity of the distributions considered for process uncertainty and measurements. However, the current results show that this estimation error for the EKF under nonlinear distributions may not propagate significantly if engaged with MHE since this framework is able to provide the EKF with a proper initialization at every time interval.

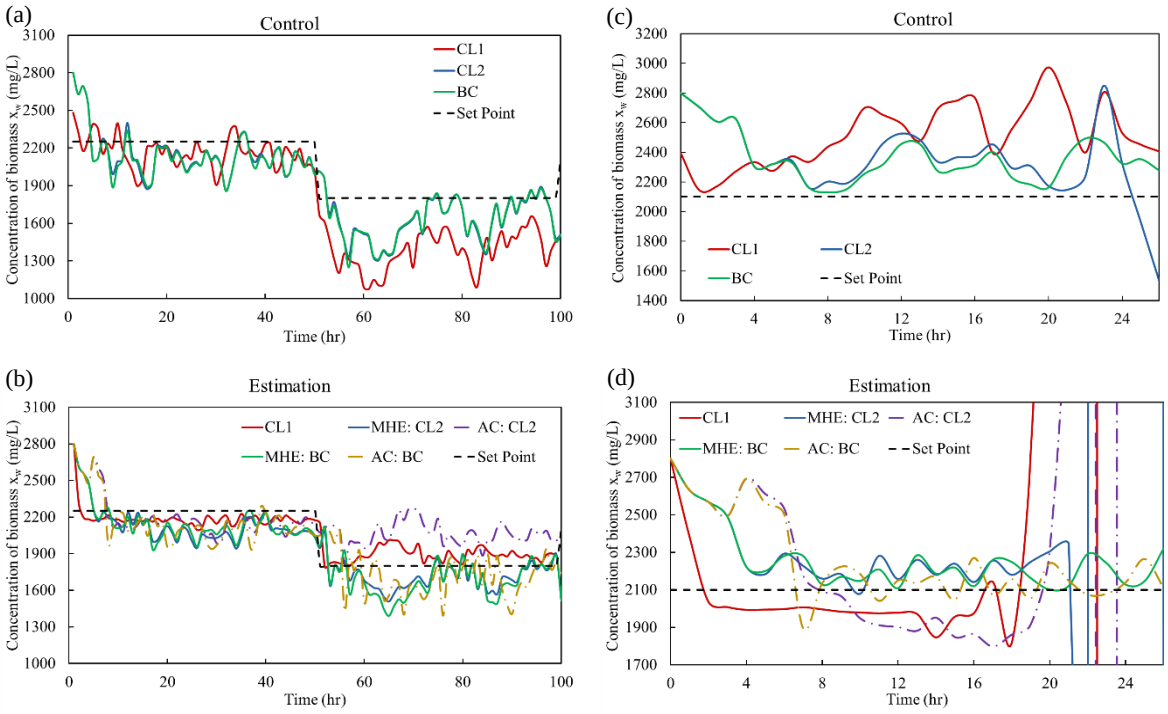


Figure 6-8: Control and estimation of biomass in the closed-loop system based on the role of EKF in the framework under (a), (b) *Scenario I*; (c), (d) *Scenario II*

6.2.2 High Impact Polystyrene Process (HIPS)

Based on the outcomes obtained from the previous case study, EKF provided an acceptable approximation of AC for an industrial-scale system even in the presence of non-symmetric bounded

process uncertainties. Hence, this section further investigates the potential of EKF as the AC estimator on a challenging control and estimation problem such as an industrial HIPS system.

In this system, the non-isothermal polymerization process takes place in a CSTR with highly sensitive nonlinear reaction kinetics. The key states to control are the monomer concentration and the reactor temperature to maintain stability and obtain different grades of polystyrene conversion based on the product demands. Equation (6-8) represents the dynamic model of the HIPS, including the rate of initiator, monomer, butadiene, radicals, branched radicals concentrations (i.e., C_i, C_m, C_b, C_r, and C_{br}), reactor temperature (T), jacket temperature (T_j), Zeroth moment dead polymer (μ_r^0), zeroth moment of the chain-length distribution for growing polymer chains with a terminal butadiene unit (μ_b^0). Equations (6-9)-(6-10) show the constraints considered for this process, i.e., bounds on the monomer concentration and reactor temperature, respectively.

$$\frac{dC_i}{dt} = \frac{Q_i C_i^f - Q C_i}{V} - k_d C_i \tag{6-8}$$

$$\frac{dC_m}{dt} = \frac{Q(C_m^f - C_m)}{V} - K_P C_m\left(\mu_r^0 + \mu_b^0\right)$$

$$\frac{dC_b}{dt} = \frac{Q(C_b^f - C_b)}{V} - C_b\left(K_{i2} C_r + K_{fs}\mu_r^0 + K_{fb}\mu_b^0\right)$$

$$\frac{dC_r}{dt} = 2e_f K_d C_i - C_r(K_{i1} C_m + K_{i2} C_b)$$

$$\frac{dC_{br}}{dt} = C_b\left[K_{i2} C_r + K_{fb}\left(\mu_r^0 + \mu_b^0\right)\right] - C_{br}\left[K_{i3} C_m + K_t\left(\mu_r^0 + \mu_b^0 + C_{br}\right)\right]$$

$$\frac{dT}{dt} = \frac{Q(T^f - T)}{V} + \frac{\Delta H_r K_P C_m\left(\mu_r^0 + \mu_b^0\right)}{\rho_s C_{ps}} - \frac{U A_H (T - T_j)}{\rho_s C_{ps} V}$$

$$\frac{dT_j}{dt} = \frac{Q_{cw}(T_j^f - T_j)}{V_c} + \frac{UA(T - T_j)}{\rho_{ps_w} C_{pcw} V_c}$$

$$\frac{d\mu_b^0}{dt} = K_{i3} C_{br} C_m - \left[K_P C_m + K_t\left(\mu_r^0 + \mu_b^0 + C_{br}\right) + K_{fs} C_m + K_{fb} C_b\right]\mu_b^0 + K_P C_m \mu_b^0$$

$$\frac{d\mu_r^0}{dt} = 2K_{i0}{C_m}^3 + K_{i1}C_rC_m + C_mK_{fs}\left(\mu_r^0 + \mu_b^0\right) - \left[K_PC_m + K_t\left(\mu_r^0 + \mu_b^0 + C_{br}\right) + K_{fs}C_m + K_{fb}C_b\right]\mu_r^0 + K_PC_m\mu_r^0$$

$$5.0 \leq C_m(t) \leq 7.5 \text{ (mol/L)} \quad (6\text{-}9)$$

$$330 \leq T(t) \leq 240 \text{ (K)} \quad (6\text{-}10)$$

Tables D-1 in Appendix D lists the nominal state values and model parameter values adopted for this case study. The kinetic parameters can be found elsewhere[181],[183]. Note that the scaled mechanistic model is used in this study as the state variables differ by orders of magnitude. HIPS includes 3 input variables: the initiator, outlet and the cooling water flow rates (i.e., Q_i, Q, and Q_{cw}) and 9 states variables (i.e., C_i, C_m, C_b, C_r, C_{br}, T, T_j, μ_r^0, and μ_b^0). The goal of this case study is to show the competence of EKF as the AC estimator for challenging cases such as the HIPS process.

Previous studies have shown that this process is challenging to control since it exhibits a complex dynamic behaviour[181,182,193]. For instance, HIPS is an open-loop unstable system that involves multiple state variables that evolve at different time scales thus making this process quite challenging. In addition to the well-known challenges in terms of controlling an unstable system, developing a control strategy that considers an online state estimation scheme makes the closed-loop framework more challenging, e.g., in the case of unstable systems, a reliable estimation of the states becomes critical to achieve stability and good NMPC performance. That is, small deviations in the initial state estimation provided to the NMPC can lead to inadequate control actions that can eventually lead to errors in the state estimation and in due course to a loss in closed-loop performance. Thus, both MHE and NMPC schemes rely on the accuracy of the AC estimation. As indicated above, a closed-loop framework involving NMPC and MHE for this challenging process has not been reported.

As in the case of the WTP, process uncertainties have been explicitly considered in the plant model and MHE to account for plant-model mismatch thus making the present closed-loop HIPS implementation more realistic. The performance of closed-loop operation of the HIPS was investigated under *Scenario II*, i.e., fixed plant design and non-symmetric bounded distribution for the process uncertainties. This study assumes that the online measurements are available for the concentration of butadiene (C_b), radicals (C_r), branched radicals (C_{br}), reactor temperature (T), and zeroth moment of the chain-length distribution for growing polymer chains with a terminal butadiene unit (μ_b^0). Although online access to some of the key states such as C_r and μ_b^0 may not be available in practice, these assumptions have been

made to simplify the analysis. Future works involve the development of new approaches that can remove this requirement. Moreover, the process measurements in this study follow a normal distribution with a standard deviation (σ_{meas}) set to 5% of the nominal scaled steady-state values of the states (see Table D-1). Moreover, σ_{unc} is set to 0.01% of the nominal scaled values of the states and is assumed to follow non-symmetric bounded distribution, i.e., *Scenario II*. Note that NMPC considers two CVs (i.e., C_m and T) and two MVs (i.e., Q, and Q_{cw}). Moreover, similar to the WTP, EKF as the AC estimator evaluates the posterior estimation of the states at time $k-N$, i.e., $\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{EKF}}$, based on the estimated states calculated from MHE at the time step *k-N*-1 ($\hat{\boldsymbol{x}}_{k-N-1}$). Then, the posterior estimation $\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{EKF}}$ is assigned as the approximation of the mean value in AC distribution ($\overline{\boldsymbol{x}}_{k-N}$). The covariance matrix for AC has been set to 1% of the nominal values of the scaled steady-state values of the states presented in Table D-1 of the appendix. The remaining underlying assumptions to perform this study are presented in Appendix D.

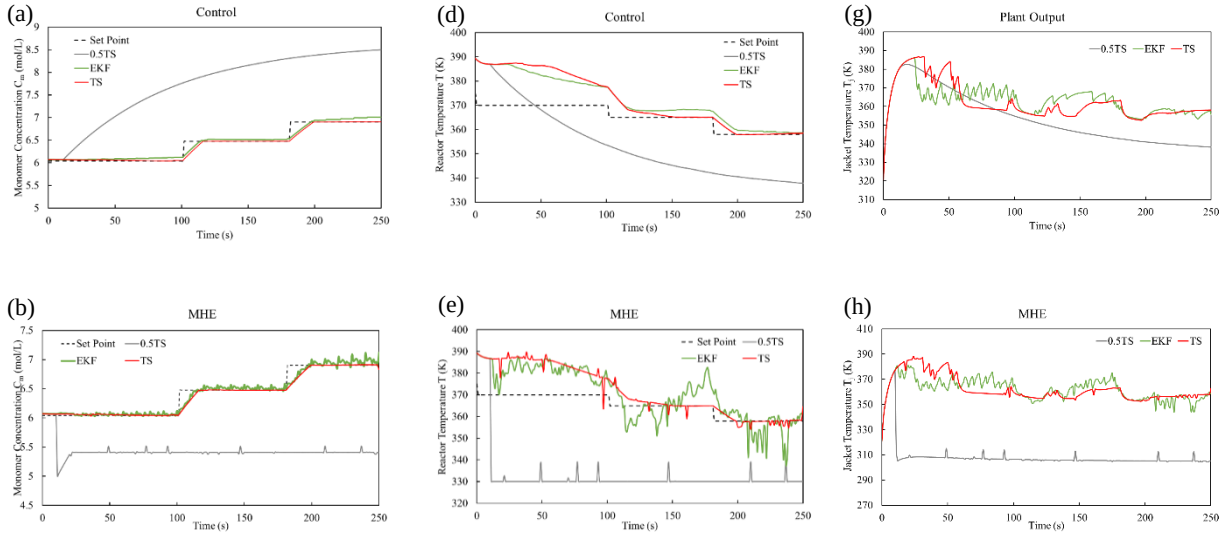
Three different approximations of AC have been considered featuring true states (TS) and EKF, i.e., $n \in \{\text{TS}, 0.5\text{TS}, \text{EKF}\}$. Table 6-3 provides the AC distribution for each of these instances. As shown in Table 6-3, in the case of $n = \{\text{EKF}\}$, the posterior estimation of the states provided by EKF was used as the approximation of the mean value ($\overline{\boldsymbol{x}}_{k-N}$) in the AC distribution at time $k-N$. Moreover, the AC covariance matrix, i.e., $\boldsymbol{P}_{k-N}$, is obtained from the scaled values of the states (denoted as $\boldsymbol{SCV} \in \mathbb{R}^{n_x}$ in Table 6-3) at the nominal operating point. Similar to WTP, $n = \{\text{TS}\}$ and $n = \{0.5\text{TS}\}$ represent the benchmark AC estimation methods, i.e., these two instances use explicitly the true value of the state and a -50% of the true value of the states as the approximation of the mean value of AC distribution $\overline{\boldsymbol{x}}_{k-N}$, respectively.

Table 6-3: Characteristic of AC estimation methods, HIPS

AC estimation method (*n*)	Approximation of AC distribution
EKF	$\mathcal{N}\left(\hat{\boldsymbol{x}}_{k-N}^{\boldsymbol{EKF}}, \boldsymbol{SCV} \times 10^{-2}\mathbf{I}\right)$
TS	$\mathcal{N}\left(\boldsymbol{x}_{k-N}^{\boldsymbol{TS}}, \boldsymbol{SCV} \times 10^{-2}\mathbf{I}\right)$
0.5TS	$\mathcal{N}\left(\boldsymbol{x}_{k-N}^{\boldsymbol{0.5TS}}, \boldsymbol{SCV} \times 10^{-2}\mathbf{I}\right)$

The results shown in this section are for both C_m and T as they are the most critical states for the purpose of monitoring and controlling the HIPS process[182]. In addition, the plant output and the estimation results for the jacket temperature (T_j) are shown to represent the performance of closed-loop for an unknown state that is not part of the set of controlled variables (CV). Figure 6-9 depicts the outcomes

for this case study. Similar to the results drawn from the WTP case study, the poor control and estimation provided by closed-loop in the case of 0.5TS highlights the need to provide a proper approximation of AC to MHE. On average, the $\widehat{SSE}_{C_m}^{(TS)}$, $\widehat{SSE}_{T}^{(TS)}$, $\widehat{SSE}_{T_j}^{(TS)}$ decreased by 99% compared to $\widehat{SSE}_{C_m}^{(0.5TS)}$, $\widehat{SSE}_{T}^{(0.5TS)}$, $\widehat{SSE}_{T_j}^{(0.5TS)}$ for C_m and T and T_j, respectively. Moreover, Figures 6-9(a) and 6-9(d) depict an adequate control in the case of using EKF as the AC estimator for both C_m and T. As shown in Figures 6-9(b), 6-9(e), and 6-9(h), MHE provided adequate estimations of C_m, T and T_j where EKF was employed as the AC approximation method. Note that C_m and T_j are unknown states, whereas online measurements are assumed to be available for T. According to these results, the MHE state estimation can provide acceptable estimations during transitions between the product grades and at each of the corresponding product grades. The results show that $\widehat{SSE}_{T_j}^{(EKF)}$ and $\widehat{SSE}_{T}^{(EKF)}$ are only 2% higher than those observed for $\widehat{SSE}_{T_j}^{(TS)}$ and $\widehat{SSE}_{T}^{(TS)}$, respectively. Note that in the case of monomer concentration, $\widehat{SSE}_{C_m}^{(EKF)}$ is only 0.2% larger than $\widehat{SSE}_{C_m}^{(TS)}$. Moreover, the good performance of EKF as the AC estimator can be observed from Figure 6-9(c), 6-9(f), and 6-9(i) when compared to the case of using TS as the AC estimator. A similar quality of the closed-loop performance was observed for the remaining unknown state variables and are not shown for brevity. For instance, Figure 6-9(g) shows an acceptable tracking performance of the jacketed temperature (T_j) when $n = \{\text{EKF}\}$.



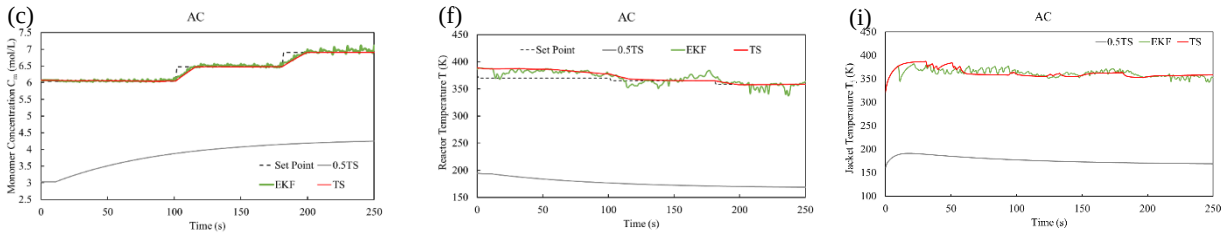


Figure 6-9: Impact of using different AC estimators on (a), (b), (c) control and estimation (MHE) of monomer concentration; (d), (e), (f) control and estimation of reactor temperature; (g), (h), (i) plant output and estimation of jacket temperature

6.2.2.1 HIPS: EKF for state estimation vs EKF for AC estimation

The results presented above have shown the reliability of EKF as the AC approximation method even though the EKF algorithm cannot explicitly take into account process constraints (i.e., Equations (6-9)-(6-10)) and non-symmetric bounded distribution on the process uncertainties, as considered in *Scenario II*. The aim of this section is to further investigate the performance of EKF as an AC estimator for this case study under different closed-loop configurations while using *Scenario II*. Similar to the WTP, three control strategies involving EKF were considered, i.e., *CL1*, *CL2* and *BC* (see descriptions in section 6.2.1.2). As shown in Figures 6-10(a) and 6-10(c), control scheme *CL1* (which uses EKF as the state estimator) failed to provide an acceptable initial condition to NMPC, which led to poor control for both C_m and T. Moreover, Figures 6-10(b) and 6-10(d) show that NMPC failed to compute acceptable control actions for the manipulated variables for control scheme *CL2*. Similar to the discussion presented in section 6.2.1.2, this is due to the error propagation in EKF framework, which is also observed in the estimation provided by MHE for *CL2* (see Figures 6-10(b) and 6-10(d)). Nevertheless, Figures 6-10(a)-6-10(d) present good control and estimation for the *BC* scheme. Since EKF in *BC* receives an acceptable initial condition that takes into account the process constraints (i.e., $\hat{\boldsymbol{x}}_{k-N-1}$ provided by MHE), EKF is able to provide an acceptable posterior estimation of the AC distribution.

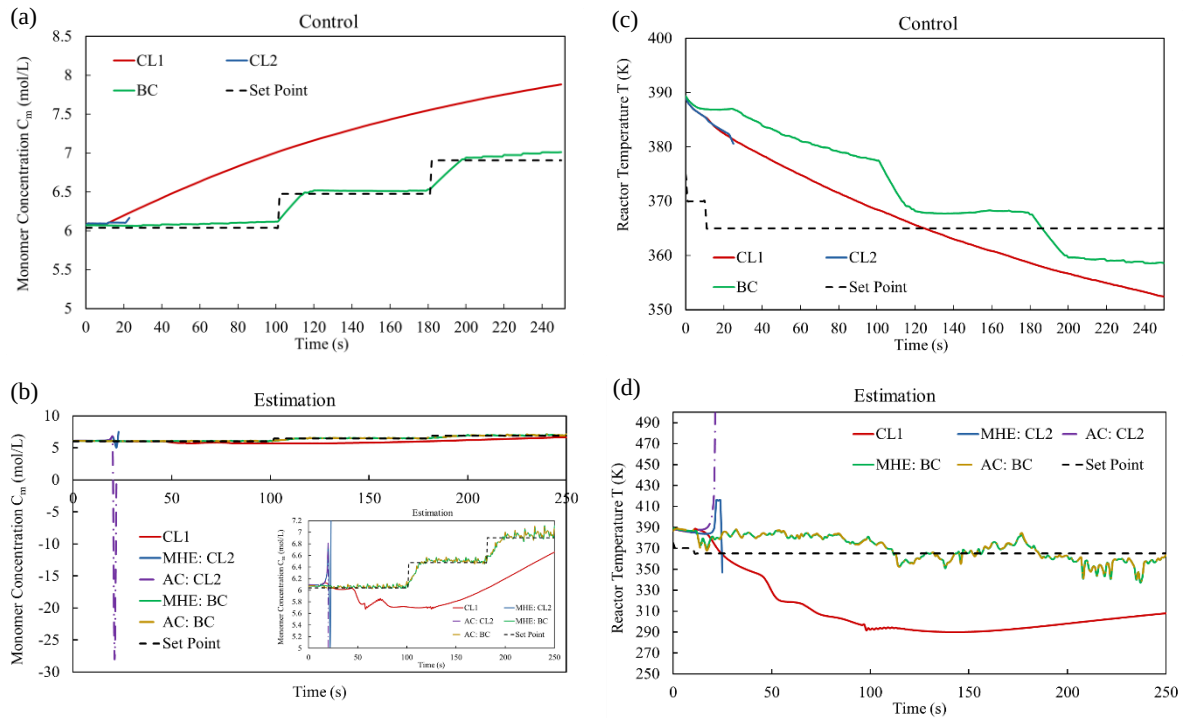


Figure 6-10: Control and estimation results in the closed-loop system based on the role of EKF in the framework for (a), (b) of monomer concentration; (c), (d) reactor temperature;

Note that errors in the EKF estimation will only be propagated for just one time interval. That is, given that the prior distribution at every time interval k is the output of MHE, which considers the constraints and bounds in the process, the estimation error caused by the lack of knowledge about the constraints in the EKF formulation only propagates one iteration, which based on the above, will more likely represent an adequate estimation of the states as AC in MHE. Thus, even though the EKF algorithm by itself may lead to instability in the presence of bounded nonlinear constraints by propagating the estimation error over time, this algorithm is an adequate AC estimator in the MHE framework. This observation agrees with the results presented in sections 6.2.1.2. An error analysis of the EKF framework is presented next to further clarify this insight.

6.2.3 Error Analysis

This section presents an analysis on the error in the estimation of the states in the EKF framework. Considering a general nonlinear model Equations (3-1)-(3-3), the priori and posterior estimation step in the standard EKF are indicated in Equations (3-4) and (3-5), respectively. Note that $\boldsymbol{u}$ is assumed to remain piecewise constant in between time intervals. The error in the estimation (i.e., $\boldsymbol{e}_{k|k}$) can be defined as follows:

$$\boldsymbol{e}_{k|k} = \boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}^{\boldsymbol{EKF}} \tag{6-11}$$

Introducing $\widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}$ into Equation (6-11) gives:

$$\boldsymbol{e}_{k|k} = \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}\right) - \left(\widehat{\boldsymbol{x}}_{k|k}^{\boldsymbol{EKF}} - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}\right) \tag{6-12}$$

By substituting Equations (3-1) and (3-4) for $\boldsymbol{x}_k$ and $\widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}$, respectively, the term $(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}})$ can be evaluated as follows:

$$\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}} = f(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}, k-1) + \boldsymbol{w}_{k-1} - f\left(\widehat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, k-1\right) \tag{6-13}$$

Thus, by substituting Equation (3-5) for $\left(\widehat{\boldsymbol{x}}_{k|k}^{\boldsymbol{EKF}} - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}\right)$, the estimation error is as follows:

$$\begin{aligned}\boldsymbol{e}_{k|k} = \big(& f(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}, k-1) - f\left(\widehat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, k-1\right) + \boldsymbol{w}_{k-1}\big) \\ & - \boldsymbol{K}_k \left(\boldsymbol{y}_k - h\left(\widehat{\boldsymbol{x}}_{k|k-1}^{(\boldsymbol{EKF})}, \boldsymbol{u}_k, k\right)\right)\end{aligned} \tag{6-14}$$

where $\boldsymbol{y}_k$ is defined by Equation (3-2). Thus, Equation (6-14) can be simplified as follows:

$$\begin{aligned}\boldsymbol{e}_{k|k} = \big(& f(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}, k-1) - f\left(\widehat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, k-1\right) + \boldsymbol{w}_{k-1}\big) \\ & - \boldsymbol{K}_k \left(h(\boldsymbol{x}_k, \boldsymbol{u}_k, k) - h\left(\widehat{\boldsymbol{x}}_{k|k-1}^{(\boldsymbol{EKF})}, \boldsymbol{u}_k, k\right) + \boldsymbol{v}_k\right)\end{aligned} \tag{6-15}$$

According to the first order Euler Taylor series expansions:

$$f(\boldsymbol{x}_k, \boldsymbol{u}_k, k) - f\left(\widehat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, k-1\right) \approx \boldsymbol{A}_{k-1}\left(\boldsymbol{x}_{k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1}^{\boldsymbol{EKF}}\right) \tag{6-16}$$

$$h(\boldsymbol{x}_k, \boldsymbol{u}_k, k) - h\left(\widehat{\boldsymbol{x}}_{k|k-1}^{(\boldsymbol{EKF})}, \boldsymbol{u}_k, k\right) \approx \boldsymbol{H}_k\left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}\right) \tag{6-17}$$

Thus, the estimation of error is as follows:

$$\boldsymbol{e}_{k|k} = \boldsymbol{A}_{k-1}\left(\boldsymbol{x}_{k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1}^{\boldsymbol{EKF}}\right) + \boldsymbol{w}_{k-1} - \boldsymbol{K}_k\boldsymbol{H}_k\left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}^{\boldsymbol{EKF}}\right) - \boldsymbol{K}_k\boldsymbol{v}_k \tag{6-18}$$

where $\boldsymbol{x}_{k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1}^{\boldsymbol{EKF}}$ is the estimation error at time interval $k-1$, i.e., $\boldsymbol{e}_{k-1|k-1}$. Thus, given (6-13) and (6-18), the error is as follows:

$$\boldsymbol{e}_{k|k} = \boldsymbol{A}_{k-1}(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{e}_{k-1|k-1} + (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{w}_{k-1} - \boldsymbol{K}_k\boldsymbol{v}_k \tag{6-19}$$

This equation can be re-written as follows:

$$\boldsymbol{e}_{k|k} = \boldsymbol{\Pi}_k\boldsymbol{e}_{k-1|k-1} + \boldsymbol{\psi}_k\boldsymbol{w}_{k-1} - \boldsymbol{K}_k\boldsymbol{v}_k \tag{6-20}$$

where $\boldsymbol{\Pi}$ and $\boldsymbol{\psi}$, i.e., error factor and process uncertainty factor, are defined as follows:

$$\boldsymbol{\Pi}_k = \boldsymbol{A}_{k-1}(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k) \quad (6\text{-}21)$$

$$\boldsymbol{\psi}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)$$

The error function shown in Equation (6-20) can be used to analyze the results of this study for both WTP and HIPS. As mentioned in sections 6.2.1.2 and 6.2.2.1, EKF in *BC* can be an acceptable AC estimator while it fails to provide an accurate estimation of states in *CL2*. This can be explained through the term $\boldsymbol{e}_{k-1|k-1}$ in Equation (6-20), that is a poor initialization provided by EKF in *CL2* framework suggests a relatively large $\boldsymbol{e}_{k-1|k-1}$ that may increase and propagate the error significantly at each iteration step. On the other hand, $\boldsymbol{e}_{k-1|k-1}$ in *BC* remains small and bounded (as shown in Figures (6-8) and (6-10) and reported by their corresponding $\widehat{SSE}_m^{(n)}$). The sources of these errors can be identified by inspecting the evolution of $\boldsymbol{\Pi}$ and $\boldsymbol{\psi}$ for each case study presented in sections 6.2.1.2 and 6.2.2.1. Figures 6-11(a)-(b) compares the average eigenvalue of $\boldsymbol{\Pi}$ for *CL2* and *BC* in WTP. The same comparison is made in Figures 6-11(c)-(d) for the HIPS process. According to Figures 6-11(a) and 6-11(c), the average eigenvalue of the error factor $\boldsymbol{\Pi}$, remains bounded throughout the closed-loop simulation; whereas Figures 6-11(b) and 6-11(d) show that the average eigenvalue of $\boldsymbol{\Pi}$ becomes unbounded after a few iterations.

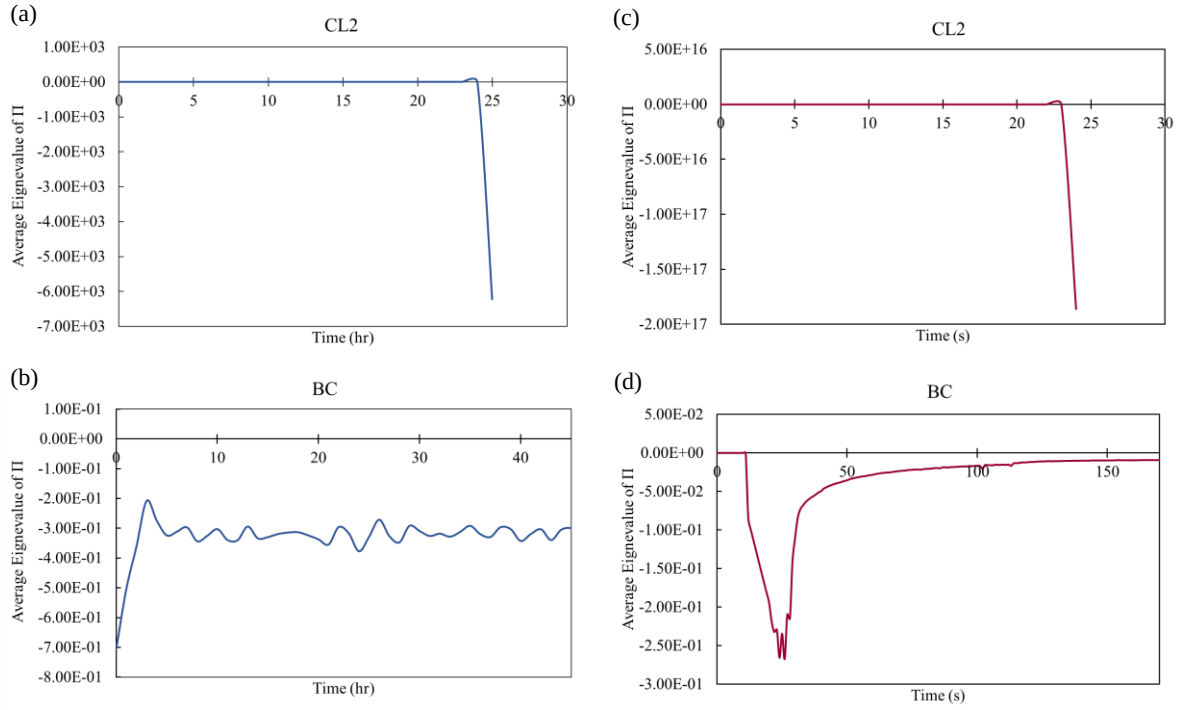


Figure 6-11: Average eigenvalue of **Π** factor in the case of *CL2* and *BC* framework for (a), (b) WTP; (c), (d) HIPS;

The same tests have been performed for $\boldsymbol{\psi}$. The results showed that the average eigenvalue for this factor remain bounded (not shown for brevity), which is an indication that this factor does not have a major role in the failure of EKF in *CL2*. Therefore, the behaviour of the eigenvalues in the error factor **Π** can be identified by evaluating the eigenvalues of the Jacobian matrix for control schemes *BC* and *CL2* at each time interval. In the control scheme *CL2*, the Jacobian matrix is updated at every time interval k based on the posterior estimation of the states at the past time interval, i.e., $\widehat{\boldsymbol{x}}_{k-1|k-1}^{\boldsymbol{EKF}}$. Therefore, as the estimation error $\boldsymbol{e}_{k-1|k-1}$ propagates over time, the effect of inaccurate estimations of $\widehat{\boldsymbol{x}}_{k-1|k-1}^{\boldsymbol{EKF}}$ provided to the Jacobian matrix becomes more significant. After several time intervals, the effect of the error may become large and produce unbounded eigenvalues in the sensitivity matrices, which eventually results in an unstable operation for the control scheme *CL2*. This behaviour agrees with the results presented for WTP and HIPS in sections 6.2.1.2 and 6.2.2.1, respectively.

6.3 Summary

This study showed that an AC approximation method is an essential component in the MHE framework to attain an acceptable closed-loop performance. Nevertheless, as long as a standard state estimation method is used to approximate the AC distribution and is engaged with MHE, open-loop and closed-loop operation (even in the presence of nonlinear process constraints) are expected to operate adequately, i.e., the states estimates are expected to follow the plant states and therefore become good educated initial guesses for the NMPC framework. This research showed that while EKF is an unconstrained state estimation scheme, this estimation method is an efficient AC estimator and expected to provide an adequate optimization-based estimation and control of industrial and/or complex systems that operate in closed-loop using MHE and NMPC. This is because EKF is provided with a proper initialization when coupled with a constrained state estimation scheme such as MHE. Moreover, the error analysis on the EKF performance was conducted to this work that provided insights on the performance of EKF as AC estimator for constrained applications. Base on this analysis, the estimation error for EKF as the AC estimator when couples with MHE remains bounded. This insight supports the capability of EKF as the AC estimator. This suggests that more computationally expensive filters such as C-PF may not be required as AC estimators since they may not seem to improve the quality of the MHE estimation when compared to a more computationally efficient estimation method such as EKF.

The outcomes of this research was validated for industrial scale systems that exhibit complex dynamics and unstable open-loop operation, and under a variety of restrictions on the process uncertainties, measurement noises, feasibility constraints and different plant designs.

In addition to the arrival cost, the MHE performance highly relies on accurate identification of the measurement noises and process uncertainties present in the system. Thus, it is essential to provide accurate models to the MHE scheme that describe the density of these random variables properly. This motivates the studies presented in the next two chapters of this thesis.

Chapter 7

Extended Moving Horizon Estimation (EMHE)

In order to achieve a high performance estimation while using the MHE framework, it is essential to identify models that can adequately capture the behaviour of the three major components of the objective function in the MHE problem, i.e., arrival cost (AC), process uncertainties, and measurement noises. While the previous chapter focused on assessing the importance of the AC in the MHE framework, this chapter aims to improve the MHE performance by improving the models representing the process uncertainties and measurement noises. A common assumption when using MHE is to consider that the distribution of the process uncertainties and measurement noises can be described by zero-mean Gaussian distributions with known covariance matrices. However, most of the real-world applications featuring noises/uncertainties that follow non-Gaussian distributions. A review of the literature presented in section 2.3.2 reveals that no study has considered an efficient modification on the standard MHE formulation to capture the non-Gaussianity of noises/uncertainties. Moreover, there might be scheduled changes occurring in the plant that may lead to a change in the density of the process uncertainties and measurement noises. As a reminder, in this work, the distribution that changes due to changes in the operational conditions is referred to as *time-dependent distribution*. Likewise, the distribution that remains unchanged during operation is referred to as *time-independent distribution*. As these changes in the plant operation are scheduled, the new distribution of the noises/uncertainties are known *a priori* to the MHE framework. However, a development on the MHE-based estimation schemes that can capture the changes in the distribution of the noises/uncertainties online is lacking from the literature.

Chapters 4 and 5 have presented a novel efficient scheme to use Gaussian mixture models to the state estimation framework (i.e., referred to as AGS-EKF) to capture the non-Gaussianity of the process uncertainties and measurement noises. Motivated by the previous works on AGS-EKF, the present work aims to make use of the Gaussian mixture models to improve the performance of MHE for the applications involving non-Gaussian process uncertainties/ measurement noises. Note that similar to the common assumption made when performing model-based state estimation schemes (e.g., MHE, EKF, etc.) the current work focuses on applications where the distributions of the process uncertainties and measurement noises are known *a priori* over the operating time. Such probability distributions can be obtained from historical data or process heuristics. The main contributions of this work are as follows:

- A novel extended version of MHE is developed. The scheme, referred to as *Extended Moving Horizon Estimation* (EMHE), relaxes the zero-mean Gaussian assumption considered for the process uncertainties and measurement noises in the standard MHE. In the proposed Extended MHE, the objective function is modified to consider non-zero mean Gaussian distributions for the noisy variables. Later in this chapter, it is shown that the standard MHE is a special case of EMHE. Also, the proposed EMHE is not expected to increase the computational load compared to the standard MHE, thus making the proposed approach computationally attractive.
- The non-Gaussian distributions for the process uncertainties and measurement noises are approximated using Gaussian mixture models. Given that the distributions of these random variables are assumed to be known *a priori*, the corresponding Gaussian mixture models can be estimated offline. Thus, this step does not increase the CPU time in the online estimation scheme. The proposed EMHE uses the mean value and covariance matrix of the overall Gaussian mixture model to approximate the non-Gaussian distribution of the process uncertainty/ measurement noises. These main characteristics of the corresponding Gaussian mixture models are introduced to the EMHE objective function, which helps to reduce the estimation error for such applications in comparison to the standard MHE. Note that the use of Gaussian mixture model in MHE framework has not been reported in the literature.
- The estimation scheme based on the proposed EMHE offers the opportunity to perform an online adaptation on the time-dependent known distributions (Gaussian/Non-Gaussian) of the process uncertainties and measurement noises. This is a practical feature for those processes that are subject to significant changes in the distribution of noises and uncertainties due to changes in their operating conditions, e.g., seasonal changes during operation. For this purpose, EMHE allows the estimation scheme to re-approximate the corresponding Gaussian mixture models of the process uncertainties and measurement noises upon any changes in their corresponding distributions. Note that expectation-maximization (EM) algorithm is an efficient method used to re-approximate these Gaussian mixture models. The additional computational costs needed to re-approximate the Gaussian mixture models highly depends on the frequency at which the scheduled change in operation happens in the plant, i.e., how often the density function of the noise/uncertainty changes. The higher the frequency at which these events take place in a system, the more often the Gaussian mixture models need to be recalculated thus increasing the CPU costs. Nevertheless, it is expected that such changes would not occur within short time intervals in an actual industrial setting. Note

that computer parallelization techniques can be adopted to simultaneously update the Gaussian mixture models of the noises/uncertainties, thus reducing the CPU costs.

Note that the mean value of the overall Gaussian mixture models of the corresponding non-zero mean non-Gaussian noises/uncertainties may also have a non-zero mean. Thus, the aforementioned modification in the MHE framework is required such that the non-zero mean and the covariance matrix of the Gaussian mixture models can be adopted within the MHE formulation. That is, the main interest of the modification considered in EMHE (i.e., relaxing the zero-mean assumption) is only to describe systems with non-zero mean Gaussian mixture models.

This chapter is organized as follows: section 7.1 presents the proposed Extended MHE (EMHE). Section 7.2 describes the scheme required to perform recursive state estimation for applications featuring process uncertainties and measurement noises that follow time-dependent distributions, i.e., the density of these random variables changes throughout the process. The computational experiments conducted to this work are presented in section 7.3, which includes three chemical engineering processes to investigate benefits offered by the proposed method in both open-loop and closed-loop when compared to standard MHE. Section 7.4 presents the chapter summary.

7.1 Extended Moving Horizon Estimation (EMHE)

The proposed Extended MHE aims to improve the MHE performance for the applications in which process uncertainties or/and measurement noises follow known non-zero mean Gaussian distributions. This chapter considers a nonlinear dynamic system given by Equations (3-1)-(3-2). MHE is a special case of the full information problem where the length of estimation horizon (N) is fixed and finite. In general, the standard full information problem (and correspondingly, standard MHE) assumes that the process uncertainties ($\boldsymbol{w}$) and measurement noises ($\boldsymbol{v}$) are mutually uncorrelated and follow the zero-mean Gaussian distributions with diagonal covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively, i.e.

$$\mathbb{E}\left[\boldsymbol{w}_{\varsigma}\boldsymbol{w}_{k}^{T}\right]=\begin{cases}\boldsymbol{Q}, & \varsigma=k\\ \boldsymbol{0}, & \varsigma\neq k\end{cases};\quad \mathbb{E}\left[\boldsymbol{v}_{\varsigma}\boldsymbol{v}_{k}^{T}\right]=\begin{cases}\boldsymbol{R}, & \varsigma=k\\ \boldsymbol{0}, & \varsigma\neq k\end{cases} \tag{7-1}$$

$$\boldsymbol{w}\in\mathbb{R}^{n_x},\boldsymbol{v}\in\mathbb{R}^{n_y},\boldsymbol{Q}\in\mathbb{R}^{n_x\times n_x},\boldsymbol{R}\in\mathbb{R}^{n_y\times n_y}$$

Moreover, $\mathbb{E}\left[\boldsymbol{w}_{\varsigma}\boldsymbol{x}_{0}^{T}\right]=0$ for all k where $\boldsymbol{x}_0$ represents the vector of initial states. The full information problem provides the optimal estimations of the state variables. That is, the optimal estimation is determined by maximizing the posterior PDF of the states over the estimation horizon. The full

information problem for the nonlinear model presented in Equations (3-1), (3-2), and (7-1) is formulated as follows[194]:

$$\{\hat{\boldsymbol{x}}_{0|k}, \dots, \hat{\boldsymbol{x}}_{k|k}\} := arg \max_{\{\boldsymbol{x}_j\}_{j=0}^{k}} p(\boldsymbol{X}_0^k | \boldsymbol{Y}_0^k) \tag{7-2}$$

where p denotes the probability density function, $\hat{\boldsymbol{x}}$ represents the optimal estimates, $\boldsymbol{Y}_0^k = \{\boldsymbol{y}_0, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k\}$ and $\boldsymbol{X}_0^k = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}$ represents the set of the measurements and the states of the systems, respectively. Given the underlying assumptions in the full information problem, i.e., the states follow a first order Markov process, the process uncertainties and the measurement noises are independent, and the available measurements are mutually independent, Equation (7-2) can be rewritten as follows:

$$\{\hat{\boldsymbol{x}}_{0|k}, \dots, \hat{\boldsymbol{x}}_{k|k}\} := arg \max_{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=0}^{k}} \log p(\boldsymbol{x}_0) + \sum_{j=0}^{k} \log p(\boldsymbol{v}_j) + \sum_{j=0}^{k-1} \log p(\boldsymbol{w}_j) \tag{7-3}$$

The details of the derivation of the Equation (7-3) from Equations (3-1),(3-2),(7-2) can be found elsewhere[194]. Given $\boldsymbol{x}_0 \sim \mathcal{N}(\hat{\boldsymbol{x}}_{0|0}, \boldsymbol{P}_0)$, $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$, and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$, Equation (7-3) can be expressed as follows:

$$\{\hat{\boldsymbol{x}}_{0|k}, \dots, \hat{\boldsymbol{x}}_{k|k}\} := arg \min_{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=0}^{k}} \|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}_0^{-1}} + \sum_{j=0}^{k-1} \|\boldsymbol{w}_j\|^2_{\boldsymbol{Q}^{-1}} + \sum_{j=0}^{k} \|\boldsymbol{v}_j\|^2_{\boldsymbol{R}^{-1}} \tag{7-4}$$

From Equation (7-4), i.e., the objective function of the general full information formulation, the standard MHE formulation can be derived. That is, MHE considers only the last *N* measurements to perform the estimation. In the standard MHE formulation, as shown in Equations (6-1)-(6-2), a penalty term φ_{k-N}, referred to as the *arrival cost*, is introduced to represent a metric of the past information that is discarded from the finite estimation horizon in the MHE scheme.

The proposed EMHE relaxes the zero-mean assumption for the uncertainties and noises in the nonlinear system and complies with the assumptions mentioned above. That is, the PDF of the non-zero mean Gaussian process uncertainties and measurement noises can be described by $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{Q})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}, \boldsymbol{R})$, respectively. Thus, the full information objective function in Equation (7-3) for the nonlinear system (i.e., Equations (3-1)-(3-2)) can be rewritten as follows:

$$\{\hat{\boldsymbol{x}}_{0|k}, \dots, \hat{\boldsymbol{x}}_{k|k}\} := arg \min_{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=0}^{k}} \left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_{0|0}\right\|^2_{\boldsymbol{P}_0^{-1}} + \sum_{j=0}^{k-1} \left\|\boldsymbol{w}_j - \boldsymbol{\mu}\right\|^2_{\boldsymbol{Q}^{-1}} + \sum_{j=0}^{k} \left\|\boldsymbol{v}_j - \boldsymbol{\tau}\right\|^2_{\boldsymbol{R}^{-1}} \quad (7\text{-}5)$$

Consequently, the proposed EMHE formulation can be derived from Equation (7-5) as follows:

$$\min_{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=k-N}^{k-1}} \sum_{j=k-N}^{k-1} \left\|\boldsymbol{w}_j - \boldsymbol{\mu}\right\|^2_{\boldsymbol{Q}^{-1}} + \sum_{j=k-N+1}^{k} \left\|\boldsymbol{v}_j - \boldsymbol{\tau}\right\|^2_{\boldsymbol{R}^{-1}} + \varphi_{k-N} \quad (7\text{-}6)$$

s.t.

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

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

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

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

where:

$\boldsymbol{\mu} \in \mathbb{R}^{n_x}, \boldsymbol{\tau} \in \mathbb{R}^{n_y}$

Note that the standard MHE is a special case of the proposed EMHE. That is, if process uncertainties and measurement noises belong to zero-mean Gaussian distributions, then, $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\tau} = \mathbf{0}$; hence, Equation (7-6) is equivalent to the standard MHE formulation presented in Equation (6-1).

This modification in the MHE formulation presented in Equation (7-6) allows EMHE to adopt uncertainties and noises that follow non-zero mean Gaussian distributions to the state estimation framework at no additional computational costs. Moreover, EMHE combined with the Gaussian mixture models can outperform standard MHE for applications involving non-zero mean non-Gaussian process uncertainties and measurement noises. To pursue this goal, the proposed EMHE uses the mean value and covariance of the Gaussian mixture models, as a suitable approximation of the non-zero mean non-Gaussian process uncertainties and measurement noises present in the process. This EMHE-based state estimation scheme is presented next.

7.2 State estimation scheme: time-dependent non-Gaussian distributions

This section presents the state estimation scheme proposed in this study to deal with systems involving non-Gaussianity and time-dependency of the densities of the random noises and uncertainties. The non-Gaussian random variables present in the process can be described by the main characteristics of their

corresponding Gaussian mixture models, as shown in Equation (4-13). Substituting the mean vector and covariance of the Gaussian mixture models, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{GM}}, \boldsymbol{Q}^{\boldsymbol{GM}})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}^{\boldsymbol{GM}}, \boldsymbol{R}^{\boldsymbol{GM}})$, in the proposed EMHE formulation shown in Equation (7-6) can introduce an adequate approximation of the non-Gaussian process uncertainties and measurement noises to the state estimation. The Gaussian mixture models can be updated online once the system moves to operating conditions where the noises' densities change, i.e., on cases when the distribution of the noises are time-dependent.

Table 7-1 presents the step-by-step procedure required to provide the point estimates using the proposed estimation scheme. As shown in this table, Equations (7-7)-(7-12) are used to evaluate the state estimation at each time interval. The notation used in these equations is provided below. *Steps I* and *II* in Table 7-1 (i.e., Equations (7-7)-(7-10)) use *a priori* knowledge of the process uncertainties and measurement noises at the current time interval k to approximate time-dependent distributions with adequate Gaussian mixture models. Note that this work considers that the random variables present in the system are independent from each other. Thus, in *Step I*, Equations (7-7),(7-8) perform the univariate EM algorithm based on *a priori* knowledge of the process uncertainties and measurement noises at the current time interval k to approximate the parameters of Gaussian components in the Gaussian mixture model of each process uncertainty/measurement noise. The main entries to the EM algorithm (denoted as the EM function in Equations (7-7) and (7-8)) are the densities of the non-Gaussian process uncertainties and measurement noises at the current time interval k, which are described with a large number of samples. This information is used at each time interval to update the parameters of the corresponding Gaussian mixture model. In *Step II*, the overall mean value and covariance of each Gaussian mixture model are obtained from Equations (7-9) and (7-10). *Step III* provides the mean vector and covariance matrix of the process uncertainties and measurement noises based on the characteristics of the process uncertainties and noises obtained from *Step II*. Note that the notation used to represent the mean vector and covariance of the corresponding Gaussian mixture models of these time-dependent distributions of the process uncertainties at the current time interval (k) are $\boldsymbol{\mu}_k^{\boldsymbol{GM}}$ and $\boldsymbol{Q}_k^{\boldsymbol{GM}}$, respectively. Likewise, $\boldsymbol{\tau}_k^{\boldsymbol{GM}}$ and $\boldsymbol{R}_k^{\boldsymbol{GM}}$ denote the mean vector and covariance matrix to approximate the time-dependent distributions of the measurement noises at the current time interval k. As shown in *Step IV*, this information (i.e., $\boldsymbol{\mu}_k^{\boldsymbol{GM}}$, $\boldsymbol{Q}_k^{\boldsymbol{GM}}$, $\boldsymbol{\tau}_k^{\boldsymbol{GM}}$, and $\boldsymbol{R}_k^{\boldsymbol{GM}}$) represents the main inputs to perform the state estimation at each time interval k using the EMHE framework, as per Equation (7-12).

Table 7-1: EMHE-based state estimation algorithm under time-dependent non-Gaussian distribution of noises/uncertainties

At each time interval k:

Step I: Update Gaussian components based on the current PDFs, i.e., $p(w_{l_k})$ and $p(v_{c_k})$

$$\left(\mu_l^{ip}{}_k, Q_l^{ip}{}_k, \alpha p_l^{ip}{}_k\right) = EM\left(p(w_{l_k}), k\right);\ w_{l_k} \in \boldsymbol{w_k};\ \forall k;\ \forall l = 1, \dots n_x \tag{7-7}$$

$$\left(\tau_c^{im}{}_k, R_c^{im}{}_k, \alpha m_c^{im}{}_k\right) = EM\left(p(v_{c_k}), k\right);\ v_{c_k} \in \boldsymbol{v_k};\ \forall k;\ \forall c = 1, \dots n_y \tag{7-8}$$

Step II: Update the characteristics of Gaussian mixture models

$$\begin{aligned}
&\mu_l^{GM}{}_k = \textstyle\sum_{ip=1}^{ngp(l)} \alpha p_l^{ip}{}_k \mu_l^{ip}{}_k \\
&Q_l^{GM}{}_k = \textstyle\sum_{ip=1}^{ngp(l)} \alpha p_l^{ip}{}_k Q_l^{ip}{}_k + \sum_{ip=1}^{ngp(l)} \alpha p_l^{ip}{}_k \left(\mu_l^{ip}{}_k - \mu_l^{GM}{}_k\right)\left(\mu_l^{ip}{}_k - \mu_l^{GM}{}_k\right)^{\boldsymbol{T}} \\
&\textstyle\sum_{ip=1}^{ngp(l)} \alpha p_l^{ip} = 1; \quad \alpha p_l^{ip} \geq 0 \qquad \forall k; \forall l = 1, \dots n_x
\end{aligned} \tag{7-9}$$

$$\begin{aligned}
&\tau_c^{GM}{}_k = \textstyle\sum_{im=1}^{ngm(c)} \alpha m_c^{im}{}_k \tau_c^{im}{}_k \\
&R_c^{GM}{}_k = \sum_{im=1}^{ngm(c)} \alpha m_c^{im}{}_k R_c^{im}{}_k + \sum_{im=1}^{ngm(c)} \alpha m_c^{im}{}_k \left(\tau_c^{im}{}_k - \tau_c^{GM}{}_k\right)\left(\tau_c^{im}{}_k - \tau_c^{GM}{}_k\right)^{\boldsymbol{T}} \\
&\textstyle\sum_{im=1}^{ngm(c)} \alpha m_c^{im} = 1; \quad \alpha m_c^{im} \geq 0 \quad \forall k; \forall c = 1, \dots n_y
\end{aligned} \tag{7-10}$$

Step III: Update the mean-vectors and covariance matrices

$$\begin{aligned}
&\boldsymbol{\mu_k^{GM}} = \left\{\mu_l^{GM}{}_k\right\}_{l=1}^{l=n_x};\ \boldsymbol{Q_k^{GM}} = diag\left(\left[Q_l^{ip}{}_k\right]_{l=1}^{l=n_x}\right); \qquad \forall k \\
&\boldsymbol{\tau_k^{GM}} = \left\{\tau_c^{GM}{}_k\right\}_{c=0}^{c=n_y};\ \boldsymbol{R_k^{GM}} = diag\left(\left[R_c^{im}{}_k\right]_{c=1}^{c=n_y}\right); \qquad \forall k
\end{aligned} \tag{7-11}$$

Step IV: Perform the estimation using the proposed EMHE

$$\min_{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=k-N}^{k-1}} \sum_{j=k-N}^{k-1} \left\|\boldsymbol{w}_j - \boldsymbol{\mu_k^{GM}}\right\|_{\boldsymbol{Q_k^{GM}}^{-1}}^2 + \sum_{j=k-N+1}^{k} \left\|\boldsymbol{v}_j - \boldsymbol{\tau_k^{GM}}\right\|_{\boldsymbol{R_k^{GM}}^{-1}}^2 + \varphi_{k-N} \tag{7-12}$$

s.t.

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

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

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

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

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

As shown in Table 7-1, n_x and n_y are the number of state variables and measurements in the system, respectively. As presented in section 5.1.2, scalar v_{c_k} refers to the c^{th} element of vector $\boldsymbol{v}_k$, which represents the measurement noise associated with c^{th} measurement at time interval k; scalar w_{l_k} refers to the l^{th} element of vector $\boldsymbol{w}_k$. ip, and im denote the indexes for the Gaussian component in the Gaussian mixture model of process uncertainties and measurement noises, respectively. For instance, αm_c^{im} is the weight assigned to the im^{th} component in the Gaussian mixture model of c^{th} measurement noise, whereas αp_l^{ip} is the corresponding weight for the ip^{th} Gaussian component in the Gaussian mixture model of the l^{th} process uncertainty variable, respectively; likewise, $ngp(l)$ and $ngm(c)$ are scalars that denote the number of Gaussian components in the corresponding mixture model of the l^{th} process uncertainty and c^{th} measurement noise, respectively. The function EM in Equations (7-7)-(7-8) represents the univariate expectation-maximization method. Note that similar to the standard MHE, the EMHE problem is expected to present the optimal estimations of the unknown states that satisfy all the process constraints. That is, from a mathematical point of view, the estimations provided by both the standard MHE and EMHE are expected to remain within the feasibility region of the process. In practice, however, the estimates resulting from EMHE and the standard MHE may become infeasible for the actual process if there is significant plan-model mismatch.

Remark 1: To simplify the analysis, the distribution of the noises within the estimation horizon are assumed to be the same as the distribution of the noises at the current time interval k, i.e., $\{\boldsymbol{w}_i\}_{i=k-N}^{k} \sim \mathcal{N}(\boldsymbol{\mu}_k^{\boldsymbol{GM}}, \boldsymbol{Q}_k^{\boldsymbol{GM}})$ and $\{\boldsymbol{v}_i\}_{i=k-N}^{k} \sim \mathcal{N}(\boldsymbol{\tau}_k^{\boldsymbol{GM}}, \boldsymbol{Q}_k^{\boldsymbol{GM}})$. Nevertheless, the proposed approach can be extended to cases when this assumption is relaxed. In that case, Equation (7-11) in *Step III* needs to be modified to keep track of all the historical distributions of the noises/uncertainties within the estimation horizon. This can be done by considering an additional dimension in the matrix of the mean values and covariance, e.g., $\boldsymbol{\mu}_{[k-N:k]}^{\boldsymbol{GM}} = \left[\{\mu_l^{GM}{}_{k-N}\}_{l=1}^{l=n_x}, \dots\dots, \{\mu_l^{GM}{}_{k}\}_{l=1}^{l=n_x}\right]$. Note that the subscripts for these terms in objective function in Equation (7-12) would need to change from k to j, i.e.,

$$\min_{\{\boldsymbol{x}_j,\boldsymbol{w}_j\}_{j=k-N}^{k-1}} \sum_{j=k-N}^{k-1} \left\|\boldsymbol{w}_j - \boldsymbol{\mu}_j^{\boldsymbol{GM}}\right\|^2_{\boldsymbol{Q}_j^{\boldsymbol{GM}^{-1}}} + \sum_{j=k-N+1}^{k} \left\|\boldsymbol{v}_j - \boldsymbol{\tau}_j^{\boldsymbol{GM}}\right\|^2_{\boldsymbol{R}_j^{\boldsymbol{GM}^{-1}}} + \varphi_{k-N} \tag{7-12A}$$

Remark 2: For the scenarios where the distributions of the process uncertainties and measurement noises are assumed to remain constant during operation (i.e., time-independent distributions), the inputs to the EMHE framework would not change. That is, the estimation scheme does not require to perform *Steps I-III* at each time interval. For instance, the scenario featuring time-independent non-zero mean Gaussian distributed random noises, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{Q})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{\tau}, \boldsymbol{R})$, is a special case of the estimation scheme presented in Table 7-1 in which $\boldsymbol{\mu}_k^{\boldsymbol{GM}} = \boldsymbol{\mu}, \boldsymbol{Q}_k^{\boldsymbol{GM}} = \boldsymbol{Q}, \boldsymbol{\tau}_k^{\boldsymbol{GM}} = \boldsymbol{\tau}$, and $\boldsymbol{R}_k^{\boldsymbol{GM}} = \boldsymbol{R}$. Likewise, the scenario involving time-independent zero-mean Gaussian distributions, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$, is a special case of estimation scheme presented above in which $\boldsymbol{\mu}^{\boldsymbol{GM}}{}_k = \boldsymbol{0}, \boldsymbol{Q}^{\boldsymbol{GM}}{}_k = \boldsymbol{Q}, \boldsymbol{\tau}^{\boldsymbol{GM}}{}_k = \boldsymbol{0}$, and $\boldsymbol{R}^{\boldsymbol{GM}}{}_k = \boldsymbol{R}$.

Remark 3: The current study assumes that changes in the operation are scheduled in advance and therefore, the corresponding probability densities of the measurement noises and process uncertainties are known *a priori*. The proposed EMHE scheme is expected to converge in the estimation as long as an adequate representation of the non-Gaussian distributions is provided using Gaussian mixture models. However, when unexpected events happen in the plant, e.g., drastic changes in the operating conditions or equipment malfunction, the process uncertainties and the measurement noises may drastically drift from their original distributions. Under that scenario, the *a priori* knowledge of the non-Gaussian distributions may no longer be valid and thus, the corresponding Gaussian mixture models may no longer provide an accurate approximation of the noises/ uncertainties in the system. Consequently, the present EMHE-scheme may lose performance or diverge as this framework heavily relies on the accuracy of these Gaussian mixture models. Future research studies (presented in chapter 8) would focus on this limitation and provide robust moving horizon schemes that can accommodate unknown densities in the measurement noises or process uncertainties.

Remark 4: In the present work, the process uncertainties (and measurement noises) associated with each state variable are mutually independent and consequently, each process uncertainty variable (measurement noise variable) is approximated to a univariate Gaussian mixture model, as shown in Table 7-1 (Equations (7-7)-(7-11)). Nevertheless, the present EMHE estimation scheme can be readily extended to consider applications featuring multivariate distributions of the process uncertainties/ measurement noises, i.e., correlated process uncertainties or measurement noises. Under those

scenarios, the EMHE scheme makes use of multivariate Gaussian mixture models, i.e., Equations (7-13A)-(7-13D), as a substitute to univariate Gaussian mixture models shown in Equations (7-7)-(7-11).

$$\left(\boldsymbol{\mu}_k^{ip}, \boldsymbol{Q}_k^{ip}, \alpha p_k^{ip}\right) = EM(p(\boldsymbol{w}_k), k); \quad \forall k \tag{7-13A}$$

$$\left(\boldsymbol{\tau}_k^{im}, \boldsymbol{R}_k^{im}, \alpha m_k^{im}\right) = EM(p(\boldsymbol{v}_k), k); \quad \forall k \tag{7-13B}$$

$$\boldsymbol{\mu}_k^{\boldsymbol{GM}} = \sum_{ip=1}^{ngp} \alpha p_k^{ip} \boldsymbol{\mu}_k^{ip} \tag{7-13C}$$

$$\boldsymbol{Q}_k^{\boldsymbol{GM}} = \sum_{ip=1}^{ngp} \alpha p_k^{ip} \boldsymbol{Q}_k^{ip} + \sum_{ip=1}^{ngp} \alpha p_k^{ip} \left(\boldsymbol{\mu}_k^{ip} - \boldsymbol{\mu}_k^{\boldsymbol{GM}}\right)\left(\boldsymbol{\mu}_k^{ip} - \boldsymbol{\mu}_k^{\boldsymbol{GM}}\right)^{\boldsymbol{T}}$$

$$\sum_{ip=1}^{ngp} \alpha p_l^{ip} = 1; \quad \alpha p_l^{ip} \geq 0 \qquad \forall k$$

$$\boldsymbol{\tau}_k^{\boldsymbol{GM}} = \sum_{im=1}^{ngm} \alpha m_k^{im} \boldsymbol{\tau}_k^{im} \tag{7-13D}$$

$$\boldsymbol{R}_k^{\boldsymbol{GM}} = \sum_{im=1}^{ngm} \alpha m_k^{im} \boldsymbol{R}_k^{im} + \sum_{im=1}^{ngm} \alpha m_k^{im} \left(\boldsymbol{\tau}_k^{im} - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)\left(\boldsymbol{\tau}_k^{im} - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)^{\boldsymbol{T}}$$

$$\sum_{im=1}^{ngm} \alpha m^{im} = 1; \quad \alpha m^{im} \geq 0 \quad \forall k$$

Note that *Step III* in the EMHE-based estimation algorithm is not required when random noises/uncertainties follow multivariate distributions. Moreover, *Step IV* (Equation (7-12)) remains unchanged under both correlated and uncorrelated noises scenarios.

Remark 5: Preliminary tests showed that the arrival cost exhibits similar effects in the performance of the EMHE scheme to those observed for the standard MHE (tests not shown for brevity). Longer estimation horizons improve the estimation but also increase the CPU time. Conversely, short estimation horizons reduce the CPU time to perform the point estimates, but it may impact the convergence of the state estimation scheme because the performance of both EMHE and standard MHE heavily relies on the accuracy offered by the arrival cost estimator (see chapter 6). As the present work is focused on the effect of noises/uncertainties in the state estimation schemes, a sufficiently long estimation horizon has been considered for each case study, which tends to minimize the impact of the arrival cost term in the standard MHE and EMHE schemes. Nevertheless, the proposed EMHE approach is expected to offer a similar performance if an appropriate arrival cost estimator is available.

7.3 Computational Experiments

This section presents the results of performing the proposed EMHE for three chemical engineering case studies, i.e., a Gas-phase reactor, an industrial-scale wastewater treatment plant (WTP), and a network of continuous-stirred tank reactors (CSTRs). This study used the mean squared error (MSE) defined in Equation (4-22) as the main metric to analyze the performance of the proposed EMHE scheme in open-

loop. Note that in the current chapter $n \in \{\text{MHE}, \text{EMHE}\}$. As mentioned in the previous chapters, this study considered the output of the mechanistic plant model (associated with the additive process uncertainties), i.e., "Plant Output", as the true value of states. Moreover, Equation (7-14) represents MSE used to evaluate the performance of the system in closed-loop using the n^{th} estimation scheme.

$$MSE_{x_m}^{(n)} = \frac{1}{t_f}\sum_{k=0}^{t_f}\left(x_{k,m} - y_{ref}{}_{k,m}\right)^2 \tag{7-14}$$

where $y_{ref}{}_{k,m}$ represents the set-points for the m^{th} state (controlled variable) at k^{th} time interval. The computational experiments were performed in Pyomo 5.2 (and Python 3.6) on a computer running Microsoft Windows Server 2016 standard. The computer specification is 96 GB RAM and Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz 2.10 GHz (2 processors). The backward method was used to discretize the nonlinear dynamic process models. Interior-point method was used to solve the optimization problems, i.e., MHE, EMHE, and NMPC. Note that the true distributions of the uncertainties/ noises for each case study were defined as a function (i.e., percentage) of the nominal steady-state value of their corresponding state variables. This was done to test the performance of the proposed estimation framework under operating conditions that are likely to occur in an actual industrial setting. Hence, multiple scenarios involving Gaussian and non-Gaussian densities for simple systems, and those involving complex highly nonlinear processes often found in industrial settings, are presented in this section. The results obtained for each of the case studies are presented next.

7.3.1 Gas-phase Reactor

The proposed EMHE was used to perform state estimation in a gas-phase reactor presented in Equation (4-26), except that the assumptions for the process uncertainties and the measurement noise is different than that shown in Equation (4-26). That is, the current study considers that the process uncertainties and the measurement noise present in this process follow the non-Gaussian distributions shown in Figure 7-1. As shown in this figure, the process uncertainties considered in this study follow uniform distributions while the random measurement noise is assumed to follow a multi-modal distribution (with three modes). The red plot in this figure represents the Gaussian mixture model provided by the EM algorithm that approximates each of the non-Gaussian distributions. Each of the Gaussian mixture models involves three Gaussian components, which have been chosen offline after performing multiple tests to achieve an acceptable level of accuracy of the mixture model without an unnecessary increase in the Gaussian components.

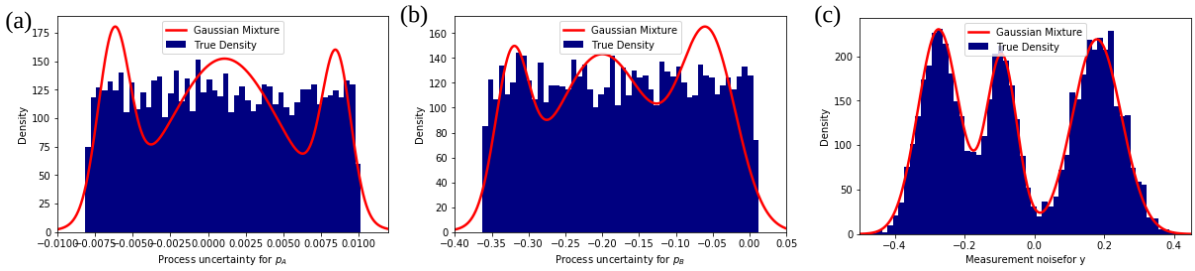


Figure 7-1: Histogram for the true non-Gaussian distribution and the Gaussian mixture approximation of (a) the process uncertainty associated with p_A; (b) process uncertainty associated with p_B; (b) measurement noise

For comparison purposes, the performance of the standard MHE, i.e., Equation (6-1), was also considered for this case study. In this case, zero-mean Gaussian process uncertainties and measurement noise with the following specifications are considered:

$$v_k \sim \mathcal{N}(0, R);\ R = 0.0169;\ w_k \sim \mathcal{N}(0, Q);\ Q = diag\{9.0 \times 10^{-6}, 2.5 \times 10^{-3}\} \qquad (7\text{-}15)$$

To make a fair comparison, preliminary tests were performed to adjust the covariance matrices for these random variables in the interest of obtaining the highest estimation accuracy using the standard MHE. The initial guess used to perform the estimation is assumed to be 10% larger than the true initial states, i.e., $1.1 \times p_{A_0}$ and $1.1 \times p_{B_0}$. Both of these states are unknown and need to be estimated online, i.e., the total pressure (y_k) is the only available measurement. The sampling interval considered for the gas-phase reactor is 0.1s. The length of estimation horizon is chosen to be 10, which is long enough to minimize the effect of the arrival cost and help to focus only on the effect of the noises in the estimation accuracy.

Figure 7-2 presents the estimates provided by both MHE and EMHE in open-loop. As shown in this figure, the proposed EMHE was able to reduce the estimation error significantly when compared to the standard MHE. The MSE for each state is reported in Table 7-2. According to this table, the estimation error in the case of using the EMHE scheme is reduced by two orders of magnitude than that reported for the case of using standard MHE. This improvement in the state estimation offered by EMHE is because the Gaussian mixture models of the non-zero mean non-Gaussian noises (and uncertainties) have been considered in the EMHE objective function; therefore, the mismatch between the process and the estimation scheme is smaller than that observed in the standard MHE. The current case study assumed that the distribution of the process uncertainties and measurement noises were time-independent, i.e., the corresponding Gaussian mixture models remain unchanged throughout the

operation. As a result, the EMHE scheme only requires performing the EM algorithm once (i.e., at the beginning of the estimation), which can be performed offline to avoid any additional CPU time in the estimation scheme. The improvement in the estimation accuracy achieved by performing the proposed EMHE is attractive as it comes at no additional computational costs.

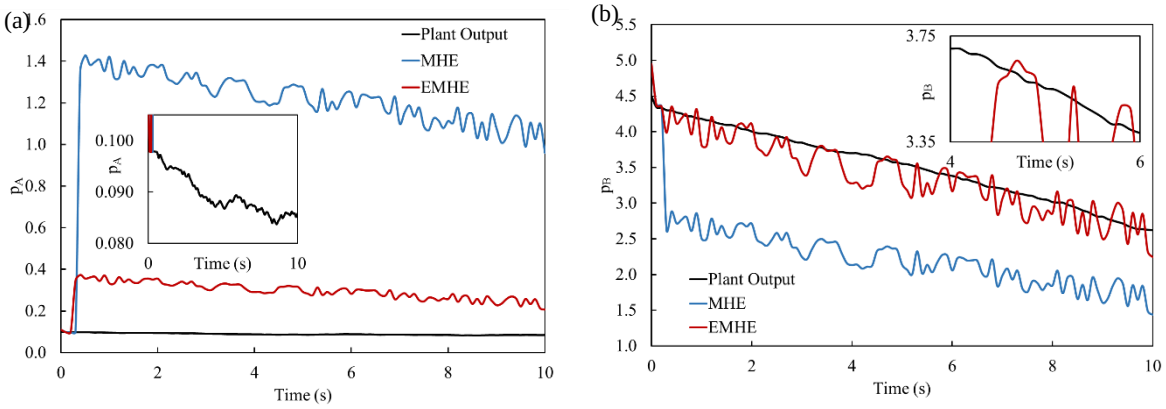


Figure 7-2: Estimation provided by standard MHE and EMHE under open-loop operation: (a) p_A; (b) p_B

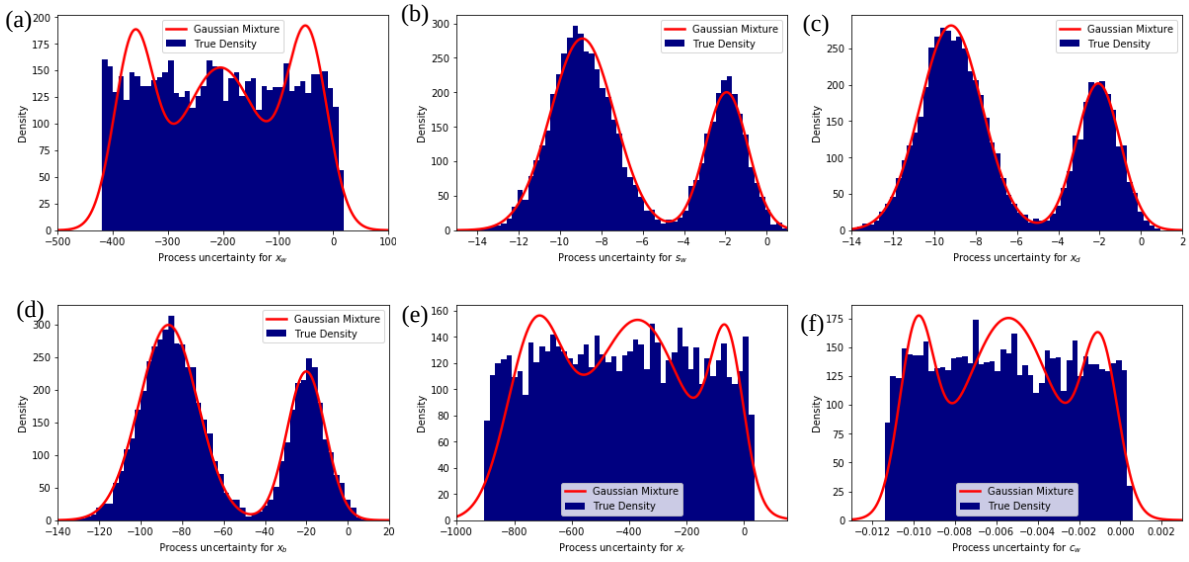
Table 7-2: MSE for p_A and p_B using different estimation schemes under open-loop operation

Estimation method (n)	$MSE_{p_A}^{(n)}$	$MSE_{p_B}^{(n)}$
MHE	1.26	1.68
EMHE	0.04	0.05

7.3.2 Wastewater Treatment Plant (WTP)

The second case study considers an industrial-scale wastewater treatment plant (WTP) presented in Equations (4-29),(6-6),(6-7). Ass mentioned in section 4.2.4, WTPs are often exposed to the external perturbations that may affect the operability of this process[192]. Moreover, it is difficult to develop a mechanistic model that can completely capture the dynamics[192]. Thus, WTPs are often associated with process uncertainties and is essential to approximate these uncertainties properly and introduce to the estimation scheme. Moreover, due to the complexities mentioned above, the distribution of the process uncertainties and consequently the distribution of the states are more likely to follow a non-Gaussian behaviour. Hence, WTP is a good candidate for the goal of assessing the performance of the proposed estimation scheme. Similar to the assumptions mentioned in sections 4.2.4 and 6.2.1, online

measurements are available for x_d, s_w, and c_w whereas x_w, x_b, and x_r represent the unknown states for this plant. The sampling time of this process is 1 hr. Figure 7-3 presents the distributions of the process uncertainty and measurement noises considered for this process. As shown in this figure, the process uncertainties associated with states x_w, x_r, and c_w, as well as the corresponding measurement noises to states s_w and x_d are assumed to follow uniform distributions; whereas bimodal densities were considered to represent the remaining process uncertainty and measurement noise variables present in the system. As mentioned above, these assumptions were made to test the performance of the proposed estimation scheme under practical scenarios involving noises and uncertainties with arbitrary non-Gaussian densities, which are expected to be observed in practice. These distributions were assumed to remain constant during operation, i.e., time-independent non-zero mean non-Gaussian noises. The histogram represents the non-Gaussian random variables whereas the red solid lines are the Gaussian mixture model for each distribution. Note that the number of Gaussian components varies for each distribution based on the level of non-Gaussianity. Also note that all the process uncertainties and measurement noises are mutually independent. Thus, each process uncertainty (and each measurement noise) follows a unique non-Gaussian distribution that is different from the distribution of the other uncertainties and noises. Univariate EM has been performed to approximate the parameters of the corresponding Gaussian mixture models of these process uncertainties and measurement noises.



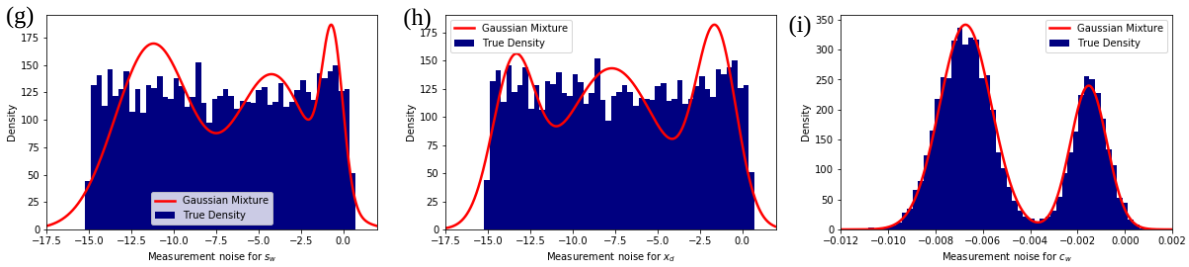


Figure 7-3: Histogram for the true non-Gaussian distribution and the Gaussian mixture approximation of the process uncertainties and the measurement noises in WTP

As in the previous case study, preliminary tests were performed to select the specifications of the noises and uncertainties in the favour of the standard MHE. The standard deviation of the process uncertainty associated with each state is set to 5% of the nominal steady-state value of the corresponding state. Similarly, the standard deviation for the measurement noises associated with each of the states is set to 5% of their nominal steady-state values. The steady-state information used as the initial condition for WTP is presented in Table B-2 in Appendix B. The Gaussian mixture models presented in Figure 7-3 represent the distribution of measurement noises and process uncertainties in EMHE framework. One of the main interests in state estimation is to provide accurate estimates of the states in closed-loop for the purpose of online control. Motivated by this, both open-loop and closed-loop scenarios have been considered for WTP.

7.3.2.1 *Scenario I:* WTP in open-loop

The preliminary tests on the standard MHE for WTP featuring zero-mean Gaussian measurement noises and process uncertainties have shown that 10hr (i.e., 10 time intervals) is an appropriate choice for the length of estimation horizon to obtain an appropriate estimation for WTP without considering the arrival cost term. Both state estimation schemes consider that the initial guesses for the estimated states are 90% of the actual initial conditions of the state variables of WTP. Figure 7-4 shows the estimations provided by EMHE and MHE for the unknown states x_w, x_b, and x_r under the non-Gaussian random variables presented in Figure 7-3. As shown in Figure 7-4, the estimation improves significantly when EMHE is used as the estimation scheme in open-loop. According to Table 7-3, the relative estimation error, i.e., $MSE_{x_w}^{(\mathrm{E}MHE)}/MSE_{x_w}^{(MHE)}$, is reduced by 20% for x_w in the case of using EMHE. Similarly, the relative estimation error for x_r, i.e., $MSE_{x_r}^{(\mathrm{E}MHE)}/MSE_{x_r}^{(MHE)}$, is decreased by 32% when EMHE is

used as the estimation scheme. Note that the estimation error for online measured states, i.e., s_w, x_d, and c_w, are almost the same using both estimation schemes.

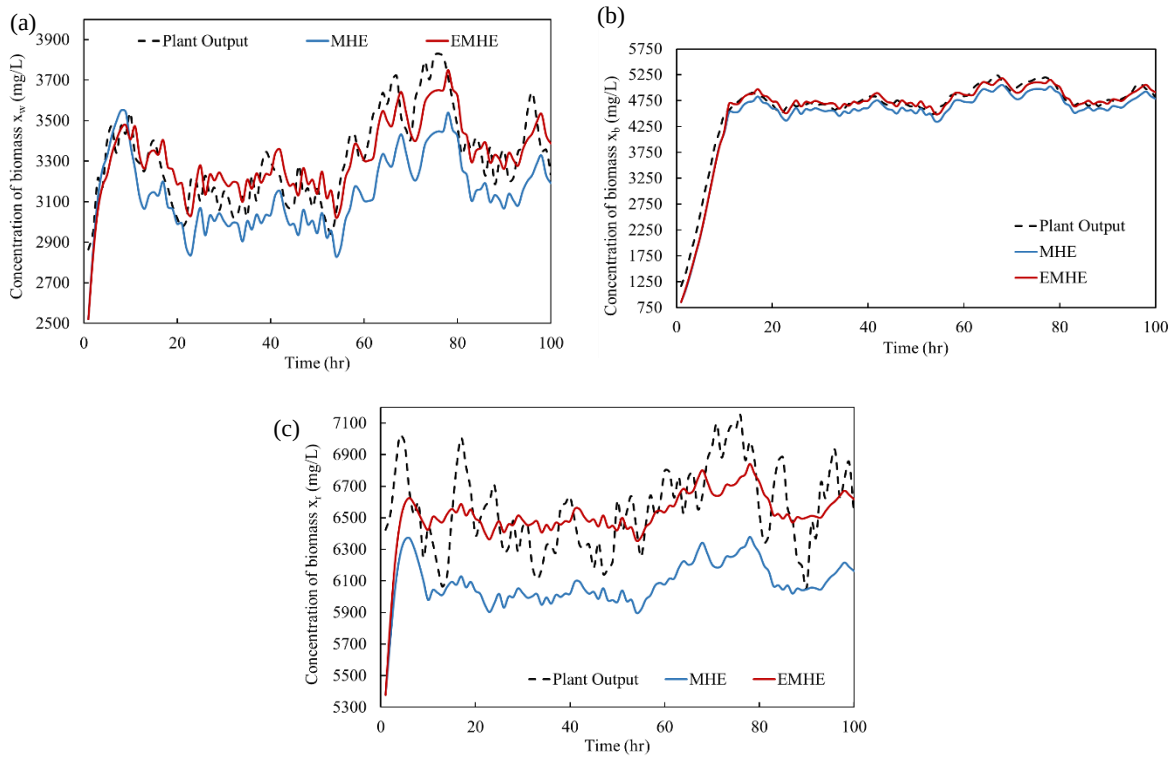


Figure 7-4: Estimation provided by standard MHE and EMHE for (a) x_w; (b) x_b; (b) x_r

Table 7-3: MSE for x_w, x_b ,and x_r using different estimation schemes

Estimation method (n)	$MSE_{x_w}^{(n)}$	$MSE_{x_b}^{(n)}$	$MSE_{x_r}^{(n)}$
MHE	4.66e4	3.39e4	2.61e5
EMHE	1.50e4	1.78e4	5.22e4

7.3.2.2 *Scenario II:* WTP in closed-loop using NMPC

NMPC has been engaged to this work to show the performance of the WTP in closed-loop using both EMHE and MHE. That is, the closed-loop framework shown in Figure 6-1 is used in the current scenario. As shown in this figure, an estimation scheme is necessary to provide the initial conditions to the NMPC problem at each time interval. Moreover, NMPC provides the optimal control actions to both process plant and the state estimator at each time interval. A detailed description on this closed-

loop framework can be found in section 6.1. The general formulation of NMPC is presented in Equation (6-3). In addition to the assumptions mentioned above for this case study, the prediction and control horizons in the NMPC framework are set to 10 hr and 5 hr, respectively. These assumptions have been previously tested to ensure acceptable NMPC performance for the WTP. As mentioned in the previous chapters, the key states for the purpose of controlling the system are the concentration of biomass, substrate and oxygen inside the bioreactor, i.e., x_w, s_w, and c_w, respectively; the weight of these controlled variables are set to 1, 200, 1, in the NMPC framework, respectively. The manipulated variables are the turbine speed (f_k), the flowrate of the decanter outlet (q_2) and the purge flowrate (q_p). These flowrates are mainly responsible to control the biomass and substrate concentrations, whereas f_k is used to keep the oxygen concentration on target. Bounds on the manipulated variables are presented in Table B-2 in appendix B. Figure 7-5 highlights the results obtained by performing estimation and control using both frameworks. Moreover, Table 7-4 represents the estimation error, i.e., the error between the plant output and the estimated states evaluated based on Equation (4-22), as well as the control error, i.e., the error between the plant output and the set-point calculated based on Equation (7-14), for the controlled variables x_w and s_w. According to Figure 7-5(a), the closed-loop performance for the controlled variable s_w is the same from both schemes, i.e., using MHE and EMHE as the estimation scheme in the closed-loop framework. That is, the error between the plant output and the set-point for s_w, i.e., Equation (7-14), is the same with a tolerance of 0.4% as reported in Table 7-4. Figure 7-5(b) presents the results for biomass concentration inside the bioreactor (x_w), which is one of the most critical states for this system for which online measurements are not available. The estimation error, i.e., the error between the estimation and the plant output for x_w, , resulted in the case of using EMHE is approximately 70% of the estimation error when using the standard MHE (see Table 7-4). According to Table 7-4, the relative error in the closed-loop control, i.e., the error between the plant output and the set-point for x_w, is approximately 59% smaller when EMHE was used in the feedback control framework. The results for remaining states have not been presented for brevity. As shown in Figure 7-5(c)-(d), the manipulated variables q_2 and q_p reach their saturation limits (i.e., bounds) at specific time intervals, regardless of the estimation scheme. During these time intervals, NMPC is computing the exact same control actions to the plant, regardless of the estimation method (i.e., standard MHE or EMHE); this leads to observe similar plant outputs from both methods. Nevertheless, the plant output is closer to its set-point in the case of using EMHE when the constraints on the manipulated variables are not active. This observation holds even though the estimates provided by standard MHE are closer to the set-point than the estimates provided by EMHE. This is because the main connection

in closed-loop is between the state estimation scheme and NMPC, as shown in Figure 6-1 (see section 6.1). Therefore, NMPC focuses on finding the optimal control actions that keep the estimation of the controlled variables at their target. These control actions are sent to the plant. With this in mind, the smaller error between the estimations and the set-points does not necessary lead to better control of the plant as the estimations may not be accurate representatives of the plant outputs. Thus, the errors between the estimations and both the plant output and the set-points need to be reduced to improve the closed-loop operation. As shown in both *Scenario I* and *Scenario II*, the proposed EMHE framework provides more accurate estimations of the plant outputs than those provided by standard MHE, which consequently leads to better set-point tracking performance.

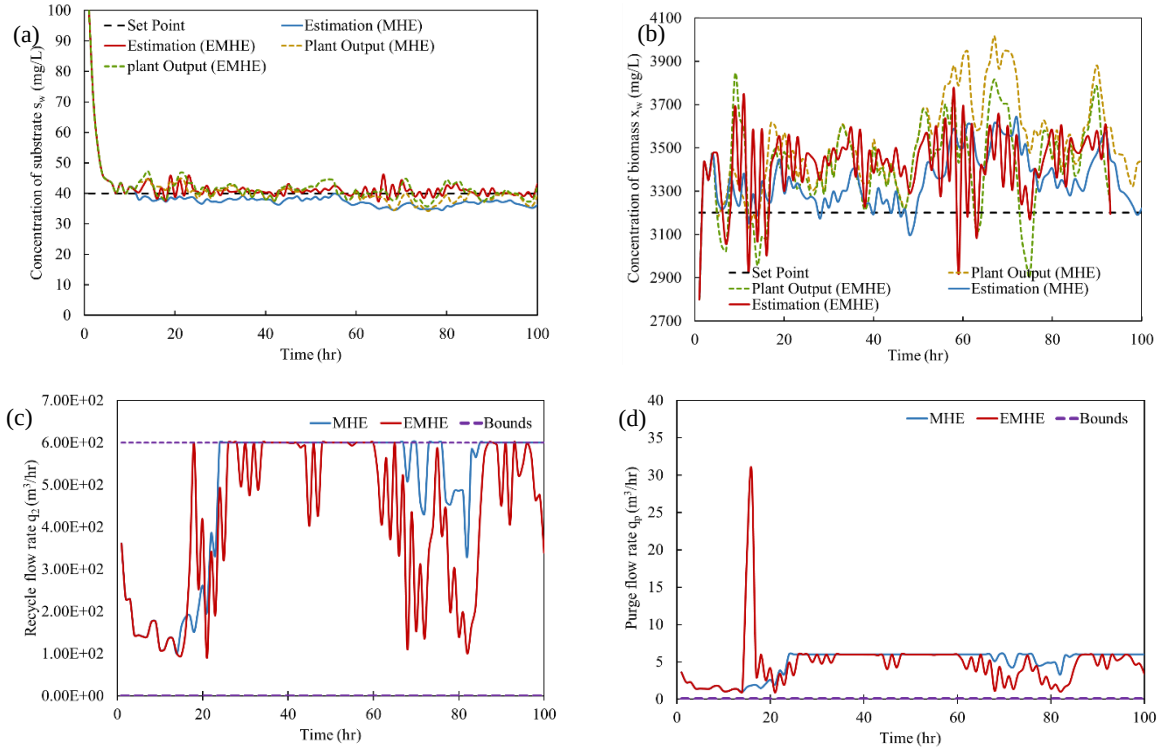


Figure 7-5: Impact of using MHE and EMHE in closed-loop: (a) estimation and plant output for s_w; (b) estimation and plant output for x_w; (c) control actions for q_2; (d) control actions for q_p

Table 7-4: MSE in the estimation and control for x_w and s_w using different estimation schemes under closed-loop

Estimation method (n)	Estimation error (Equation (4-22))		Control error (Equation (7-14))	
	$MSE_{x_w}^{(n)}$	$MSE_{s_w}^{(n)}$	$MSE_{x_w}^{(n)}$	$MSE_{s_w}^{(n)}$

MHE	4.31e+4	9.87	1.43e+5	53.50
EMHE	3.02e+4	5.38	8.47e+4	53.76

7.3.3 A series of Continuous-stirred Tank Reactors (CSTR)

The EMHE scheme proposed in section 7.2 is applied on a chemical process involving two connected continuous-stirred tank reactors (CSRT) in series. This series of CSTRs has been used in the literature for robust MHE and control applications[140,195,196]. Figure 7-6 shows a flowsheet of this process. The symbols F,T, and C denotes the flowrate, temperature, and concentration of the corresponding stream in Figure 7-6, respectively. Three parallel reactions take place at each CSTR, i.e., $A \xrightarrow{k_{10}} B$, $A \xrightarrow{k_{20}} C$ and $A \xrightarrow{k_{30}} D$. The goal is to produce C while B and D are by-products. The unreacted A at the outlet stream of the second CSTR is partially recycled to the first CSTR at a flowrate F_r, temperature of T_2 and with the concentration of C_{A_2}. Each CSTR is equipped with a jacket to control the temperature of the reactor by adding (removing) heat to (from) the reactor.

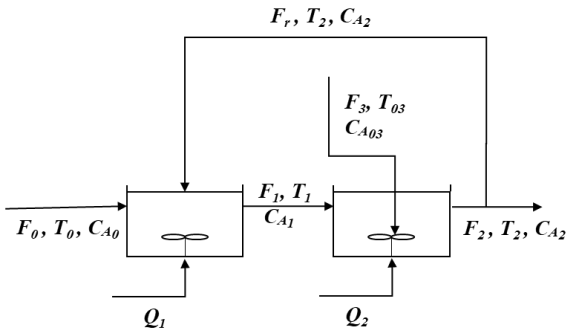


Figure 7-6: Series of CSTRs flowsheet

The states of the system are the temperature and concentration of species A inside each CSTR, i.e., T_1, C_{A_1}, T_2, and C_{A_2}. The process for this system can be modelled as follows[140,195,196]:

$$\frac{dT_1}{dt} = \frac{F_0}{V_1}(T_0 - T_1) + \frac{F_r}{V_1}(T_2 - T_1) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_1}} C_{A_1} + \frac{Q_1}{\rho c_p V_1}$$

$$\frac{dC_{A_1}}{dt} = \frac{F_0}{V_1}\left(C_{A_0} - C_{A_1}\right) + \frac{F_r}{V_1}\left(C_{A_2} - C_{A_1}\right) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_1}} C_{A_1} \quad (7\text{-}16)$$

$$\frac{dT_2}{dt} = \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_3}{V_2}(T_{03} - T_2) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_2}} C_{A_2} + \frac{Q_2}{\rho c_p V_2}$$

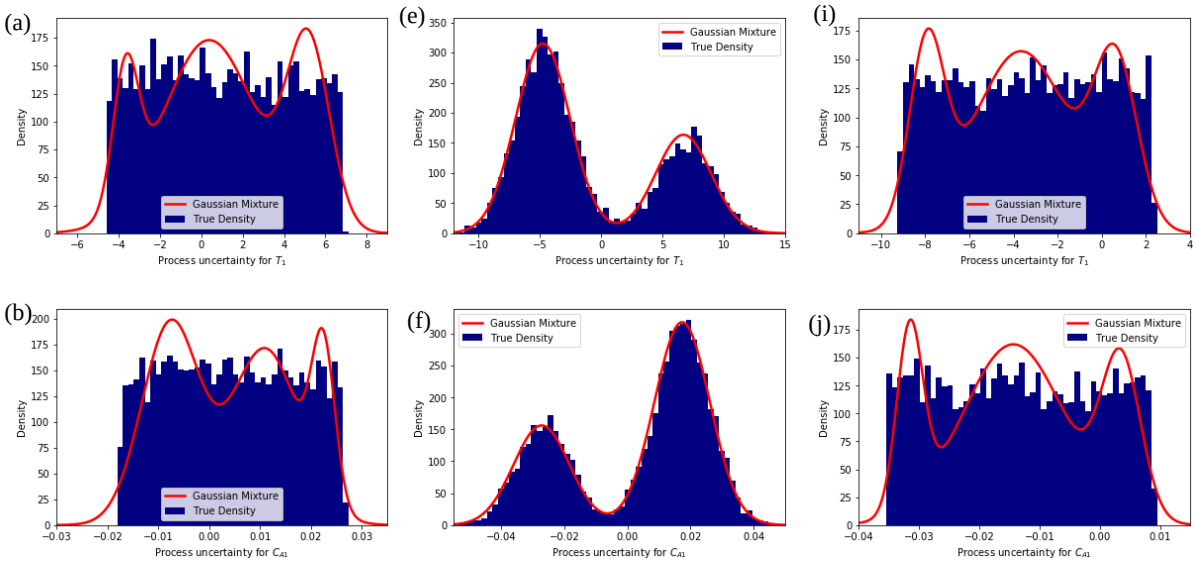
$$\frac{dC_{A_2}}{dt} = \frac{F_1}{V_2}\left(C_{A_1} - C_{A_2}\right) + \frac{F_3}{V_2}\left(C_{A_{03}} - C_{A_2}\right) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_2}} C_{A_2}$$

$$x_0 = [T_1^s, C_{A_1}^s, T_2^s, C_{A_2}^s]^T = [457.943\ \text{K}, 1.770\ \text{kmol}/m^3, 415.459\ K, 1.752\ \text{kmol}/m^3]^T$$

The model parameters for this system have been previously reported[140,195,196]. Note that the superscript s in Equation (7-16) denotes the steady-state value of the corresponding state. The process has three steady-states and only one of these steady-states is unstable. That is, using the unstable steady-state as the initial condition in open-loop also leads to instability in the state estimation, which is the interest of this work (x_0 in Equation (7-16)). The current study performed the state estimation using both EMHE and standard MHE in closed-loop using the framework presented in Figure 6-1 (section 6.1). Online measurements are assumed to be available for temperatures T_1 and T_2. The time interval considered for this case study is 0.1 hr. All four state variables are considered controlled variables in the NMPC framework. To simplify the analysis, an equal weight set to the unity was used in the NMPC framework for all the controlled variables. The manipulated variables are the inlet concentration of A entering the CSTRs, i.e., C_{A_0} and $C_{A_{03}}$, and the heat inputs, i.e., Q_1 and Q_2. A ramp change has been considered in the set-points of each of the controlled variables C_{A_1} and C_{A_2}, which are expected to move the operation of this process from 1.75 kmol/m^3 to 1.0 kmol/m^3 for both C_{A_1} and C_{A_2}. In this present case study, the process uncertainties are assumed to follow time-dependent non-Gaussian distributions. That is, the distributions of the process noises are initially assumed to follow a uniform distribution as shown in Figure 7-7 (a)-(d). The characteristics of these distributions are assumed to change during the transition time (i.e., the ramp change in the set-point of C_{A_1} and C_{A_2}). The bimodal distributions shown in Figure 7-7(e)-(h) represent the distribution of the process uncertainties associated with the states of the system during the transition time. Once the system reaches a new steady-state operating condition, the process uncertainties are assumed to follow a new uniform distribution that is different from their initial densities, as shown in Figure 7-7(i)-(l). Note that the red solid lines in Figure 7-7 denote the corresponding Gaussian mixture model to each non-Gaussian distribution, which have been identified using the EM method. Note that the Gaussian mixture models for the initial distributions of the random noises/uncertainties (shown in Figures 7-7(a)-(d)) were evaluated offline; whereas the re-approximations of the Gaussian mixture models presented in Figures 7-7(e)-(h) and Figures 7-7(i)-(l) were performed online upon the set-point changes during the process operation. That is, the proposed

estimation scheme performed the EM-algorithm twice to update the characteristics of the noises and uncertainties distributions online. The averaged CPU time required to execute the EM algorithm for the non-Gaussian distributions shown in Figures 7-7(e)-(h) and Figures 7-7(i)-(l) is 0.14 s and 0.71 s, respectively. Thus, the additional CPU time required in the proposed estimation scheme to provide the point estimates is relatively low, i.e., the CPU time for this scenario increased by 0.0043s on average per sampling interval when compared to the standard MHE. This result justifies the selection of the EM algorithm in the proposed framework since the additional CPU time needed to update the Gaussian mixture models was not significant. Note however that the CPU time required to perform EM algorithm increases by increasing the number of Gaussian components in the mixture. As indicated above, computer parallelization techniques can be implemented to reduce the computational costs associated with this calculation.

To simplify the analysis, the measurement noises are assumed to be randomly chosen from time-independent zero-mean Gaussian distributions. The standard deviation of the measurement noise associated with each measurement is set to 0.5% of the true steady-state value of the corresponding state.



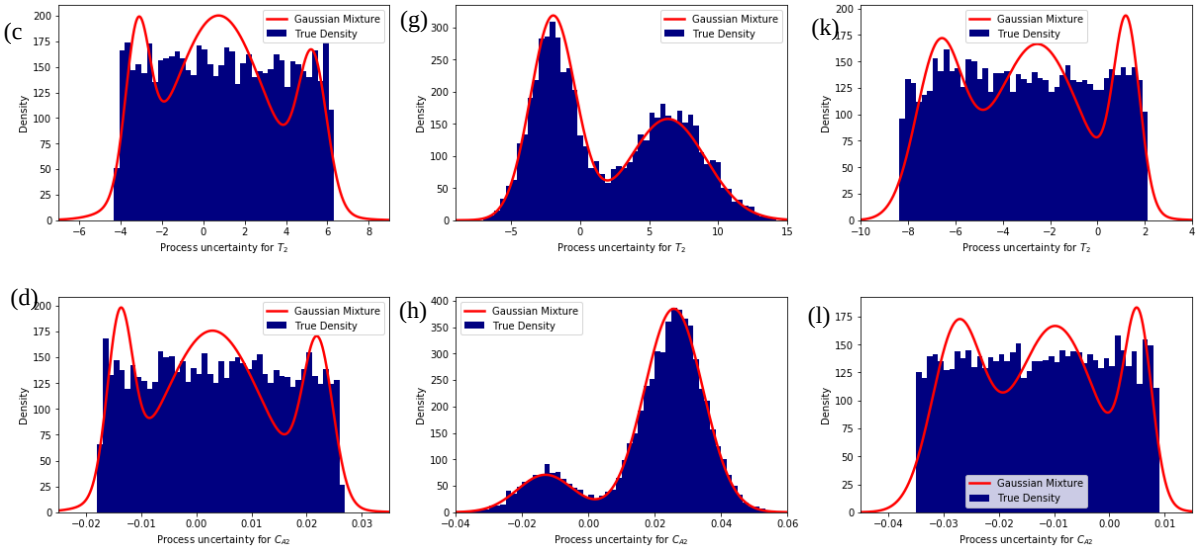


Figure 7-7: Histogram for the true non-Gaussian distribution and the Gaussian mixture approximation of the process uncertainties in CSTRs for (a)-(d) before the ramp; (e)-(h) during the ramp; (i)-(l) after the ramp

Note that in the standard MHE scheme, the standard deviation for the process uncertainty associated with each state is set to 1% of the nominal steady-state value of the corresponding states. The closed-loop framework shown in Figure 6-1 is implemented for the series of CSTRs. Figure 7-8 compares the estimations and plant outputs obtained by performing both standard MHE and the proposed EMHE scheme whereas Figure 7-9 illustrates the control actions obtained from the closed-loop operation. As shown in Figure 7-8, the estimation provided by standard MHE fluctuates around its set-point. A similar behaviour can be observed in the control actions shown in Figure 7-9 for the case of using standard MHE in the closed-loop framework. The fluctuations in the estimations resulted in an inappropriate initialization of NMPC problem at each time interval. Moreover, noisy control actions are sent to the process, which deviates the process plant from their set-points. Hence, NMPC was not capable of keeping the controlled variables at their targets under the standard MHE framework. On the other hand, the proposed EMHE scheme is capable of adopting the non-Gaussian distributions of the process uncertainties throughout the operation of the process, which resulted in a significant improvement in the state estimation. The proposed EMHE scheme was not only able to capture the non-Gaussianity of the process uncertainties adequately, but it was also able to capture the changes in these distributions during the closed-loop operation. Obviously, the acceptable estimates provided by EMHE led to a proper initialization of NMPC and therefore an acceptable operation of this process in closed-loop. Note that large peaks observed in the estimates provided by MHE between the time interval 15 and 20

are due to the random noises observed in the plant. Several tests were conducted to confirm the latter and are not shown here for brevity.

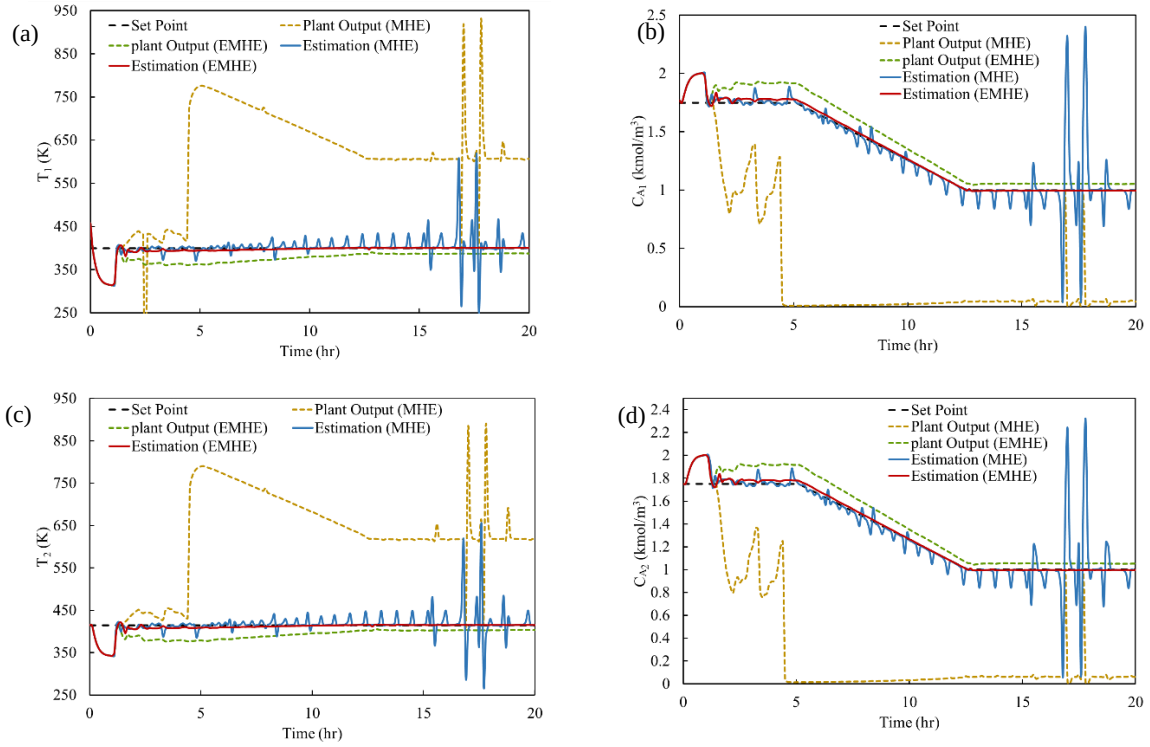
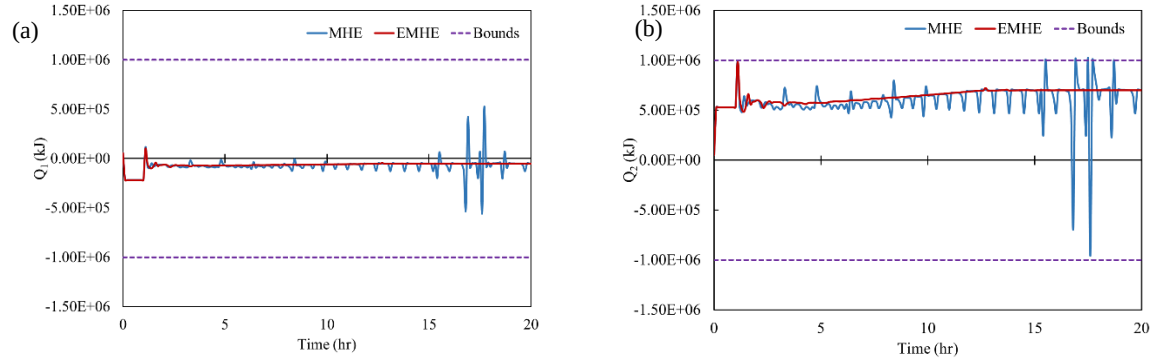


Figure 7-8: Estimation and plant output obtained using MHE and EMHE in the closed-loop for (a) T_1; (b) C_{A_1};(c) T_2; (d) C_{A_2}



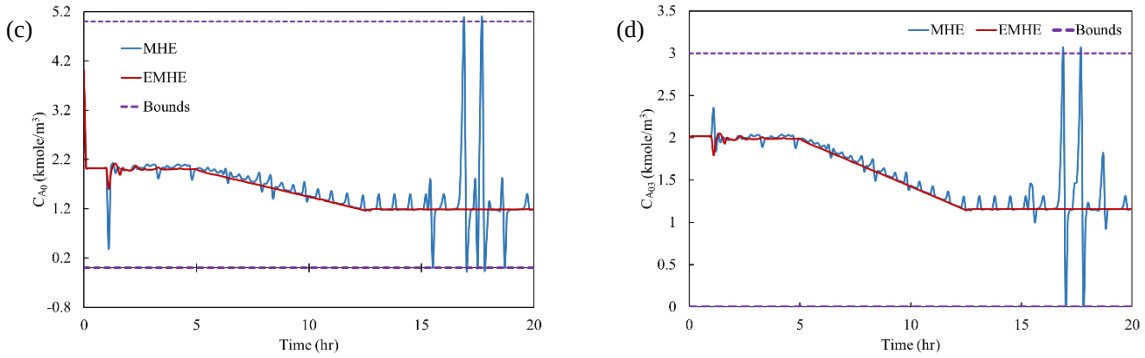


Figure 7-9: Control actions obtained by considering MHE and EMHE in the closed-loop for (a) Q_1; (b) Q_2;(c) C_{A_0}; (d) $C_{A_{03}}$

7.4 Summary

This study presents a novel modification to the conventional MHE framework, namely the extended MHE (EMHE) scheme, for applications where the process uncertainties and measurement noises follow non-zero mean non-Gaussian distributions that may change online due to issues with measurement devices or changes in the operating conditions, e.g., seasonal changes or switching to another grade of the product. In the proposed EMHE framework, the non-Gaussian densities are approximated by Gaussian mixture models, which are used to specify the EMHE objective function. As a result, EMHE can take into account a proper approximation of the process uncertainties and measurement noises thereby improving the accuracy in the estimation when compared to standard MHE. Moreover, the proposed EMHE scheme provides the opportunity to adopt online possible changes in the distribution of the process uncertainties and measurement noises, i.e., time-dependent distributions. EMHE was developed neglecting the zero-mean Gaussian distribution assumption for the noises/uncertainties considered in standard MHE formulation derivation. Thus, EMHE is a more comprehensive version of the standard MHE. EMHE improves the state estimation by considering multiple Gaussian mixture models aimed at providing an accurate representation of nonlinear probability distribution functions for process uncertainty and measurement noises. Moreover, the approximation of the non-Gaussian distributions to Gaussian mixture models are performed offline and therefore, the optimization problem in EMHE does not increase significantly the computational costs than those required by the standard MHE formulation.

The current chapter assumed that the changes in the distribution of the measurement noises and process uncertainties are scheduled a priori and thereby, the distribution of the noises/uncertainties are known

a priori when performing the state estimation. The next chapter presents a novel development on the MHE framework that relaxes this assumption, i.e., the distribution of these random variables is unknown due to the unscheduled changes in the plant operation.

Chapter 8
Robust Moving Horizon Estimation (RMHE)

The MHE-based estimation schemes often assume that the distribution of the process uncertainties and measurement noises are known *a priori* and can be described by zero-mean Gaussian densities, which may not be a valid assumption for real-world applications. In scenarios featuring scheduled operation changes, the EMHE scheme introduced in Chapter 7 can be used to capture the known non-zero mean non-Gaussian distribution of the random variables present in the process. That is, EMHE performs an external method such as Expectation-Maximization (EM) algorithm outside of the EMHE formulation to evaluate the Gaussian mixture models of the known distributions of the process uncertainties and measurement noises. Thus, EMHE does not consider additional constraints or decision variables than those used in the standard MHE formulation. However, in the case that these changes happen suddenly in the plant (e.g., drastic disturbances entering the plant, equipment malfunction or failures in plant devices), the distributions of the non-Gaussian measurement noises or process uncertainties are unknown to the estimation scheme thus losing the ability to have access to accurate estimations of the plant states during this critical point in the operation. Hence, a robust estimation scheme is required to approximate these unknown non-Gaussian distributions appropriately in favor of a higher estimation accuracy. Moreover, to the author's knowledge, studies considering a robust MHE-based estimation scheme for applications involving process uncertainties or measurement noises that their distributions are non-Gaussian and unknown to the state estimation framework are absent from the literature (see section 2.3.2).

In the present chapter, a novel robust MHE framework, i.e., referred to as *Robust Moving Horizon Estimation* (RMHE), is introduced to improve the state estimation when the process uncertainties or measurement noises follow unknown non-Gaussian distributions, e.g., when the changes in the plant operation are not schedule in advance. The proposed RMHE uses the EMHE developed in chapter 7 as basis to capture possible non-zero mean measurement noises or process uncertainties, which requires the same computational costs as in the standard MHE. The RMHE framework includes the Gaussian mixture models explicitly to the optimization problem to find the optimal estimated states as well as optimal approximation of the unknown distributions of the measurement noises (or process uncertainties). In the RMHE scheme, the unexpected (i.e., unknown) random process uncertainties or the measurement noises are modelled using Gaussian mixture models, which would consist of two types of Gaussian components. During normal operation, all the expected random noises and process

uncertainties are considered to follow a Gaussian component in the mixture, which is the same as the known Gaussian distributions considered in the standard MHE. For the remainder of this study, this Gaussian component is referred to as the *core-Gaussian component*. Alternatively, the unexpected (unknown) random noises or process uncertainties that would most likely not follow the core-Gaussian component during sudden changes in the operation are assumed to follow one of the remaining Gaussian components in the mixture; these are referred to as *perimeter-Gaussian components*. The mean value and the covariance matrix of each perimeter-Gaussian component are considered as the tuning parameters in the RMHE framework. In addition, the weights assigned to each Gaussian component are also unknown variables estimated online by the proposed RMHE scheme based on the unexpected random noises or uncertainties taking place in the process at the current estimation horizon.

As discussed above, RMHE deals with applications where the Gaussian mixture models of the process uncertainties (or measurement noises) are unknown to the estimation scheme *a priori*. To overcome this challenge, the proposed RMHE considers additional constraints and decision variables than those considered in the EMHE and standard MHE formulations. These additional constraints and decision variables are considered to model these unknown Gaussian mixture models explicitly in the RMHE formulation. Hence, the RMHE framework does not use any external method such as EM algorithm as in EMHE to evaluate the Gaussian mixture models, i.e., the evaluation of the optimal Gaussian mixture models is performed online and within the RMHE framework. The proposed RMHE requires somewhat the same computational costs as in the standard MHE thus making this estimation scheme attractive for large-scale applications.

This chapter is organized as follows: Section 8.1 introduces the proposed RMHE framework. Section 8.2 presents the computational experiments conducted to this work considering two chemical engineering processes featuring unexpected measurement noises or process uncertainties. A summary of this chapter is presented at the end.

8.1 Robust Moving Horizon Estimation (RMHE) Framework

This section describes the proposed robust moving horizon estimation (RMHE) formulation using the expressions and concepts presented in the previous section. The proposed RHME aims to improve the state estimation for those applications affected by unexpected/unplanned process uncertainties or measurement noises. As mentioned earlier, RMHE aims to model the unknown distribution of each process uncertainty or measurement noise variable with a general Gaussian mixture model consisting

of a finite number of Gaussian components (ngr) involving core and perimeter components. A core-Gaussian component represents the distributions that measurement noises (or process uncertainties) are expected to follow under the normal plant operation, i.e., the core-Gaussian component has the same mean-value and covariance matrix as the known Gaussian distribution considered in the standard MHE. The remaining Gaussian components present in the Gaussian mixture model are referred to as perimeter-Gaussian components and are used to describe the unknown uncertainty or noise distributions during sudden changes in the operation. The main objective of the RMHE formulation is to find the optimal state estimates and Gaussian mixture models that minimize the errors in the estimation due to process uncertainties and measurement noises over a user-defined estimation horizon (N). The proposed RMHE formulation is developed based on the EMHE (see Equation (7-1)) and the main characteristics of the Gaussian mixture model of the noises or uncertainties described in Equations (4-11)-(4-13). The RMHE formulation is as follows:

$$\min_{\substack{\{\boldsymbol{x}_j,\boldsymbol{w}_j\}_{j=k-N}^{k-1},\\ \{\boldsymbol{\alpha m}_r,\boldsymbol{\alpha p}_r\}_{r=1}^{ngr}}} \sum_{j=k-N}^{k-1} \left\|\boldsymbol{w}_j - \boldsymbol{\mu}_k^{\boldsymbol{GM}}\right\|^2_{\boldsymbol{Q}_k^{\boldsymbol{GM}^{-1}}} + \sum_{j=k-N+1}^{k} \left\|\boldsymbol{v}_j - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right\|^2_{\boldsymbol{R}_k^{\boldsymbol{GM}^{-1}}} + \varphi_{k-N} \tag{8-1}$$

s.t. (8-2)

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

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

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

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

$$\boldsymbol{\mu}_k^{\boldsymbol{GM}} = \begin{cases} \sum_{r=1}^{ngr} \boldsymbol{\alpha p}_k^r \odot \boldsymbol{\mu}_k^r, & \Upsilon_{pu} = 1 \\ \boldsymbol{\mu}_k^1, & \Upsilon_{pu} = 0 \end{cases} \tag{8-3}$$

$$\boldsymbol{Q}_k^{\boldsymbol{GM}} = \begin{cases} \sum_{r=1}^{ngr} \boldsymbol{\alpha p}_k^r \odot \boldsymbol{Q}_k^r + \sum_{r=1}^{ngr} \boldsymbol{\alpha p}_k^r \odot (\boldsymbol{\mu}_k^r - \boldsymbol{\mu}_k^{\boldsymbol{GM}})(\boldsymbol{\mu}_k^r - \boldsymbol{\mu}_k^{\boldsymbol{GM}})^{\boldsymbol{T}}, & \Upsilon_{pu} = 1 \\ \boldsymbol{Q}_k^1, & \Upsilon_{pu} = 0 \end{cases}$$

$$\boldsymbol{\alpha p}_k^1 = \boldsymbol{1}, \quad \Upsilon_{pu} = 0;$$

$$\sum_{r=1}^{ngr} \boldsymbol{\alpha p}_k^r = \boldsymbol{1}; \quad \forall \Upsilon_{pu};$$

$$\boldsymbol{0} \le \boldsymbol{\alpha p}_k^r \le \boldsymbol{1}; \quad \forall \Upsilon_{pu}; \forall r = 1,2, \dots, ngr$$

$$\boldsymbol{\tau}_k^{\boldsymbol{GM}} = \begin{cases} \sum_{r=1}^{ngr} \boldsymbol{\alpha m}_k^r \odot \boldsymbol{\tau}_k^r, & \Upsilon_{mn} = 1 \\ \boldsymbol{\tau}_k^1, & \Upsilon_{mn} = 0 \end{cases} \tag{8-4}$$

$$\boldsymbol{R}_k^{\boldsymbol{GM}} = \begin{cases} \sum_{r=1}^{ngr} \boldsymbol{\alpha m}_k^r \odot \boldsymbol{R}_k^r + \sum_{r=1}^{ngr} \boldsymbol{\alpha m}_k^r \odot \left(\boldsymbol{\tau}_k^r - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)\left(\boldsymbol{\tau}_k^r - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)^{\boldsymbol{T}}, & \Upsilon_{mn} = 1 \\ \boldsymbol{R}_k^1, & \Upsilon_{mn} = 0 \end{cases}$$

$$\boldsymbol{\alpha m}_k^1 = \mathbf{1}, \quad \Upsilon_{mn} = 0;$$

$$\textstyle\sum_{r=1}^{ngr} \boldsymbol{\alpha m}_k^r = \mathbf{1}; \quad \forall \Upsilon_{mn};$$

$$\mathbf{0} \leq \boldsymbol{\alpha m}_k^r \leq \mathbf{1}; \quad \forall \Upsilon_{mn}; \forall r = 1,2, \dots, ngr$$

where:

$$\boldsymbol{\mu}_k^{\boldsymbol{GM}}, \boldsymbol{\mu}_k^r, \boldsymbol{\alpha p}_k^r \in \mathbb{R}^{n_x}, \boldsymbol{\tau}_k^{\boldsymbol{GM}}, \boldsymbol{\tau}_k^r, \boldsymbol{\alpha m}_k^r \in \mathbb{R}^{n_y}; \boldsymbol{Q}_k^{\boldsymbol{GM}}, \boldsymbol{Q}_k^r \in \mathbb{R}^{n_x \times n_x}, \boldsymbol{R}_k^{\boldsymbol{GM}}, \boldsymbol{R}_k^r \in \mathbb{R}^{n_y \times n_y}$$

where r denotes the index of the Gaussian component in the Gaussian mixture model. As shown in Equations (8-1)-(8-4), in addition to the known core-Gaussian distribution that the process uncertainties and noises are expected to follow during the normal plant operation, i.e., $\mathcal{N}(\boldsymbol{\mu}^1, \boldsymbol{Q}^1)$ and $\mathcal{N}(\boldsymbol{\tau}^1, \boldsymbol{R}^1)$, the proposed formulation consider additional perimeter-Gaussian distributions that can properly capture the unexpected random uncertainties or noises, i.e., $\mathcal{N}(\boldsymbol{\mu}^r, \boldsymbol{Q}^r)$ or $\mathcal{N}(\boldsymbol{\tau}^r, \boldsymbol{R}^r)$ where $r \in \{2,3, \dots ngr\}$. Moreover, the chances that the random process uncertainties (or measurement noises) present in the system follow the core-Gaussian distribution is given by $\boldsymbol{\alpha p}^1$ (or $\boldsymbol{\alpha m}^1$**)**, respectively. If the system works under normal plant operation, i.e., with random noises and uncertainties that follow the core-Gaussian component, then the RMHE optimization formulation shown in Equation (8-1) would set the weights on the core-Gaussian component to values closer to one while the weights for the perimeter-Gaussian components in the mixture are expected to be close to zero. Note that in the MHE and RMHE schemes, the only explicit connection between plant and the estimation scheme is through the measurement algebraic equation shown in Equation (8-2) (i.e., $\boldsymbol{y}_j = h(\boldsymbol{x}_j, \boldsymbol{u}_j) + \boldsymbol{v}_j$). That is, $\boldsymbol{v}_j$ is the main resource that both schemes rely on to recognize the presence of unexpected noises or uncertainties affecting the operation. Accordingly, this limits the application of the proposed RMHE to cases involving either unexpected measurement noises or unexpected process uncertainties. With this in mind, the present approach assumes that the condition that causes a significant deviation from a normal plant operation is detected before activation of the proposed framework. In the case that this condition is not detected, the RMHE is expected to perform similar to a standard MHE. While the present approach cannot handle both conditions simultaneously, the proposed RMHE framework is flexible and can switch from one unexpected condition to another, e.g., switch from unexpected process uncertainties to unexpected measurement noises and vice versa. To account for this condition, a binary

parameter (Υ_{ϱ}) in Equation (8-5) is considered and used as an input parameter to the RMHE formulation. This parameter restricts the robustness framework to perform estimation under unexpected behaviour in process uncertainties ($\Upsilon_{pu} = 1$ and $\Upsilon_{mn} = 0$) or measurement noises ($\Upsilon_{mn} = 1$ and $\Upsilon_{pu} = 0$), i.e.,

$$\Upsilon_{\varrho} \in \{0,1\}, \forall \varrho \in \{mn, pu\}; \qquad \sum_{\varrho \in \{mn,pu\}} \Upsilon_{\varrho} = 1; \tag{8-5}$$

Note that Υ_{pu} and Υ_{mn} are user-defined parameters that are inputs to the RMHE framework as shown in Equations (8-1)-(8-4). Similarly, the number of Gaussian components present in the Gaussian mixture model (ngr) is a tuning parameter in the RMHE scheme that must be specified *a priori*.

Note that improving the approximation of the unknown non-Gaussian distributions using an adequate ngr as well as the perimeter-Gaussian components' specifications, i.e., $\boldsymbol{\mu}^r, \boldsymbol{Q}^r, \boldsymbol{\tau}^r, \boldsymbol{R}^r$ where $r \in \{2,3, \dots ngr\}$, are expected to improve the estimation accuracy. Nevertheless, a reasonable set of assumptions for these parameters may lead to accurate estimations in the proposed RMHE; those assumptions can be obtained from historical plant data or from process heuristics. Moreover, the present study assumes that $\boldsymbol{\mu}^1 = \boldsymbol{\tau}^1 = \mathbf{0}$, (i.e., zero-mean core-Gaussian components). Nevertheless, the proposed RMHE is expected to perform well for cases involving non-zero mean core-Gaussian components. Moreover, the current study assumes that the process uncertainties or measurement noises associated with each state are uncorrelated. Nonetheless, the present RMHE approach can be easily extended to cases involving multivariate distributions of the noises or uncertainties, i.e., correlated process uncertainties (or measurement noises). For those cases, the RMHE scheme needs to make use of multivariate Gaussian mixture models as a substitute to the univariate Gaussian mixture models shown in Equations (4-11)-(4-13). That is, the vector-valued of weights shown in the univariate of Gaussian mixture model of uncorrelated process uncertainties, i.e., $\boldsymbol{\alpha p}_k^r$ in Equation (8-3), have to be replaced by a scalar αp_k^r to represent the weight of the r^{th} Gaussian component in the corresponding multivariate Gaussian mixture model. Consequently, the vectors **1** and **0** shown in Equation (8-3) also have to be replaced by the scalars 1 and 0, respectively. Similar modifications must be made in Equation (8-4) when correlated measurement noises are considered, i.e., $\boldsymbol{\alpha m}_k^r$ and vectors **1** and **0** need to be replaced by αm_k^r (i.e., a scalar) and 1and 0, respectively, to consider a multi-variate Gaussian mixture model.

Remark 1: Note that the RMHE problem would not reduce to the EMHE problem even when the distributions of both process uncertainties and measurement noises are non-Gaussian and known a priori. Both EMHE and RMHE use Gaussian mixture models to capture non-Gaussianity in the

distributions of the process uncertainties and measurement noises. However, performing EMHE under the known non-Gaussian distributions assumption is likely to offer slightly more accurate estimates than those obtained by the RMHE under the unknown non-Gaussian distributions assumption. This expectation is because EMHE uses the EM algorithm that uses a large number of samples drawn from the non-Gaussian distributions to provide the corresponding Gaussian mixture models; in contrast, RMHE only has access to N (length of estimation horizon) number of samples, i.e., the random process uncertainties (or measurement noises) in the current estimation horizon, to find the optimal Gaussian mixture models of the non-Gaussian process uncertainties (or measurement noises).

Remark 2: In general, under the closed-loop operation, the state estimation schemes are responsible to provide initial conditions to advanced control schemes (e.g., model predictive controller (MPC)). Robust and accurate estimates imply a better initialization of the control scheme thus improving the online monitoring and control of the system. As the standard MPC problem is not able to capture the process uncertainties and measurement noises, the proposed RMHE can be combined with the developments of MPC presented in [197,198] to provide a feed-back control system that is robust against noises and uncertainties. Note that RMHE is expected to improve the overall closed-loop performance and will be explored in future studies.

8.2 Computational experiments

Two case studies arising in the chemical engineering that are often used in the context of state estimation have been considered in this study to investigate the performance of the proposed RMHE when compared to the standard MHE. This work uses the mean squared error (MSE) shown in Equation (4-22) as the metric to quantify the performance of the estimation schemes, i.e., $n \in \{\text{MHE}, \text{RMHE}\}$. The simulations were conducted using Pyomo 5.2 in Python 3.6 on a computer running Microsoft Windows Server 2016 standard. The computer was equipped with 96 GB RAM and Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz 2.10 GHz (2 processors). The backward method has been used to discretize the nonlinear dynamic models of the systems. Two case studies considered in this process involve an actual wastewater treatment plant (WTP) and a network of continuous-stirred tank reactors (CSTRs) and separator. Each of these processes, the assumptions made in the estimation analysis and the results are presented next.

8.2.1 Wastewater Treatment Plant (WTP)

The WTP described by Equation (4-29) have been widely used in the field of state estimation and control[172,173]. As mentioned earlier, WTP processes are often subjected to external perturbations and therefore are difficult to model properly[192]. Developing a mechanistic model that can completely represents the dynamic of the process is challenging[192]. When these highly nonlinear processes are subjected to external perturbations, the process uncertainties and measurement noises present in the system are less likely to follow their expected probability density functions. Consequently, it is quiet challenging to properly model these random variables for WTP. In addition, as shown in section 6.2.1, an inappropriate modelling of the uncertainties and noises may lead to a poor MHE performance for WTP. As RMHE aims to improve the estimation robustness by modelling these unexpected random variables, WTP may benefit from this state estimation approach. Motivated by this, an industrial-scale WTP is considered in this work to implement RMHE for practical scenarios concerning unexpected measurement noises during operation. Similar to the previous chapters, online measurements are available for the biomass concentration in the upper layer of the decanter (x_d), as well as the substrate (s_w) and the dissolved oxygen (c_w) concentrations inside the bioreactor; also, x_w, x_b, and x_r are assumed to be unknown states for this plant. The observability of the system was confirmed by checking the observability matrix to be full rank. The sampling time for this system is set to 1 hr. As the main goal in this work is to show the effect of unexpected noises/uncertainties on the state estimation, the length of estimation horizon (N) is set to a relatively large number ($N = 10$) so that reduces the effect of arrival cost and focus on the effect of the noises and uncertainties on the estimation scheme. The mechanistic model of WTP is initialized using the nominal steady-state condition presented in Table B-2 (Appendix B). All the state estimation schemes are initialized assuming that the states are 5% away from their true value. During normal operation, the process uncertainty associated with each state is assumed to follow a zero-mean Gaussian distribution with a 5% of the nominal steady-state value of the corresponding state as its standard deviation. Similarly, the measurement noise associated with each state is expected to follow a zero-mean Gaussian distribution with a standard deviation that is set to 10% of the nominal steady-state value of the corresponding state. These assumptions hold in all the scenarios performed for WTP that describe the characteristics of the distribution of the process uncertainties and measurement noises in both the standard MHE and the core-Gaussian component in RMHE (i.e., assuming normal plant operation). The different scenarios considered to test the performance of the proposed RMHE formulation are presented next.

8.2.1.1 *Scenario I:* Measurement device failure

To illustrate the benefits of this approach, this scenario considers a failure in the measurement devices at the 50th sampling time that lasts for 50th sampling intervals. This failure in the hardware sensors leads to unexpectedly random measurement noises during this time. The random noises associated with each of the measurable states over the simulation time are depicted in Figure 8-1.

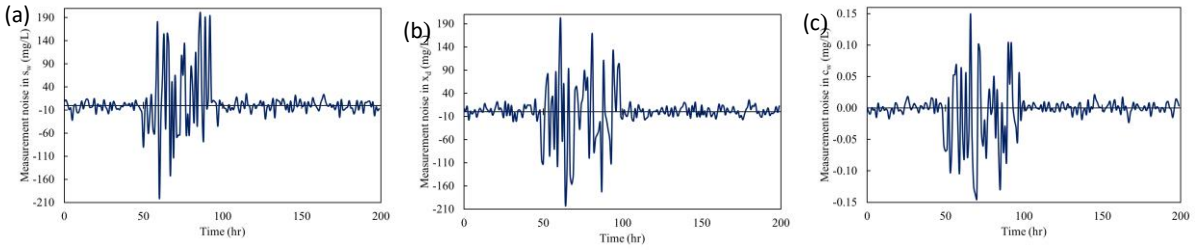


Figure 8-1: True random measurement noises associated with each measurement

Given these outliers in the measurement noises, let's set $\Upsilon_{mn} = 1$ in the RMHE formulation. As discussed in section 8.1, the number of perimeter-Gaussian components as well as the mean value and covariance matrix for each of these perimeter-Gaussian components are tuning parameters in the present RMHE framework. These parameters need to be chosen based on historical information available for the process plant. The simplest assumption is to consider that there is only one perimeter-Gaussian component in the mixture of the RMHE framework (i.e., $ngr = 2$) that exhibits the same covariance matrix as the core-Gaussian component, i.e., $\boldsymbol{R}^1 = \boldsymbol{R}^2$. Note that the noises are assumed to be uncorrelated, i.e., $\boldsymbol{R}^1 = \boldsymbol{R}^2 = diag(\left[std_1, std_2, \dots std_m, std_{n_y}\right])$; $\boldsymbol{std} = \left[std_1, std_2, \dots std_m, std_{n_y}\right]$. As for the mean value, two instances are considered, i.e., $\boldsymbol{\tau}^+$ and $\boldsymbol{\tau}^-$. These two instances are used to explore the RMHE performance when compared to the standard MHE performance. That is, $\boldsymbol{\tau}^+ = +6 * \boldsymbol{std}$ and $\boldsymbol{\tau}^- = -6 * \boldsymbol{std}$. Figure 8-2 depicts a schematic of the core-Gaussian component together with the perimeter-Gaussian component considered for each of these instances.

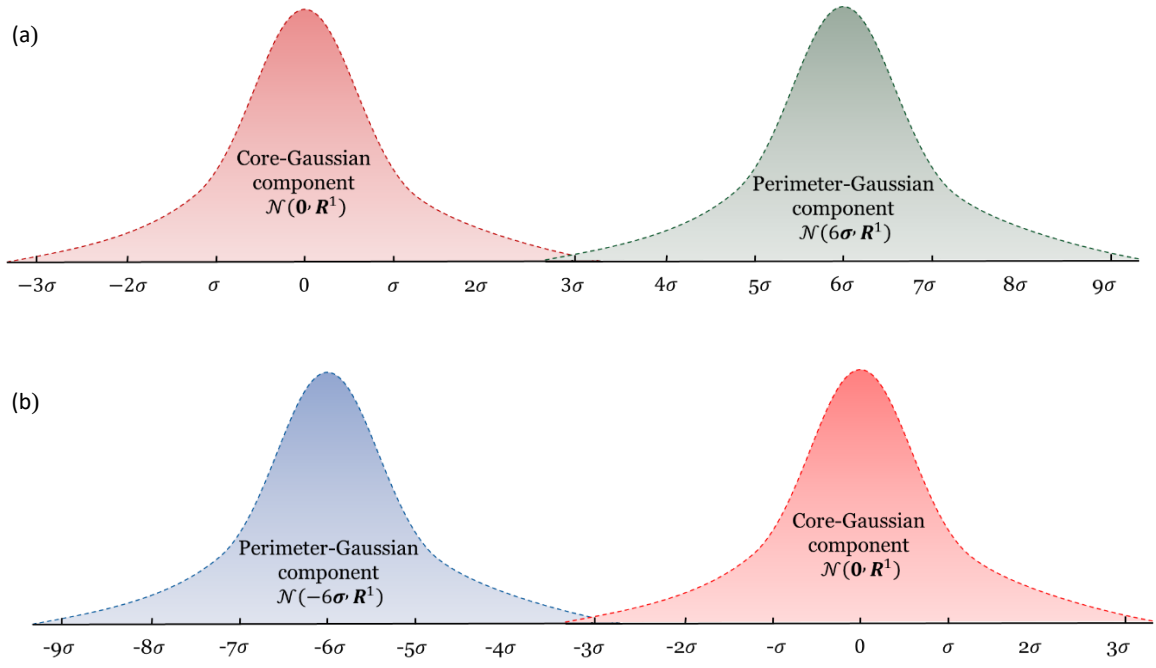


Figure 8-2: Gaussian components considered in RMHE ($ngr = 2$). The weights on the core-Gaussian and the perimeter-Gaussian are set to 0.5 for instances (a) $\boldsymbol{\tau}^+$ and (b) $\boldsymbol{\tau}^-$

Figure 8-3 presents the estimations of the unknown states provided by different estimation schemes, i.e., the standard MHE and RMHE using $\boldsymbol{\tau}^-$ and $\boldsymbol{\tau}^+$. Moreover, Table 8-1 represents the MSE obtained for the state variables using the estimation methods mentioned above. As shown in Table 8-1, the estimation error has been evaluated for four time periods, i.e., during normal operation (0-50hr), during measurement equipment failure (50hr-100hr), during the transition time (100hr-150hr), and after equipment failure (150hr-200hr). As shown in Figure 8-3 and Table 8-1, these three periods of operation are denoted as Part I, Part II, Part III, and Part IV, respectively. During Part III, it is assumed that the equipment was fixed but the effect of unexpected noises still lasts in the process and unexpected noises are still involved within the estimation horizon. According to Figure 8-3, the proposed RMHE and the standard MHE provided similar estimations for the first 50 sampling intervals (Part I) on which all the measurement devices are assumed to be working properly; hence, the *a priori* knowledge of the distribution of the measurement noises is valid. Likewise, the RMHE and standard MHE seem to have as good of a performance in the last 50hr of the operation (Part IV) in which the measurement devices were fixed and random noises are assumed to be back to normal (i.e., 150-200th sampling time). According to Table 8-1, the estimation error reported for RMHE and standard MHE are relatively the

same, i.e., $\frac{MSE_{x_w}^{(MHE)}}{MSE_{x_w}^{(RMHE)}} \approx 0.95$. During the device failure time, i.e., 50th-100th sampling time (Part II), the estimates provided by the standard MHE significantly deviate from the actual plant states, as shown in Figure 8-3. This is because the device failure led to large measurement noises that no longer follow the known zero-mean Gaussian distribution. On the other hand, the proposed RMHE was able to increase the robustness in the estimation and consequently, improved the estimation accuracy during the devices' failure period (Part II) when compared to the standard MHE. This improvement in the estimation accuracy is quantified in Table 8-1 (Part II). For instance, the $MSE_{x_r}^{(RMHE)}$ was reduced by an order of magnitude when compared to $MSE_{x_r}^{(MHE)}$. Note that during the transition time (Part III) and for the first 10 time intervals, RMHE improves the estimations when compared to the standard EKF, as shown in Figure 8-3 and Table 8-1. This is because the unexpected noises that are present in the process before repairing the equipment are still involved in the estimation horizon of both RMHE and standard RMHE. Evidently, both state estimation schemes provide similar performance after the effect of unexpected noises in the estimation horizon vanish, which is similar to Part I and Part IV.

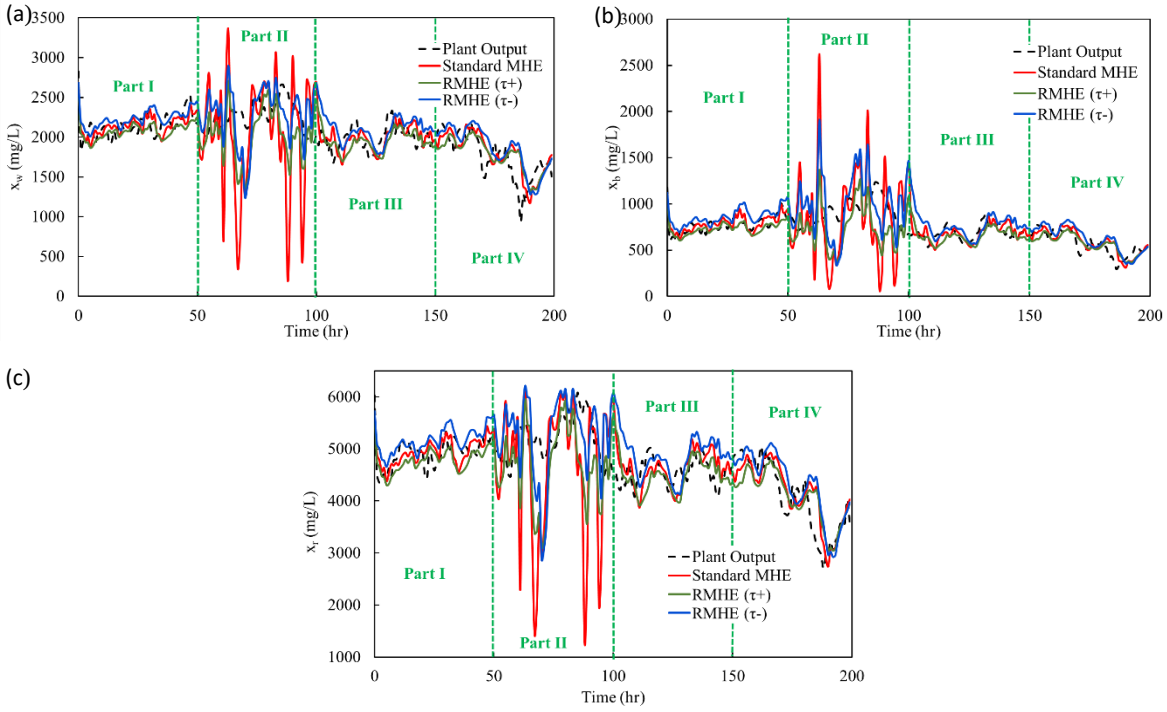


Figure 8-3: Estimation provided by the standard MHE, RMHE with $ngr = 2$ under the unexpected measurement noise scenario for unknown state (a) x_w; (b) x_b; (c) x_r

Table 8-1: MSE for unknown states x_w, x_b, x_r using different estimation schemes at various simulation time

Sampling time (hr)	Part I: Before device failure (0-50)			Part II: During device failure (50-100)			Part III: During transition (100-150)			Part IV: After device failure (150-200)		
Estimation method (n)	$MSE_{x_w}^{(n)}$	$MSE_{x_b}^{(n)}$	$MSE_{x_r}^{(n)}$	$MSE_{x_w}^{(n)}$	$MSE_{x_b}^{(n)}$	$MSE_{x_r}^{(n)}$	$MSE_{x_w}^{(n)}$	$MSE_{x_b}^{(n)}$	$MSE_{x_r}^{(n)}$	$MSE_{x_w}^{(n)}$	$MSE_{x_b}^{(n)}$	$MSE_{x_r}^{(n)}$
Standard MHE	2.29e+4	4.90e+3	6.70e+04	6.36e+5	2.91e+5	2.18e+6	2.44e+4	2.65e+4	2.11e+5	5.72e+4	8.90e+4	1.56e+5
RMHE ($\boldsymbol{\tau}^+$)	2.06e+4	6.14e+3	8.04e+04	1.86e+5	1.02e+5	8.61e+5	1.84e+4	1.32e+4	1.89e+5	5.71e+4	9.49e+4	1.79e+5
RMHE ($\boldsymbol{\tau}^-$)	2.40e+4	1.28e+4	1.14e+05	1.42e+5	1.40e+5	6.61e+5	3.60e+4	4.38e+4	3.85e+5	7.62e+4	1.38e+4	2.31e+5

These observations can be further explained through the weights ($\boldsymbol{\alpha m}^r$) presented in Figure 8-4 for each of the instances (i.e., $\boldsymbol{\tau}^-$ and $\boldsymbol{\tau}^+$). In the case of normal plant operation (i.e., Part I and IV), the RMHE should approximate the Gaussian mixture to model to be similar to the known zero-mean Gaussian distribution of the noises (i.e., core-Gaussian component in the mixture). As shown in Figure 8-4, the value of weight for the core-Gaussian component is close to 1 at the beginning and at the end of these operations (i.e., Part I and Part IV) where the measurement devices are assumed to be working properly, as shown in Figure 8-1. Hence, measurement noises are expected to follow the known Gaussian distribution (the core-Gaussian component) during these periods.

Since the number of Gaussian components in the mixture is set to 2 (i.e., zero mean core-Gaussian and non-zero mean perimeter-Gaussian), the weight associated with the core-Gaussian must be close to 1 such that the mean value of the overall Gaussian mixture model is close to zero (see Equations (8-1)-(8-4)). Note that a weight close to the unity on the core-Gaussian component suggests that the specifications of the measurement noises' distributions are similar in both RMHE and the standard MHE, which results in RMHE estimations similar to that obtained by the standard MHE. Note that during normal operation, the weights for the core-Gaussian component are not expected to be exactly 1 since the perimeter-Gaussian component slightly overlaps with the core component. Moreover, Figure 8-4 shows that the weight on the perimeter-Gaussian component increased during the devices' malfunction at the 50th sampling interval (Part II), i.e., the RMHE scheme reduced the weight of the core-Gaussian component to capture the unexpected random noises present in the process. The weight corresponding to the core-Gaussian component was increased to about 1 once the measurement devices

were fixed (Parts III and IV). These changes in the weights assigned to each Gaussian component in the mixture is a good indication that the unexpected noises have been properly identified by the RMHE scheme. This result also shows that the Gaussian mixture model considered in this framework is able to provide an appropriate approximation of the distribution of the unexpected noises to the state estimation scheme.

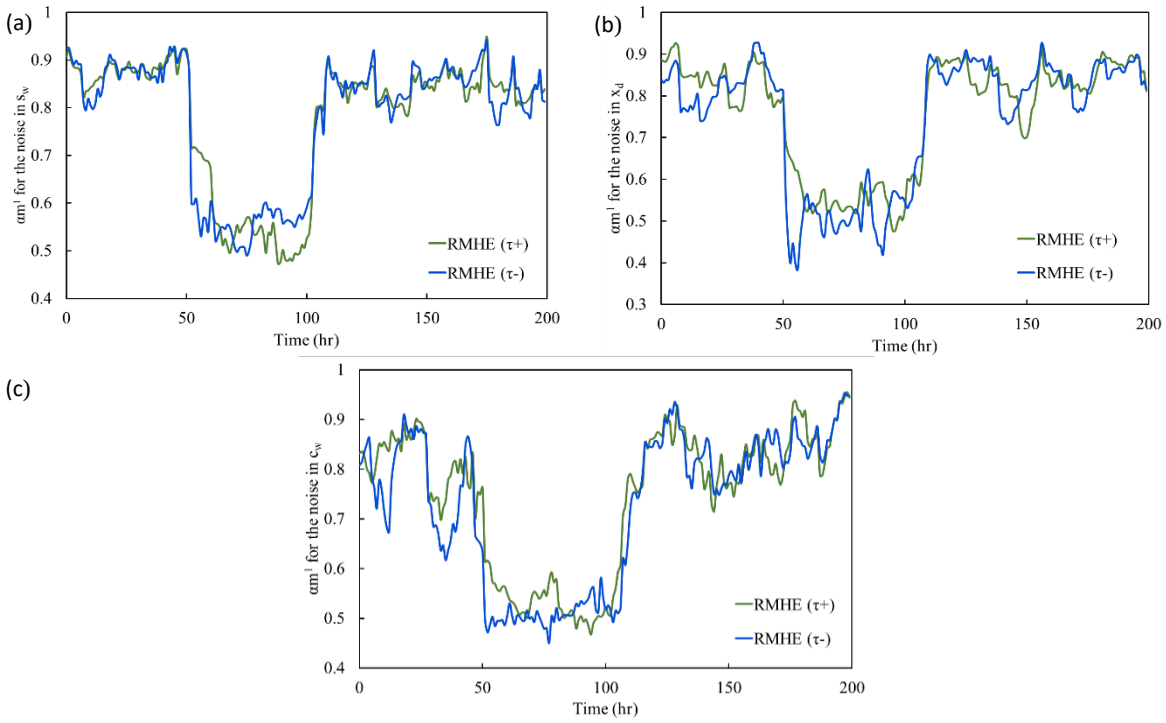


Figure 8-4: Optimal weights ($\boldsymbol{\alpha m}^r$) provided by the RMHE framework with $ngr = 2$ corresponded to the core-Gaussian component for the measurement noise associated with (a) s_w; (b) x_d; (c) c_w

8.2.1.2 *Scenario II:* Effect of the length of the estimation horizon

This scenario provides insight on the effect of the length of estimation horizon (N) on the RMHE performance. This scenario performed the state estimation using both RMHE and standard MHE considering different length of estimation horizons ($N \in \{10,50\}$). Figure 8-5 presents the results of performing this scenario. Figures 8-5(a)-(c) presents the estimates provided by the standard MHE and RMHE instances for $N = 50$ whereas Figures 8-5(d)-(f) directly compares the effect of increasing the length of estimation horizon on the RMHE performance using two different time horizons. As is the case when $N = 10$ (i.e., *Scenario I* in section 8.2.1.1), the results presented in Figures 8-5(a)-(c) shows that the proposed RMHE improves the estimation significantly when compared to the standard MHE.

For instance, the estimation error for x_w in the case of using RMHE has been reduced by one order of magnitude than that obtained from the standard MHE. Nevertheless, the accuracy in the estimation improved as the length of the estimation horizon was increased. This was expected as RMHE is a robust version of the standard MHE; hence, increasing the length of the horizon is expected to lead to an improvement in the estimation accuracy at the cost of increasing the computational costs. Moreover, a longer horizon in the RMHE framework results in a larger number of measurement noise samples considered in the formulation, thus leading to a more accurate Gaussian mixture model approximation. Nevertheless, this is a trade-off between the accuracy in the estimation and computational costs, as in the standard MHE scheme.

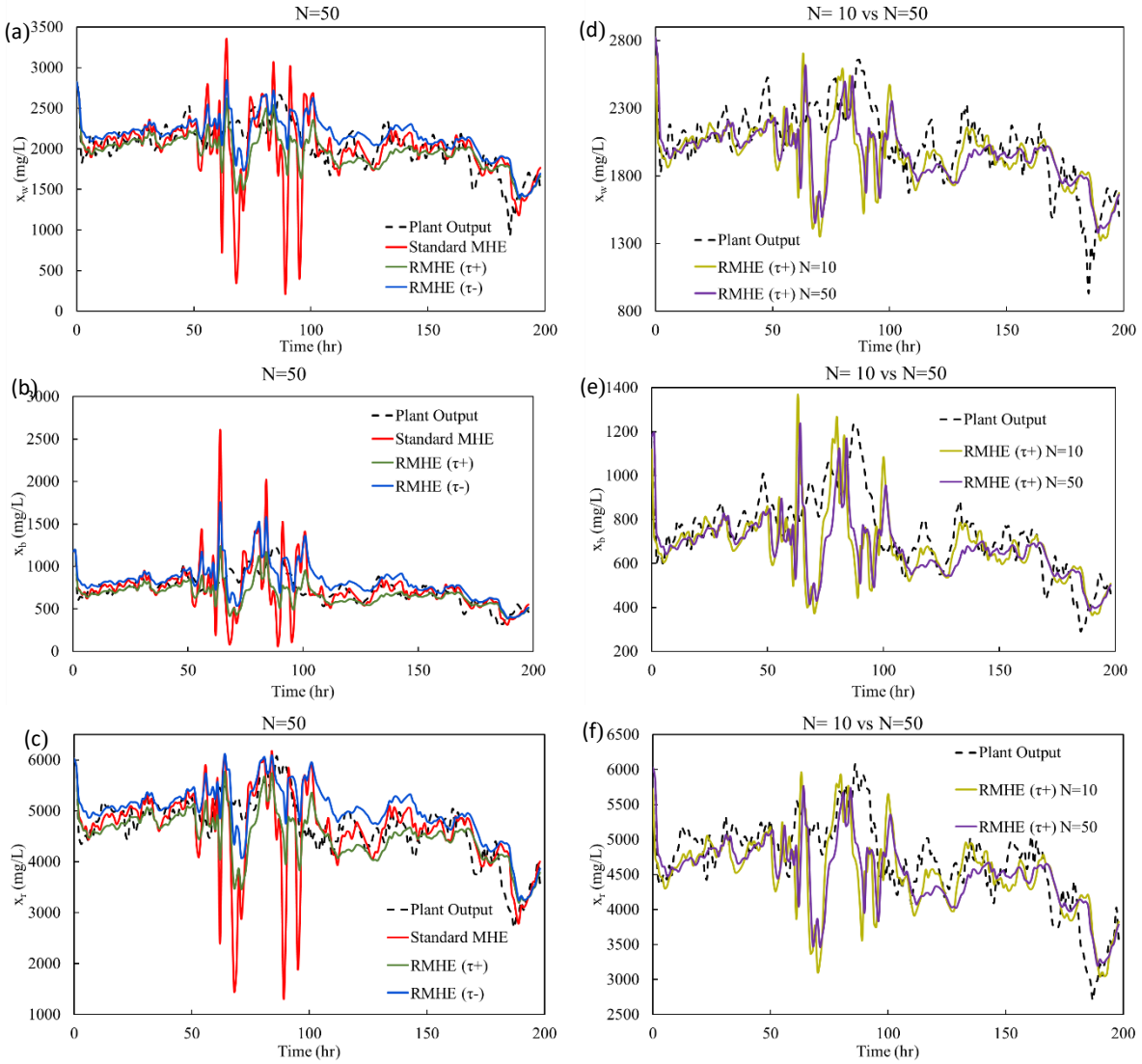


Figure 8-5: Estimations for the unknown states (a)-(c) provided by different estimation schemes considering $N = 50$; (d)-(f) provided by RMHE ($\boldsymbol{\tau}^{+}$) for both $N = 50$ and N=10

8.2.1.3 *Scenario III:* effect of number of perimeter-Gaussian components

This scenario investigates the RMHE performance using a different number of Gaussian components (ngr). In this scenario, ngr is set to 3 whereas the assumptions made for *Scenario I* (section 8.2.1.1) were also considered here. Since it is unknown *a priori* that measurement noises during the equipment failure are positive or negative, setting $ngr = 3$ is a reasonable assumption. That is, one perimeter-Gaussian component on each side of the core-Gaussian component is considered in the Gaussian mixture model of the measurement noise. The mean value for the perimeter-Gaussian components is set to positive and negative values that are six times larger than the standard deviation, i.e., both perimeter-Gaussian components shown in Figure 8-2(a) and Figure 8-2(b) are considered in the mixture model of the noises considered in this scenario. For simplicity, the same standard deviation is considered for all the Gaussian components in the mixture as in the core-Gaussian component, i.e., $\boldsymbol{R}_k^1 = \boldsymbol{R}_k^2 = \boldsymbol{R}_k^3$. Given these assumptions, the RMHE framework adapted for *Scenario III* is as follows:

$$\min_{\substack{\{\boldsymbol{x}_j, \boldsymbol{w}_j\}_{j=k-N}^{k-1}, \\ \{\boldsymbol{\alpha m}_r, \boldsymbol{\alpha p}_r\}_{r=1}^{3}}} \sum_{j=k-N}^{k-1} \left\|\boldsymbol{w}_j\right\|_{\boldsymbol{Q}_k^{1^{-1}}}^2 + \sum_{j=k-N+1}^{k} \left\|\boldsymbol{v}_j - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right\|_{\boldsymbol{R}_k^{\boldsymbol{GM}^{-1}}}^2 + \varphi_{k-N} \tag{8-6}$$

s.t. (8-7)

$WTP\ mechanistic\ model\ Equation\ (4-29)$

$\boldsymbol{y}_j = [s_{w\,j}, x_{d\,j}, c_{w\,j}]^T + \boldsymbol{v}_j\ ; \qquad \forall j = k-N+1, \dots k$

where: $\boldsymbol{w}_j \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}^1)$ and $\boldsymbol{v}_j \sim \mathcal{N}(\boldsymbol{\tau}_j^{\boldsymbol{GM}}, \boldsymbol{R}_j^{\boldsymbol{GM}})$; $\quad \forall j = k-N+1, \dots k$

$\boldsymbol{\tau}_k^{\boldsymbol{GM}} = \boldsymbol{\alpha m}_k^2 \boldsymbol{\tau}_k^- + \boldsymbol{\alpha m}_k^1(\boldsymbol{0}) + \boldsymbol{\alpha m}_k^3 \boldsymbol{\tau}_k^+$; where: $\boldsymbol{\tau}_k^+ = -\boldsymbol{\tau}_k^-$ (8-8)

$\boldsymbol{R}_k^{\boldsymbol{GM}} = \boldsymbol{R}_k^1 + \left(\boldsymbol{\alpha m}_k^2 + \boldsymbol{\alpha m}_k^3\right)\boldsymbol{E}^2$;

where: $\boldsymbol{E}^2 = \left(\boldsymbol{\tau}_k^+ - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)\left(\boldsymbol{\tau}_k^+ - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)^{\boldsymbol{T}} = \left(\boldsymbol{\tau}_k^- - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)\left(\boldsymbol{\tau}_k^- - \boldsymbol{\tau}_k^{\boldsymbol{GM}}\right)^{\boldsymbol{T}}$

$\sum_{r=1}^{3} \boldsymbol{\alpha m}_k^r = \boldsymbol{1}$;

$\boldsymbol{0} \leq \boldsymbol{\alpha m}_k^r \leq \boldsymbol{1}\ ; \forall r = 1,2,3$

where Equation (8-7) represents the corresponding models for process model (WTP model), measurements, the process uncertainties and measurement noises. Note that all the WTP equations included in Equation (8-7) consider additive process uncertainties $\boldsymbol{w}_j$. The estimations provided by the different estimation schemes considering $N = 10$ as well as the weights on the Gaussian components provided by RMHE are depicted in Figure 8-6. Similar to *Scenario I*, the proposed RMHE improved the estimation when compared to the standard MHE, as shown in Figures 8-6(a)-(c). For instance, the estimation error for x_b was reduced by one order of magnitude in the case of RMHE when compared to that obtained by the standard MHE. However, the weights assigned on the core-Gaussian component remained unchanged (and set to zero) throughout the process, as shown in Figures 8-6(d)-(f); the weights on the perimeter-Gaussian components remain at about 0.5 throughout the process. The weight on the core-Gaussian component was expected to be close the unity under the normal plant operation, as observed in *Scenario I* (i.e., Parts I and IV). This condition occurred because the RMHE aims to find the optimal Gaussian mixture model that has zero mean, i.e., $\boldsymbol{\tau}_k^{\boldsymbol{GM}} = \mathbf{0}$. To pursue this goal, the RMHE framework must set $\boldsymbol{\alpha m}_k^1 \geq 0$ and $\boldsymbol{\alpha m}_k^2 = \boldsymbol{\alpha m}_k^3 = (\mathbf{1} - \boldsymbol{\alpha m}_k^1)/2$, as shown in Equation (8-8). On the other hand, the RMHE formulation aims to minimize the sum of weighted noises and uncertainties over the estimation horizon, as shown in Equation (8-6). Considering $\boldsymbol{\tau}_k^{\boldsymbol{GM}} = \mathbf{0}$, the larger $\boldsymbol{R}_k^{\boldsymbol{GM}}$ the smaller the objective function. This means that the RMHE formulation should find the optimum value of the weight of each component that maximizes $\boldsymbol{R}_k^{\boldsymbol{GM}}$. Since $\left(\boldsymbol{\alpha m}_k^2 + \boldsymbol{\alpha m}_k^3\right)\boldsymbol{E}^2$ in Equation (8-8) is a positive term, the weights on the perimeter-Gaussian components should be non-zero to maximize $\boldsymbol{R}_k^{\boldsymbol{GM}}$. Therefore, given Equation (8-8) and $\boldsymbol{\alpha m}_k^2 = \boldsymbol{\alpha m}_k^3 = (\mathbf{1} - \boldsymbol{\alpha m}_k^1)/2$, RMHE returns $\boldsymbol{\alpha m}_k^2 \approx \boldsymbol{\alpha m}_k^3 \approx \mathbf{0.5}$ and $\boldsymbol{\alpha m}_k^1 \approx \mathbf{0}$, just as depicted in Figure 8-6 when performing *Scenario III*. Note that a similar behavior is expected for $ngr \geq 3$, regardless of the choice for the mean value and covariance of the perimeter-Gaussian components. As an MHE-based estimation scheme, the proposed RMHE framework also aims to minimize the errors in the process by maximizing the weights on these errors that are in fact the covariance matrices. Note that unnecessary large standard deviations in the noises may lead to inaccurate estimates during normal plant operation. This can be point out as the limitation of the proposed approach, i.e., even though it is possible to set ngr to any positive integer number, the best choice to avoid biased estimation (because of the large covariance matrices) is to consider only one additional perimeter-Gaussian component.

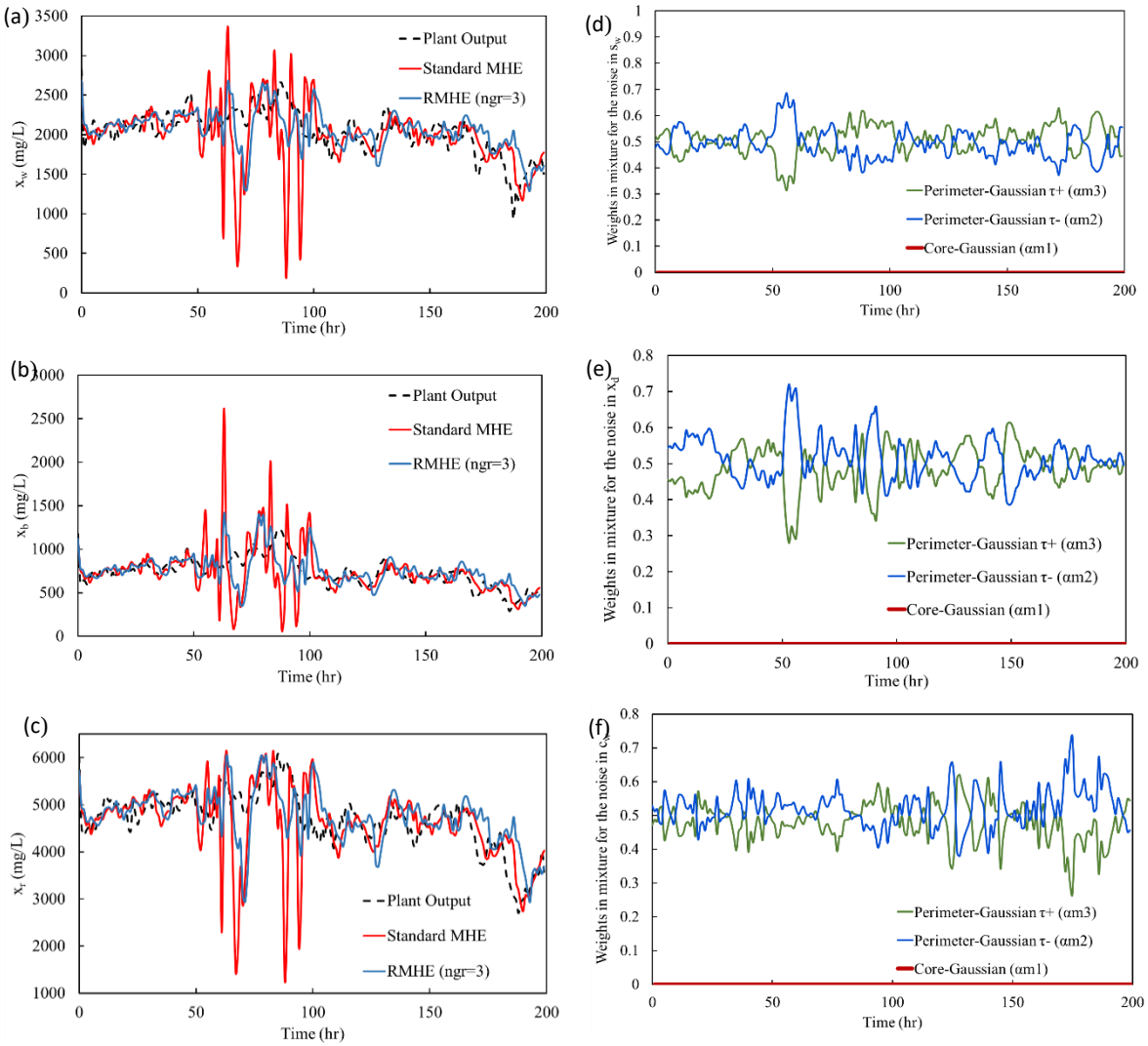


Figure 8-6: (a)-(c) Estimations for the unknown states provided by different estimation schemes considering $N = 10$ and $ngr = 3$; (d)-(f) optimal weights provided by RMHE considering $ngr = 3$

Nevertheless, the RMHE scheme performs well and avoid unnecessary large covariance matrices in the case of $ngr = 2$, because the RMHE problem requires to satisfy $\boldsymbol{\tau}_k^{\boldsymbol{GM}} = \mathbf{0}$ for the normal operating condition, and thus, the only choice is to consider $\boldsymbol{\alpha m}_k^1 \approx \mathbf{1}$. Based on the results presented in this section, it is recommended to choose $ngr = 2$ as the proposed RMHE framework can successfully improve the estimation and avoid biased estimations under this assumption.

Moreover, as mentioned in section 8.1, the additional constraints considered in the RMHE problem (Equations. (8-3) and (8-4)), does not increase considerably the computational costs when compared to

the standard MHE. For instance, given $N = 10$ considered in *Scenario I* and *Scenario III*, the averaged CPU time required to provide the point estimates is: 0.22 s for the MHE, 0.24 s for the RMHE ($ngr = 2$) and 0.25 s for the RMHE ($ngr = 3$). While increasing ngr also increases the number of additional constraints and decision variables in the RMHE problem, the value of ngr does not considerably affect the required CPU time to perform the online state estimation while using this method.

Note that the current case study considered unknown distributions for the noises due to unscheduled equipment malfunction. Thus, EMHE would not have any *a priori* knowledge of non-Gaussianity or possible changes in the distributions. That leaves the EMHE problem with the only choice of considering the same assumption considered in the standard MHE for these distributions, i.e., known zero-mean Gaussian distribution. As mentioned in Chapter 7, the EMHE problem is equivalent to the standard MHE problem if the both process uncertainties and measurement noises are assumed to follow zero-mean Gaussian distributions. Thus, the standard MHE and EMHE are expected to provide the same performance for the scenarios considered for the WTP.

8.2.2 Network of continuous-stirred tank reactors (CSTRs) and separator

To further confirm the RMHE performance for the scenarios involving unexpected random variables, a second case study involving a larger number of states than that considered in the WTP has been used to test the RMHE performance. This case study considers a network of two CSTRs in series with a separator. This process has been used in to test previous state estimation and robust MHE schemes[129,199].

Figure 8-7 shows a flowsheet of the process. A second order reaction, i.e., $A \xrightarrow{k_r} B$, takes place on each CSTR; k_r denotes the reaction rate constant. The network includes nine states, which are the temperature and the concentration of the species A and B on each CSTR and the separator. The subscript 1,2,3 refers to the corresponding state inside the first CSTR, second CSTR, and the separator, respectively. A recycle stream is used to send back unreacted species collected in the separator to the first CSTR.

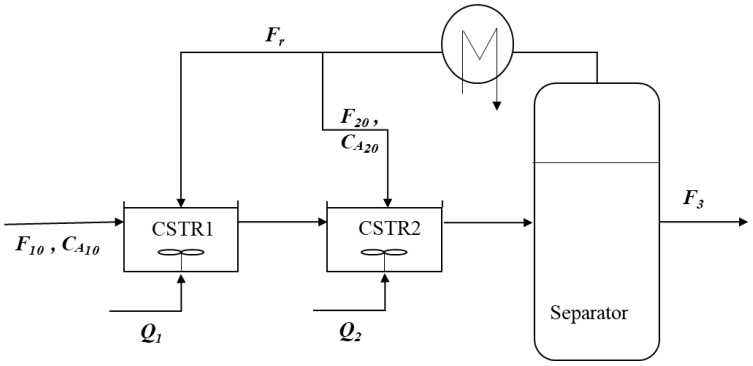


Figure 8-7: Network of CSTRs and separator flowsheet

The mechanistic model that represents this process is as follows[129,199]:

$$\frac{dT_1}{dt} = \frac{F_{10}}{V_1}(T_{10} - T_1) + \frac{F_r}{V_1}(T_3 - T_1) + \frac{-\Delta H}{\rho C_p} k e^{\frac{-E}{RT_1}} C_{A_1}^2 + \frac{Q_1}{\rho C_p V_1}$$

$$\frac{dC_{A_1}}{dt} = \frac{F_{10}}{V_1}\left(C_{A_{10}} - C_{A_1}\right) + \frac{F_r}{V_1}\left(C_{A_r} - C_{A_1}\right) - k e^{\frac{-E}{RT_1}} C_{A_1}^2$$

$$\frac{dC_{B_1}}{dt} = \frac{-F_{10}}{V_1} C_{B_1} + \frac{F_r}{V_1}\left(C_{B_r} - C_{B_1}\right) + k e^{\frac{-E}{RT_1}} C_{A_1}^2$$

$$\frac{dT_2}{dt} = \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_{20}}{V_2}(T_{20} - T_2) + \frac{-\Delta H}{\rho C_p} k e^{\frac{-E}{RT_2}} C_{A_2}^2 + \frac{Q_2}{\rho C_p V_2}$$

$$\frac{dC_{A_2}}{dt} = \frac{F_1}{V_2}\left(C_{A_1} - C_{A_2}\right) + \frac{F_{20}}{V_2}\left(C_{A_{20}} - C_{A_2}\right) - k e^{\frac{-E}{RT_2}} C_{A_2}^2 \qquad (8\text{-}9)$$

$$\frac{dC_{B_2}}{dt} = \frac{F_1}{V_2}\left(C_{B_1} - C_{B_2}\right) - \frac{F_{20}}{V_2} C_{B_2} + k e^{\frac{-E}{RT_2}} C_{A_2}^2$$

$$\frac{dT_3}{dt} = \frac{F_2}{V_3}(T_2 - T_3) + \frac{-\Delta H_{vap} F_{rm}}{\rho C_p V_3} + \frac{Q_2}{\rho C_p V_3}$$

$$\frac{dC_{A_3}}{dt} = \frac{F_2}{V_3}\left(C_{A_2} - C_{A_3}\right) - \frac{F_r}{V_3}\left(C_{A_r} - C_{A_3}\right)$$

$$\frac{dC_{B_3}}{dt} = \frac{F_2}{V_3}\left(C_{B_2} - C_{B_3}\right) - \frac{F_r}{V_3}\left(C_{B_r} - C_{B_3}\right)$$

where

$$C_{C_3} = \left(\rho - C_{A_3}M_A - C_{B_3}M_B\right)/M_C$$

$$C_{\iota_r} = \frac{\alpha_\iota \rho C_{\iota_3}}{\alpha_A C_{A_3} M_A + \alpha_B C_{B_3} M_B + \alpha_C C_{C_3} M_C}, \qquad \iota = A, B, C$$

$$F_{rm} = F_r\left(C_{A_r} + C_{B_r} + C_{C_r}\right)$$

The nominal values for all the model parameters can be found elsewhere[129,199]. The current study assumes that online measurements are only available for the temperature in the vessels, i.e., T_1, T_2, and T_3; the remaining six states representing the concentrations, i.e., C_{A_1}, C_{A_2}, C_{A_3}, C_{B_1}, C_{B_2}, and C_{B_3}, are the unknown states considered for this process. The linear observability matrix for this system was evaluated using the Jacobian matrix of this system (see appendix E) identified around the nominal steady-state condition for the states. The linear observability matrix was full rank; hence, the observability of the system around the nominal steady-state condition considered in this work was confirmed. The sampling time for this process is 0.01 hr. Similar to the WTP case study, to further focus on the effect of unexpected operating condition on the state estimation, the length of the time horizon was set to a relatively large value to reduce the effect of the arrival cost in the estimation. According to the preliminary tests performed for this case study, setting the horizon length to $N = 30$ is sufficient to reduce the effect of the arrival cost in the estimation. The steady-state values of the states reported in Table B-3 in Appendix B have been used as the true plant states at the beginning of the operation. The states in the standard MHE and RMHE schemes have been initialized assuming that they are 5% away from the true plant states. Under the normal plant operation, the process uncertainties follow zero-mean Gaussian distributions where the standard deviation for the process uncertainty associated with each state is assumed to be 1% of the nominal steady-state value of the corresponding state. Likewise, measurement noise is expected to follow zero-mean Gaussian distribution with a standard deviation set to 0.001% of the nominal steady-state value of the corresponding measured state (i.e., T_1, T_2, and T_3). Note that these specifications describe the distribution of the noises/uncertainties in the standard MHE as well as the core-Gaussian component in RMHE framework.

For this case study, a scenario involving unexpected process uncertainties has been considered to show the performance of the proposed RMHE in comparison with the standard MHE. That is, due to the external disturbances (zero mean with 0.001 standard deviation) imposed on the concentration species

A in the feed stream at the initial time of the operation, the process uncertainties present in the system are assumed to follow a bimodal distribution, as shown in Figure 8-8. The distribution of the unexpected process uncertainties present in the system is denoted as “True Density”, whereas the zero-mean Gaussian distribution considered in the standard MHE is depicted as the “MHE Density” in Figure 8-8. Note that in practice, the true distributions are not available since they are unexpected and unknown to both users and estimation schemes; hence, the Gaussian mixture model cannot be evaluated prior to the state estimation. Since *a priori* knowledge of the non-Gaussianity present in the process uncertainties’ densities is not available, the reasonable assumption when performing EMHE is to consider the same zero-mean Gaussian distributions to describe the process uncertainties as that considered in the standard MHE. Similar to the discussion presented for the first case study, under this assumption, the EMHE problem would reduce to the standard MHE problem and therefore, both MHE and EMHE are expected to provide the same estimations.

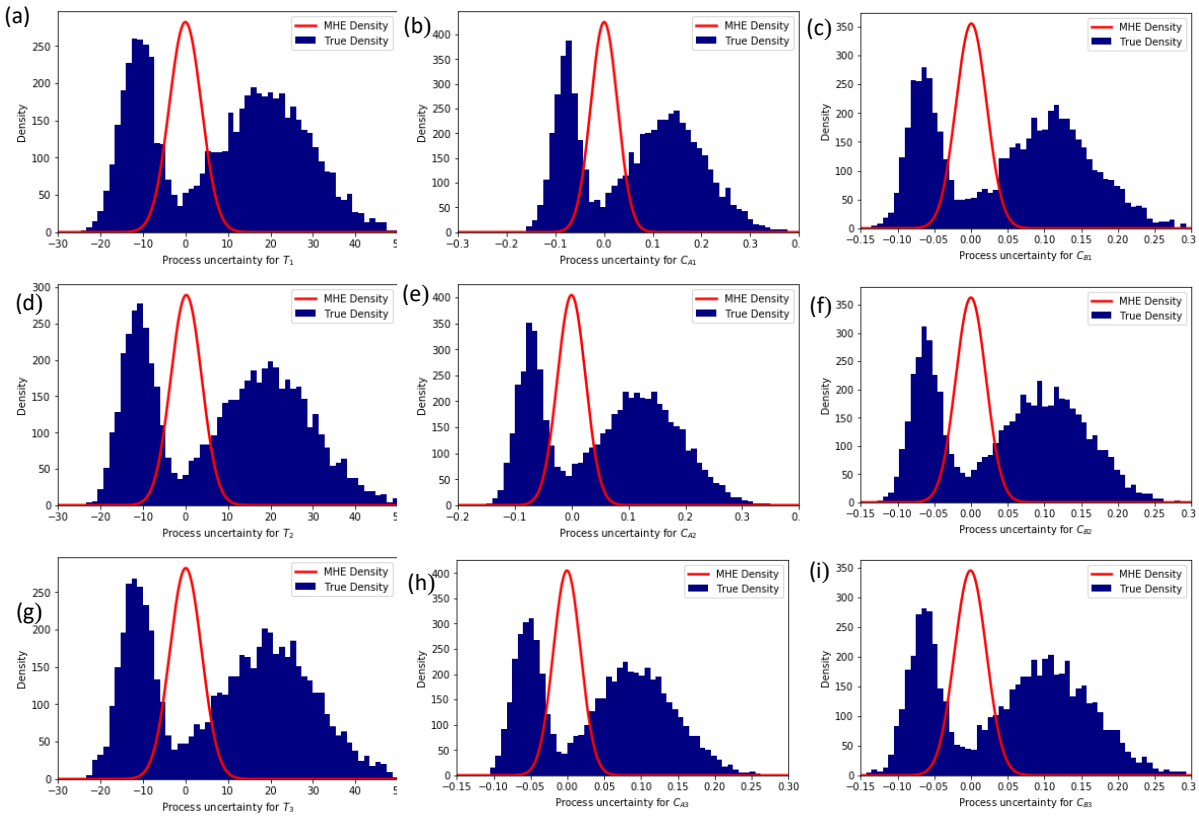


Figure 8-8: Histogram for the true distribution of the unexpected process uncertainties and the corresponding zero-mean Gaussian distribution considered in the standard MHE

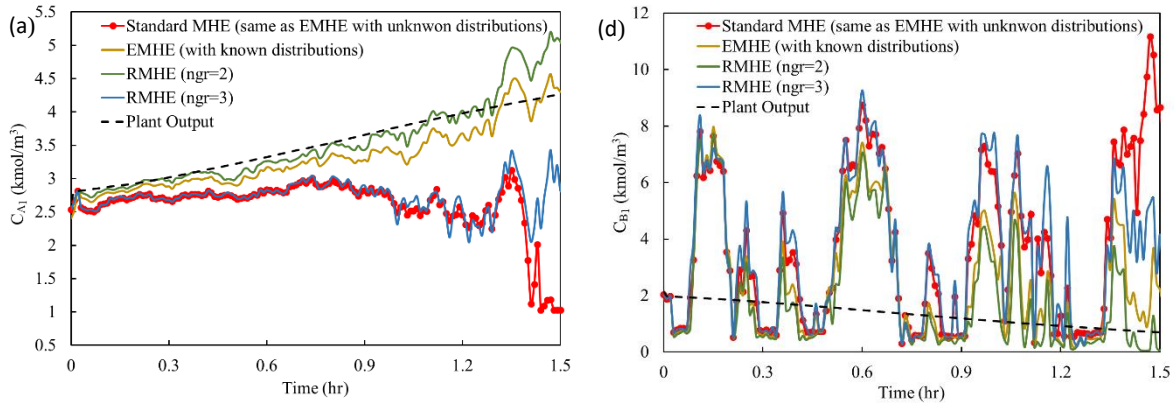
Moreover, the EMHE scheme was tested assuming an ideal scenario where the non-Gaussian distributions of the process uncertainties are known *a priori*. This was done with the sole purpose of comparing the performance of EMHE (under known non-Gaussian uncertainties) with RMHE (under unknown non-Gaussian uncertainties). In this ideal scenario, the Gaussian mixture models consisting of two Gaussian components were approximated to the non-Gaussian distributions presented in Figure 8-8. Given the ideal assumption considered in this instance, i.e., known non-Gaussian process uncertainties, the EM algorithm was performed offline (prior to the state estimation) to provide the mean-value and the covariance matrix of the overall Gaussian mixture models to perform EMHE.

As the unexpected random variables affecting the plant operation is the process uncertainty, $\Upsilon_{pu} = 1$ in the RMHE formulation, as shown in Equation (8-5). Two instances considering $ngr = 2$ and $ngr = 3$ are considered to further investigate effect of this tuning parameter while using a different case study. The covariance matrix for the perimeter-Gaussian components is the same as the core-Gaussian component covariance matrix. As for the mean value of the perimeter-Gaussian component in the $ngr = 2$ instance, the current study considered that $\boldsymbol{\mu}^{+} = +6 * \boldsymbol{stdp}$. Note that $\boldsymbol{stdp} = [stdp_1, stdp_2, \dots, stdp_{n_x}]$ and $\boldsymbol{Q}^{r} = diag([stdp_1, stdp_2, \dots stdp_m, stdp_{n_x}])$. Moreover, the mean value for the perimeter-Gaussian components in $ngr = 3$ is set to $\boldsymbol{\mu}^{+} = +6 * \boldsymbol{stdp}$ and $\boldsymbol{\mu}^{-} = -6 * \boldsymbol{stdp}$.

Figure 8-9 illustrates the estimations obtained for the unknown states by performing the standard MHE (EMHE under zero-mean Gaussian process uncertainties), ideal EMHE under known non-Gaussian distributions, and the two RMHE instances (i.e., $ngr \in \{2,3\}$). Table 8-2 compares the estimation errors obtained from each estimation scheme. As shown in Figure 8-9 and Table 8-2, the RMHE with two Gaussian components ($ngr = 2$) was able to improve the estimation accuracy significantly; for example, the estimation error for the state C_{A_1} was reduced by 94% in the case of using RMHE ($ngr = 2$) when compared to the standard MHE. The weights on the core-Gaussian and the perimeter-Gaussian components are around 0.5 over this unexpected operating condition. This result is expected as the true non-Gaussian uncertainties presented in Figure 8-8 suggest that these random variables follow bimodal distributions. According to Figure 8-9, the RMHE ($ngr = 3$) was not able to improve the robustness of the estimation scheme significantly. The estimation error reported in Table 8-2 for this instance is as large as the error in the estimates provided by the standard MHE. This happens because the RMHE formulation focuses on maximizing the standard deviation of the process uncertainties when $ngr \geq 3$,

which lead to unnecessary large covariance matrices and consequently, biased estimates, as discussed in *Scenario III* for the WTP (see section 8.2.1.3). Thus, it is recommended to perform the proposed RMHE considering $ngr = 2$ to improve the estimation accuracy and avoid biased estimates.

Moreover, according to Figure 8-9 and Table 8-2, the estimations provided by the RMHE ($ngr = 2$) considering unknown process uncertainties are only slightly different from the estimates provided by the ideal EMHE (EMHE under known non-Gaussian process uncertainties). For instance, the estimation error for C_{B_1} in the case of performing RMHE and the ideal EMHE are quite similar (i.e., with a small difference of ~ 0.004), as shown in Table 8-2. As mentioned in *Remark 1* (see section 8.1), the difference in the performance of these schemes is due to the difference in the Gaussian mixture models evaluated in the ideal EMHE instance and RMHE ($ngr = 2$) to describe the process uncertainties. The similar responses observed by these two instances supports the high performance of the proposed RMHE scheme and the fact that this approach is capable of modelling online the unknown non-Gaussian process uncertainties. Moreover, the true densities presented in Figure 8-8 are considered to remain unchanged throughout the operation. Thus, EMHE does not require to update the Gaussian mixture models online. On the other hand, RMHE provides the optimal Gaussian mixture models online and at each time interval to describe the densities of the last recent *N* process uncertainties. That is, the Gaussian mixture models of the process uncertainties changes at every time interval when performing RMHE.



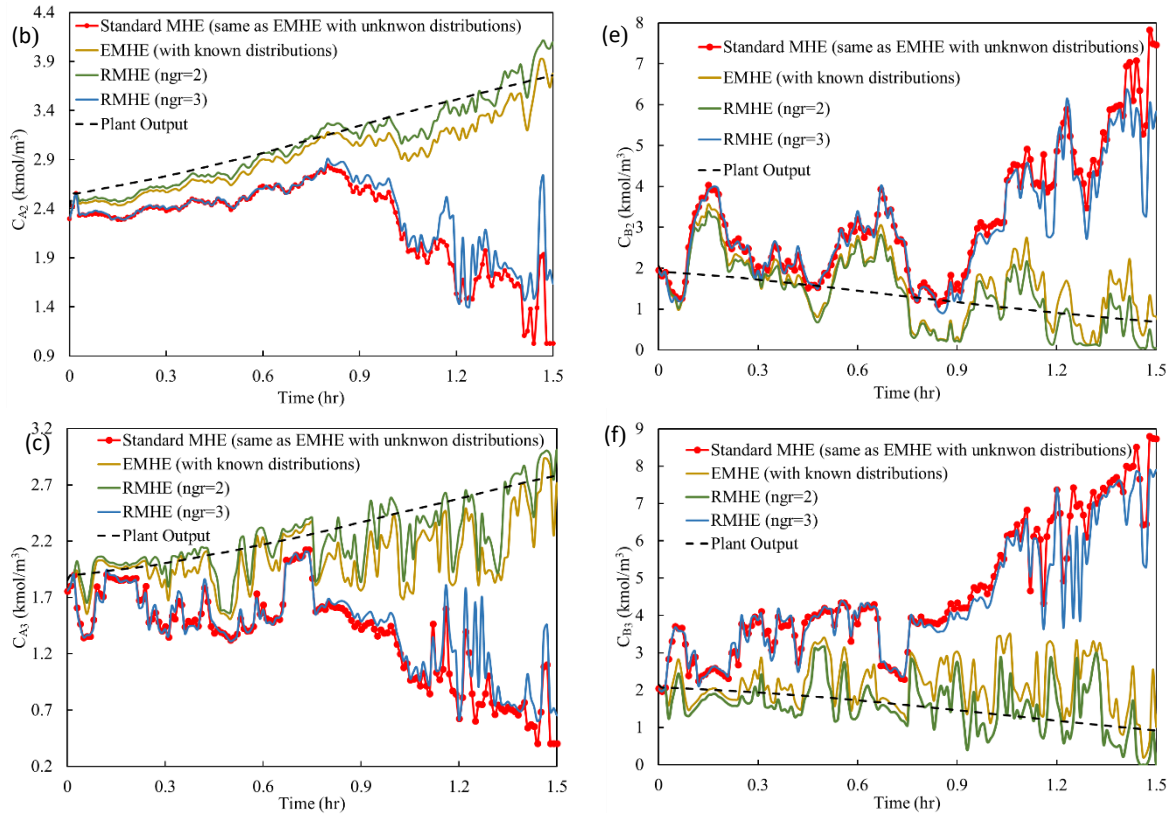


Figure 8-9: Estimation provided by the standard MHE, ideal EMHE (with known distributions), RMHE with $ngr = 2$ and $ngr = 3$ for (a) C_{A_1}; (b) C_{A_2}; (c) C_{A_3}; (d) C_{B_1}; (e) C_{B_2}; (f) C_{B_3}

Table 8-2: MSE for estimated states using different estimation schemes

Estimation method (n)	$MSE^{(n)}_{C_{A_1}}$	$MSE^{(n)}_{C_{B_1}}$	$MSE^{(n)}_{C_{A_2}}$	$MSE^{(n)}_{C_{B_2}}$	$MSE^{(n)}_{C_{A_3}}$	$MSE^{(n)}_{C_{B_3}}$
Standard MHE (same as EMHE)	1.44	0.283	1.36	7.34	1.34	14.02
EMHE (under known distributions)	0.06	0.005	0.05	0.59	0.13	0.93
RMHE ($ngr = 2$)	0.07	0.001	0.02	0.44	0.05	0.39
RMHE ($ngr = 3$)	0.87	0.204	1.05	5.66	1.13	11.75

As mentioned earlier in section 8.1, the proposed RMHE is an efficient estimation framework since it does not increase the computational costs considerably. For instance, the averaged CPU time required

by the standard MHE to provide the point estimate for the present case study was 0.66 s whereas the proposed RMHE only required 0.87 s on average to perform estimation for one time interval.

8.3 Summary

This chapter presented a novel MHE-based robust estimation scheme, referred to as *Robust Moving Horizon Estimation* (RMHE) to improve the accuracy in the estimation for practical scenarios when the distribution of the process uncertainties or measurement noises are not known *a priori* and follow a non-Gaussian behaviour. The proposed RMHE considers additional constraints and decision variables than in the standard MHE framework, which are needed to approximate the distributions of the uncertainties/noises to Gaussian mixture models online. Although the additional constraints and decision variables considered in the RMHE problem increases the computational effort than that required by the standard MHE scheme, the required CPU time in RMHE is not significantly larger than that in MHE. Therefore, RMHE offers an efficient scheme that increases the robustness of the estimation with respect to the unexpected noises/uncertainties occurring in the process. The computational experiments conducted in this work showed that the proposed RMHE offers a similar performance as that obtained in the standard MHE in the case of nominal plant operation in which both process uncertainties and measurement noises follow zero-mean Gaussian distributions.

Chapter 9
Conclusions and Recommendations

In this PhD thesis, novel methodologies for conventional and advanced state estimation schemes were developed and tested under operating scenarios that are expected to occur in practice. The key idea is to use the *abridged Gaussian sum* approach introduced in this thesis to represent the general class of non-Gaussian distribution of constrained states, process uncertainties, measurement noises. Hence, innovations have been introduced in the standard mathematical formulations of extended Kalman filter (EKF) and moving horizon estimation (MHE) to provide efficient state estimation schemes to capture non-Gaussianity involved in real-world processes. The key knowledge gaps are explained next followed by the novelties and contributions proposed to address the challenges identified in this area.

Both EKF and MHE schemes assume that the process uncertainties and measurement noises follow zero-mean Gaussian distributions; however, constrained nonlinear dynamic models involving non-zero mean non-Gaussian process uncertainties and measurement noises are better suited to capture the actual behavior of key chemical engineering systems. Studies considering such random uncertainties/ random noises in the EKF framework are lacking from the literature. Moreover, current constrained-EKF approaches for applications featuring constrained states are not computationally efficient because they often require online solution of optimization problems, implementation of sampling-based approaches or implementation of multiple EKFs (Gaussian sum filters).

Similarly, studies on MHE considering non-zero mean non-Gaussian distributions for process uncertainties and measurement noises are scarce in the literature. Moreover, the distribution of the random noises/uncertainties may change during the plant operation due to the scheduled or/and unscheduled changes occurring in the process plant. The current MHE-based schemes are not able to capture time-dependent known distributions for process uncertainties and measurement noises, while this is crucial for the scenarios involving the scheduled operating condition changes. Under the unscheduled operating condition changes, assuming that the distributions of the noises/uncertainties are known *a priori* may no longer hold. However, a robust MHE scheme that performs state estimation considering unknown non-Gaussian distributions for the process uncertainties or measurement noises is not available.

This PhD thesis sought to contribute addressing the knowledge gaps described above and improving widely used model-driven state estimation methods such as KF/EKF and MHE. The major outcomes and insights gained from this work are outlined next.

- This study investigated the challenges of state estimation for a large-scale gasification system using KF. An actual pilot-scale entrained-flow gasifier involving more than 200 states was considered. A plant-model mismatch, additive uncertainty in the prior estimation, and load-following scenarios have been considered. The results from this study showed that KF is capable of estimating the unknown states for a large variety of changes in the gasifier's inputs, even though online temperature sensors are only available in limited locations across the gasifier. However, the success of KF directly depends on the accuracy of the linear dynamic model that is used for the prior estimation of the states. The outcome of this study revealed that, as long as an accurate linear state space model of the process is available, KF can be employed as a practical tool for the purpose of online state estimation for large-scale applications.
- A novel approach referred to as *Abridged Gaussian Sum-Extended Kalman Filter* (AGS-EKF), and its extension referred to as *constrained AGS-EKF*, were developed to improve EKF performance for constrained nonlinear systems with non-Gaussian process uncertainties and measurement noises. The methods employ Gaussian mixture models to approximate non-Gaussian distributions present in the process, i.e., constrained states, non-Gaussian uncertainties and noises. The prior and posterior estimation steps in EKF are modified to capture non-Gaussian process uncertainties and measurement noises, respectively. An intermediate step is considered where the constraints are explicitly applied on the states, i.e., the non-Gaussian distribution of the prior states are approximated using Gaussian mixture models. While the modifications in the prior and posterior estimation step do not require any additional computational costs than those needed in the standard EKF, the intermediate estimation step requires additional CPU time to approximate the corresponding Gaussian mixture model of the states. Despite this additional cost, the constrained AGS-EKF is an efficient approach as the additional costs are not significant when compared to conventional GSFs. The proposed approach avoids projecting the Gaussian components in the mixture individually; accordingly, this estimation scheme resolves the issue of extensive computational burdens and biased estimates often observed in the GSF scheme. Supported by the computational experiments conducted to this work, the proposed constrained AGS-EKF scheme is computationally efficient and provides appropriate estimates for applications involving constraints on states, non-Gaussian process uncertainties and measurement noises.
- The performance of EKF as an arrival cost (AC) estimator for industrial and/or complex applications under both open-loop and closed-loop conditions was also investigated in this PhD thesis. The results showed that EKF provides adequate performance as AC estimator in the presence

of process constraints if it is properly initialized using a model-based estimation framework such as MHE, which can explicitly account for process constraints. Hence, EKF is an efficient and suitable AC estimator (even under non-symmetric bounded distributions in the process uncertainty) to maintain the operation of large and challenging systems in closed-loop using an MHE-NMPC framework. An error analysis on the convergence of the EKF-based AC estimator was performed. According to this error analysis, even though the estimation error in EKF scheme becomes unbounded in the presence of nonlinear process constraints in the system, the estimation error for EKF as the AC estimator remains bounded. The bounded estimation error for EKF in the closed-loop NMPC-MHE framework confirms the reliability of EKF to approximate the AC distribution.

- An extended version of MHE (EMHE) was developed to improve the estimation in the presence of general non-Gaussian process uncertainties and measurement noises. The EMHE problem requires the same computational costs as the standard MHE problem. To capture the non-Gaussian densities of noises/uncertainties, Gaussian mixture models are used in the EMHE scheme. The proposed EMHE-based estimation scheme can be updated online when the distributions of the noises/uncertainties change due to sudden or seasonal changes in the process operating conditions. Although the re-approximation of the Gaussian mixture models to adapt corresponding changes in the noises/uncertainties requires additional computation costs, the averaged CPU time of the proposed estimation scheme does not increase considerably. Thus, EMHE is an efficient state estimation scheme for the purpose of online monitoring and control of practical applications. Numerical studies showed that EMHE was able to capture both non-Gaussianity and scheduled changes in the distributions of the noises/uncertainties, which resulted in significant improvements in the estimation and online control.
- A novel robust MHE (RMHE) scheme that approximates the unknown non-Gaussian distributions of the random uncertainties/noises using an optimal Gaussian mixture model that is updated online was also developed in this thesis. The proposed RMHE includes a Gaussian mixture model explicitly into the MHE formulation to model the unknown non-Gaussian distribution of these random variables affecting the process. Therefore, the RMHE not only provides the optimal estimation of the states, but also returns an optimal Gaussian mixture model for each measurement noise (or process uncertainty) in the nonlinear dynamical system. The RMHE formulation involves additional constraints that represent the mean value and covariance matrix of the Gaussian mixture model of the noise/uncertainty. The computational experiments used to test the performance of these estimation framework revealed that the additional constraints and decision variables included

in RMHE do not significantly increase the required computational effort when compared to the standard MHE problem. Thus, the proposed RMHE is an efficient version of the MHE that offers robustness in the state estimation without a significant increase in the computational costs.

9.1 Recommendations for Future Research

The insights gained through this thesis have opened new research avenues that can be potentially considered for future work. The major recommendations derived from this research are as follows:

- The results presented in chapter 3 showed that using a full dynamic ROM in the EKF estimation for a complex system such as the gasifier is computationally demanding due to the high computational cost. Thus, model reduction techniques can be explored in the future to enable the implementation of advanced state estimation strategies in the gasification system, and their potential integration with optimization-based control strategies. For instance, an optimization-based closed-loop scheme, i.e., NMPC coupled with MHE, for a gasification system can be an interesting area of future research. To the author's knowledge, this aspect has not been explored in the open literature.
- The performance of the constrained AGS-EKF for a large-scale application (i.e., the gasifier involving ~ 500 states) was not assessed in this research. In such cases, the effect of non-Gaussian measurement noises and more importantly, non-Gaussian additive process uncertainties may intensify through the highly nonlinear complex process model, which makes the constrained AGS-EKF scheme more sensitive to the accuracy of the corresponding Gaussian mixture models. In addition, the multivariate Gaussian mixture model of constrained states for large-scale applications may become complex and consequently, approximating one unique multivariate Gaussian mixture model that can satisfy the constraints on all the states becomes a sophisticated and challenging task. The numerical complexity increases if the large-scale system is badly scaled. For those cases, one potential solution is to normalize the process model so that the covariance matrix of the constrained multivariate distribution of the states represent a smaller range of numerical values, which may help the EM algorithm to find the proper Gaussian mixture model while satisfying all the process constraints. Future research studies may consider the application of the constrained AGS-EKF framework for large-scale constrained nonlinear processes involving non-Gaussian measurement noises and process uncertainties.
- The constrained AGS-EKF introduced in this PhD thesis holds the underlying assumption that the distribution of the non-Gaussian measurement noises and process uncertainties can be known *a*

priori, which may not always be a valid assumption. As discussed in chapter 8, developing the state estimation schemes to increase the robustness is essential for applications involving unknown distributions of the noises or uncertainties. To further improve the features offered by the constrained AGS-EKF, a potential future work can focus on extending the AGS-EKF formulation and consider a robust AGS-EKF scheme to improve accuracy of the estimation when the process uncertainties and measurement noises follow unknown non-Gaussian distributions. Robust control tools successfully used for applications involving optimal process design and control[200,201] can be borrowed to consider the development of such robust estimation techniques.

- The outcome of the proposed EMHE framework have focused only on the non-Gaussian process uncertainties and measurement noises. The effect of the arrival cost was disregarded by considering a large estimation horizon. Future works can consider the state estimation for applications involving known non-Gaussian noises and uncertainties using EMHE coupled with the constrained AGS-EKF to evaluate the arrival cost parameters in the EMHE framework. The expecting outcome is to improve the accuracy of the estimation using short estimation horizon for general constrained applications with non-Gaussian process uncertainties and measurement noises.
- The RMHE framework introduced in this research is capable of providing an optimal Gaussian mixture model to describe the unknown distributions of either the process uncertainties or measurement noises. Future work may consider the development of a RMHE framework where both process uncertainty and measurement noise that follow unknown non-Gaussian distributions can be explicitly identified and handled by the estimation scheme. While this approach may imply the solution of mixed-integer nonlinear optimization problems, recent studies have shown that such problems could be transformed into nonlinear optimization problems using switching time techniques[202]. Hence, that approach could be potentially explored to develop computationally efficient formulations that may not require the need to solve online mixed-integer problems.
- One key step in the MHE framework is to have access to accurate models describing measurement noises, process uncertainties and arrival cost information. While multiple approaches have been suggested, including those presented in this PhD thesis, the application of machine learning algorithms using active learning can be a potential area of future development that can further improve the online estimation for large-scale applications, particularly those that emerge in the energy sectors[154,203,204,205,206]. Hence, future research can be focused on exploring the implementation of active learning methods for MHE applications.

- Integration of design and control is an area that has been widely explored over the last decades. While several simultaneous design and control methods are available in the literature[207,208,209,210,211], there is not a single study that has considered the impact of state estimation in the context of optimal design and control. Thus, it is recommended to perform such a study as a future work, which can provide new insights on the impact of state estimation for the optimal design of chemical systems.

Bibliography

(1) Becerra, V. M.; Roberts, P. D.; Griffiths, G. W. Applying the Extended Kalman Filter to Systems Described by Nonlinear Differential-Algebraic Equations. *Control Eng. Pract.* **2001**, *9* (3), 267.

(2) Karimipour, H.; Dinavahi, V. Extended Kalman Filter-Based Parallel Dynamic State Estimation. *IEEE Trans. Smart Grid* **2015**, *6* (3), 1539.

(3) Kandepu, R.; Foss, B.; Imsland, L. Applying the Unscented Kalman Filter for Nonlinear State Estimation. *J. Process Control* **2008**, *18* (7–8), 753.

(4) Wan, E. A.; Van Der Merwe, R.; Haykin, S. The Unscented Kalman Filter. *Kalman Filter. neural networks* **2001**, *5* (2007), 221.

(5) Romanenko, A.; Santos, L. O.; Afonso, P. A. Unscented Kalman Filtering of a Simulated PH System. *Ind. Eng. Chem. Res.* **2004**, *43* (23), 7531.

(6) Fuchigami, S.; Niina, T.; Takada, S. Particle Filter Method to Integrate High-Speed Atomic Force Microscopy Measurements with Biomolecular Simulations. *J. Chem. Theory Comput.* **2020**, *16* (10), 6609.

(7) Carpenter, J.; Clifford, P.; Fearnhead, P. Improved Particle Filter for Nonlinear Problems. *IEE Proceedings-Radar, Sonar Navig.* **1999**, *146* (1), 2.

(8) Zhou, D.; Wu, Y.; Gao, F.; Breaz, E.; Ravey, A.; Miraoui, A. Degradation Prediction of PEM Fuel Cell Stack Based on Multiphysical Aging Model with Particle Filter Approach. *IEEE Trans. Ind. Appl.* **2017**, *53* (4), 4041.

(9) Iori, T.; Ohtsuka, T. Recursive Elimination Method in Moving Horizon Estimation for a Class of Nonlinear Systems and Non-Gaussian Noise. *SICE J. Control. Meas. Syst. Integr.* **2020**, *13* (6), 282.

(10) Chen, B.; Hu, G. Nonlinear State Estimation under Bounded Noises. *Automatica* **2018**, *98*, 159.

(11) Monin, A. Modal Trajectory Estimation Using Maximum Gaussian Mixture. *IEEE Trans. Automat. Contr.* **2012**, *58* (3), 763.

(12) Calafiore, G. Reliable Localization Using Set-Valued Nonlinear Filters. *IEEE Trans. Syst. man, Cybern. A Syst. humans* **2005**, *35* (2), 189.

(13) Amor, N.; Bouaynaya, N. C.; Shterenberg, R.; Chebbi, S. On the Convergence of Constrained Particle Filters. *IEEE Signal Process. Lett.* **2017**, *24* (6), 858.

(14) Ungarala, S.; Dolence, E.; Li, K. Constrained Extended Kalman Filter for Nonlinear State Estimation. *IFAC Proc. Vol.* **2007**, *40* (5), 63.

(15) Yang, F.; Wang, Z.; Lauria, S.; Liu, X. Mobile Robot Localization Using Robust Extended H∞ Filtering. *Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng.* **2009**, *223* (8), 1067.

(16) Liu, X.; Chen, B.; Zhao, H.; Qin, J.; Cao, J. Maximum Correntropy Kalman Filter with State Constraints. *IEEE Access* **2017**, *5*, 25846.

(17) Kotecha, J. H.; Djuric, P. M. Gaussian Sum Particle Filtering. *IEEE Trans. signal Process.* **2003**, *51* (10), 2602.

(18) Afshar, P.; Nobakhti, A.; Wang, H. Joint Process and Quality Control of a Continuously Stirred Tank Reactor with Non-Gaussian Process Noise. *IET Control theory Appl.* **2012**, *6* (5), 651.

(19) Luders, B.; How, J. Probabilistic Feasibility for Nonlinear Systems with Non-Gaussian Uncertainty Using RRT. In *Infotech@ Aerospace 2011*; 2011; p 1589.

(20) Wang, G.; Li, N.; Zhang, Y. Maximum Correntropy Unscented Kalman and Information Filters for Non-Gaussian Measurement Noise. *J. Franklin Inst.* **2017**, *354* (18), 8659.

(21) Xu, H.; Xie, W.; Yuan, H.; Duan, K.; Liu, W.; Wang, Y. Fixed-Point Iteration Gaussian Sum Filtering Estimator with Unknown Time-Varying Non-Gaussian Measurement Noise. *Signal Processing* **2018**, *153*, 132.

(22) Stojanovic, V.; He, S.; Zhang, B. State and Parameter Joint Estimation of Linear Stochastic Systems in Presence of Faults and Non-Gaussian Noises. *Int. J. Robust Nonlinear Control* **2020**, *30* (16), 6683.

(23) An, T.; Yin, X.; Liu, J.; Forbes, J. F. Coordinated Distributed Moving Horizon State Estimation for Linear Systems Based on Prediction-driven Method. *Can. J. Chem. Eng.* **2017**, *95* (10), 1953.

(24) Karg, B.; Lucia, S. Approximate Moving Horizon Estimation and Robust Nonlinear Model Predictive Control via Deep Learning. *Comput. Chem. Eng.* **2021**, *148*, 107266.

(25) Yin, X.; Liu, J. Event-Triggered State Estimation of Linear Systems Using Moving Horizon Estimation. *IEEE Trans. Control Syst. Technol.* **2020**, *29* (2), 901.

(26) Elsheikh, M.; Hille, R.; Tatulea-Codrean, A.; Krämer, S. A Comparative Review of Multi-Rate Moving Horizon Estimation Schemes for Bioprocess Applications. *Comput. Chem. Eng.* **2021**, 107219.

(27) Schneider, R.; Georgakis, C. How to Not Make the Extended Kalman Filter Fail. *Ind. Eng. Chem.*

Res. **2013**, *52* (9), 3354.

(28) Haseltine, E. L.; Rawlings, J. B. Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation. *Ind. Eng. Chem. Res.* **2005**, *44* (8), 2451.

(29) Müller, M. A. Nonlinear Moving Horizon Estimation in the Presence of Bounded Disturbances. *Automatica* **2017**, *79*, 306.

(30) Liu, J. Moving Horizon State Estimation for Nonlinear Systems with Bounded Uncertainties. *Chem. Eng. Sci.* **2013**, *93*, 376.

(31) Ji, L.; Rawlings, J. B.; Hu, W.; Wynn, A.; Diehl, M. Robust Stability of Moving Horizon Estimation Under Bounded Disturbances. *IEEE Trans. Automat. Contr.* **2016**, *61* (11), 3509.

(32) Diehl, M.; Ferreau, H. J.; Haverbeke, N. Efficient Numerical Methods for Nonlinear MPC and Moving Horizon Estimation. In *Nonlinear model predictive control*; Springer, 2009; pp 391–417.

(33) Zavala, V. M.; Laird, C. D.; Biegler, L. T. A Fast Moving Horizon Estimation Algorithm Based on Nonlinear Programming Sensitivity. *J. Process Control* **2008**, *18* (9), 876.

(34) Valipour, M.; Toffolo, K. M.; Ricardez-Sandoval, L. A. State Estimation and Sensor Location for Entrained-Flow Gasification Systems Using Kalman Filter. *Control Eng. Pract.* **2021**, *108*, 104702.

(35) Valipour, M.; Ricardez-Sandoval, L. A. Abridged Gaussian Sum Extended Kalman Filter for Nonlinear State Estimation under Non-Gaussian Process Uncertainties. *Comput. Chem. Eng.* **2021**, *155*, 107534.

(36) Valipour, M.; Ricardez-Sandoval, L. A. Constrained Abridged Gaussian Sum Extended Kalman Filter: Constrained Nonlinear Systems with Non-Gaussian Noises and Uncertainties. *Ind. Eng. Chem. Res.* **2021**.

(37) Valipour, M.; Ricardez-Sandoval, L. A. Assessing the Impact of EKF as the Arrival Cost in the Moving Horizon Estimation under Nonlinear Model Predictive Control. *Ind. Eng. Chem. Res.* **2021**, *60* (7), 2994.

(38) Valipour, M.; Ricardez-Sandoval, L. A. Extended Moving Horizon Estimation for Chemical Processes under Non-Gaussian Noises. *AIChE J.* **2021**. https://doi.org/10.1002/aic.17545

(39) Valipour, M.; Ricardez-Sandoval, L. A. A Robust Moving Horizon Estimation under Unknown Distributions of Process or Measurement Noises. *Comput. Chem. Eng.* **2021**, 107620. https://doi.org/10.1016/j.compchemeng.2021.107620

(40) Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. **1960**.

(41) Kalman, R. E.; Bucy, R. S. New Results in Linear Filtering and Prediction Theory. **1961**.

(42) Roy, S. K.; Nicolson, A.; Paliwal, K. K. DeepLPC: A Deep Learning Approach to Augmented Kalman Filter-Based Single-Channel Speech Enhancement. *IEEE Access* **2021**, *9*, 64524.

(43) Ahmadi Jeyed, H.; Ghaffari, A. Nonlinear Estimator Design Based on Extended Kalman Filter Approach for State Estimation of Articulated Heavy Vehicle. *Proc. Inst. Mech. Eng. Part K J. Multi-body Dyn.* **2019**, *233* (2), 254.

(44) Meurer, F.; Neubert, M.; Werner, N. Nonlinear State Estimation for the Czochralski Process Based on the Weighing Signal Using an Extended Kalman Filter. *J. Cryst. Growth* **2015**, *419*, 57.

(45) Sahoo, S. R.; Yin, X.; Liu, J. Optimal Sensor Placement for Agro-hydrological Systems. *AIChE J.* **2019**, *65* (12), e16795.

(46) Yin, X.; Liu, J. State Estimation of Wastewater Treatment Plants Based on Model Approximation. *Comput. Chem. Eng.* **2018**, *111*, 79.

(47) Patwardhan, S. C.; Narasimhan, S.; Jagadeesan, P.; Gopaluni, B.; L. Shah, S. Nonlinear Bayesian State Estimation: A Review of Recent Developments. *Control Eng. Pract.* **2012**, *20* (10), 933.

(48) Mukherjee, T.; Varshney, D.; Kottakki, K. K.; Bhushan, M. Broyden's Update Based Extended Kalman Filter for Nonlinear State Estimation. *J. Process Control* **2021**, *105*, 267.

(49) Ricker, N. L.; Lee, J. Nonlinear Modeling and State Estimation for the Tennessee Eastman Challenge Process. *Comput. Chem. Eng.* **1995**, *19* (9), 983.

(50) 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* (1), 119.

(51) Agyeman, B. T.; Bo, S.; Sahoo, S. R.; Yin, X.; Liu, J.; Shah, S. L. Soil Moisture Map Construction by Sequential Data Assimilation Using an Extended Kalman Filter. *J. Hydrol.* **2021**, *598*, 126425.

(52) Calderón, C. J.; Ricardez-Sandoval, L.; Ancheyta, J. State Estimation for Heavy Oil Hydroprocessing Reactors Using Extended Kalman Filters. *Fuel* **2020**, *262*, 116565.

(53) Kallenberger, C.; Hamedović, H.; Zoubir, A. M. Comparison of the Extended Kalman Filter and the Unscented Kalman Filter for Parameter Estimation in Combustion Engines. *Eur. Signal Process. Conf.* **2007**, No. Eusipco, 2424.

(54) Zeng, J.; Liu, J.; Zou, T.; Yuan, D. Distributed Extended Kalman Filtering for Wastewater

Treatment Processes. *Ind. Eng. Chem. Res.* **2016**, *55* (28), 7720.

(55) Zapata, J. I. State Estimation of a Solar Direct Steam Generation Mono-Tube Cavity Receiver Using a Modified Extended Kalman Filtering Scheme. *Sol. Energy* **2015**, *114*, 152.

(56) Jayant Baliga, B. The IGBT Device. *IGBT Device* **2015**, *Chapter 19*.

(57) Zondervan, E.; Nawaz, M.; de Haan, A. B.; Woodley, J. M.; Gani, R. Optimal Design of a Multi-Product Biorefinery System. *Comput. Chem. Eng.* **2011**, *35* (9), 1752.

(58) Chávez-Islas, L. M.; Vásquez-Medrano, R.; Flores-Tlacuahuac, A. Optimal Synthesis of a High Purity Bioethanol Distillation Column Using Ionic Liquids. *Ind. Eng. Chem. Res.* **2011**, *50* (9), 5175.

(59) Mansouri, S. S.; Huusom, J. K.; Gani, R.; Sales-Cruz, M. Systematic Integrated Process Design and Control of Binary Element Reactive Distillation Processes. *AIChE J.* **2016**, *62* (9), 3137.

(60) López-Guajardo, E.; Ortiz-Nadal, E.; Montesinos-Castellanos, A.; Nigam, K. D. P. Process Intensification of Biodiesel Production Using a Tubular Micro-Reactor (TMR): Experimental and Numerical Assessment. *Chem. Eng. Commun.* **2017**, *204* (4), 467.

(61) Yuan, Y.; Dong, X.; Ricardez-Sandoval, L. A Multi-Scale Simulation of Syngas Combustion Reactions by Ni-Based Oxygen Carriers for Chemical Looping Combustion. *Appl. Surf. Sci.* **2020**, *531*, 147277.

(62) Yuan, Y.; You, H.; Ricardez-Sandoval, L. Recent Advances on First-Principles Modeling for the Design of Materials in CO2 Capture Technologies. *Chinese J. Chem. Eng.* **2019**, *27* (7), 1554.

(63) Wang, T.; Stiegel, G. J. *Integrated Gasification Combined Cycle (IGCC) Technologies*; Woodhead Publishing, 2016.

(64) Heidenreich, S.; Müller, M.; Foscolo, P. U. *Advanced Biomass Gasification: New Concepts for Efficiency Increase and Product Flexibility*; Academic Press, 2016.

(65) Nuamah, A.; Malmgren, A.; Riley, G.; Lester, E. 5.05 - Biomass Co-Firing. In *Comprehensive Renewable Energy; Oxford:heid*; Elsevier, 2012; pp 55–73.

(66) Hossein Sahraei, M.; McCalden, D.; Hughes, R.; Ricardez-Sandoval, L. A. A Survey on Current Advanced IGCC Power Plant Technologies, Sensors and Control Systems. *Fuel* **2014**, *137*, 245.

(67) Basu, P. Chapter 5 - Gasification Theory and Modeling of Gasifiers. In *Biomass gasification and pyrolysis: Practical design and theory*; Academic Press, 2010; pp 117–165.

(68) Miller, B. G. 4 - Introduction to Coal Utilization Technologies. In *Clean coal engineering*

technology; Elsevier, 2017; pp 147–229.

(69) Roberts, D. G.; Badwal, S. P. S.; Wibberley, L. J.; Bhattacharya, S. 9 - Gasification, DICE, and Direct Carbon Fuel Cells for Power, Fuels, and Chemicals Production from Low Rank Coals. In *Low-Rank Coals for Power Generation, Fuel and Chemical Production*; Woodhead Publishing, 2017; pp 217–237.

(70) Harris, D. J.; Roberts, D. G. Coal Gasification and Conversion. In *The coal handbook: towards cleaner production*; Elsevier, 2013; pp 427–454.

(71) Zhu, Y.; Frey, H. C. 3 - Integrated Gasification Combined Cycle (IGCC) Power Plant Design and Technology. In *Advanced Power Plant Materials, Design and Technology*; Woodhead Publishing, 2010; pp 54–88.

(72) Sahraei, M. H.; Duchesne, M. A.; Hughes, R. W.; Ricardez-Sandoval, L. A. Dynamic Reduced Order Modeling of an Entrained-Flow Slagging Gasifier Using a New Recirculation Ratio Correlation. *Fuel* **2017**, *196*, 520.

(73) Xu, C.; He, D.; Wang, H.; Guan, J.; Liu, C.; Peng, F.; Wang, W.; Kou, Z.; He, K.; Yan, X.; et al. Nano-Polycrystalline Diamond Formation under Ultra-High Pressure. *Int. J. Refract. Met. Hard Mater.* **2013**, *36*, 232.

(74) Reed, T. B.; Fahey, R. E.; Strauss, A. J. Sealed Crucible Technique for Thermal Analysis of Volatile Compounds up to 2500 °C: Melting Points of EuO, EuS, EuSe and EuTe. *J. Cryst. Growth* **1972**, *15* (3), 174.

(75) Sun, K.; Sur, R.; Chao, X.; Jeffries, J. B.; Hanson, R. K.; Pummill, R. J.; Whitty, K. J. TDL Absorption Sensors for Gas Temperature and Concentrations in a High-Pressure Entrained-Flow Coal Gasifier. *Proc. Combust. Inst.* **2013**, *34* (2), 3593.

(76) THORLABS. InGaAs Free-Space Amplified Photodetectors https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=4.

(77) Paul, P.; Bhattacharyya, D.; Turton, R.; Zitney, S. E. State Estimation of an Acid Gas Removal (AGR) Plant as Part of an Integrated Gasification Combined Cycle (IGCC) Plant with CO2 Capture. *29th Annu. Int. Pittsburgh Coal Conf. 2012, PCC 2012* **2012**, *1*, 758.

(78) Carrasco, R.; Sanchez, E. N.; Carlos-Hernandezy, S. Neural Network Identification for Biomass Gasification Considering Inputs. In *World Automation Congress 2012*; IEEE, 2012; pp 1–6.

(79) Huang, Q.; Paul, P.; Bhattacharyya, D.; Pillai, R. C.; Sabolsky, K.; Sabolsky, E. M. Estimations of

Gasifier Wall Temperature and Extent of Slag Penetration Using a Refractory Brick with Embedded Sensors. *Ind. Eng. Chem. Res.* **2017**, *56* (35), 9858.

(80) Marafioti, G.; Olaru, S.; Hovd, M. State Estimation in Nonlinear Model Predictive Control, Unscented Kalman Filter Advantages. In *Nonlinear Model Predictive Control*; Springer, 2009; pp 305–313.

(81) Prakash, J.; Patwardhan, S. C.; Shah, S. L. State Estimation and Nonlinear Predictive Control of Autonomous Hybrid System Using Derivative Free State Estimators. *J. Process Control* **2010**, *20* (7), 787.

(82) Dewasme, L.; Fernandes, S.; Amribt, Z.; Santos, L. O.; Bogaerts, P.; Wouwer, A. Vande. State Estimation and Predictive Control of Fed-Batch Cultures of Hybridoma Cells. *J. Process Control* **2015**, *30*, 50.

(83) Rao, C. V; Rawlings, J. B.; Mayne, D. Q. Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations. *IEEE Trans. Automat. Contr.* **2003**, *48* (2), 246.

(84) Rawlings, J. B.; Bakshi, B. R. Particle Filtering and Moving Horizon Estimation. *Comput. Chem. Eng.* **2006**, *30* (10–12), 1529.

(85) Simon, D.; Chia, T. L. Kalman Filtering with State Equality Constraints. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38* (1), 128.

(86) Vachhani, P.; Narasimhan, S.; Rengaswamy, R. Robust and Reliable Estimation via Unscented Recursive Nonlinear Dynamic Data Reconciliation. *J. Process Control* **2006**, *16* (10), 1075.

(87) Kolås, S.; Foss, B. A.; Schei, T. S. Constrained Nonlinear State Estimation Based on the UKF Approach. *Comput. Chem. Eng.* **2009**, *33* (8), 1386.

(88) López-Negrete, R.; Patwardhan, S. C.; Biegler, L. T. Constrained Particle Filter Approach to Approximate the Arrival Cost in Moving Horizon Estimation. *J. Process Control* **2011**, *21* (6), 909.

(89) Prakash, J.; Huang, B.; Shah, S. L. Recursive Constrained State Estimation Using Modified Extended Kalman Filter. *Comput. Chem. Eng.* **2014**, *65*, 9.

(90) Teixeira, B. O. S.; Tôrres, L. A. B.; Aguirre, L. A.; Bernstein, D. S. On Unscented Kalman Filtering with State Interval Constraints. *J. Process Control* **2010**, *20* (1), 45.

(91) Zhao, Z.; Huang, B.; Liu, F. Constrained Particle Filtering Methods for State Estimation of

Nonlinear Process. *AIChE J.* **2014**, *60* (6), 2072.

(92) Li, B.; Liu, C.; Chen, W.-H. An Auxiliary Particle Filtering Algorithm with Inequality Constraints. *IEEE Trans. Automat. Contr.* **2016**, *62* (9), 4639.

(93) Sorenson, H. W.; Alspach, D. L. Recursive Bayesian Estimation Using Gaussian Sums. *Automatica* **1971**, *7* (4), 465.

(94) Alspach, D.; Sorenson, H. Nonlinear Bayesian Estimation Using Gaussian Sum Approximations. *IEEE Trans. Automat. Contr.* **1972**, *17* (4), 439.

(95) Schoenberg, J. R.; Campbell, M.; Miller, I. Posterior Representation with a Multi-modal Likelihood Using the Gaussian Sum Filter for Localization in a Known Map. *J. F. Robot.* **2012**, *29* (2), 240.

(96) Terejanu, G.; Singla, P.; Singh, T.; Scott, P. D. Adaptive Gaussian Sum Filter for Nonlinear Bayesian Estimation. *IEEE Trans. Automat. Contr.* **2011**, *56* (9), 2151.

(97) Kottakki, K. K.; Bhartiya, S.; Bhushan, M. State Estimation of Nonlinear Dynamical Systems Using Nonlinear Update Based Unscented Gaussian Sum Filter. *J. Process Control* **2014**, *24* (9), 1425.

(98) Kottakki, K. K.; Bhushan, M.; Bhartiya, S. Constrained Unscented Gaussian Sum Filter for State Estimation of Nonlinear Dynamical Systems. *Comput. Chem. Eng.* **2016**, *91*, 352.

(99) Mohammadi, A.; Plataniotis, K. N. Complex-Valued Gaussian Sum Filter for Nonlinear Filtering of Non-Gaussian/Non-Circular Noise. *IEEE Signal Process. Lett.* **2014**, *22* (4), 440.

(100) Horwood, J. T.; Poore, A. B. Adaptive Gaussian Sum Filters for Space Surveillance. *IEEE Trans. Automat. Contr.* **2011**, *56* (8), 1777.

(101) Horwood, J. T.; Aragon, N. D.; Poore, A. B. Gaussian Sum Filters for Space Surveillance: Theory and Simulations. *J. Guid. Control. Dyn.* **2011**, *34* (6), 1839.

(102) DeMars, K. J.; Bishop, R. H.; Jah, M. K. Entropy-Based Approach for Uncertainty Propagation of Nonlinear Dynamical Systems. *J. Guid. Control. Dyn.* **2013**, *36* (4), 1047.

(103) Psiaki, M. L. Gaussian Mixture Nonlinear Filtering with Resampling for Mixand Narrowing. *IEEE Trans. Signal Process.* **2016**, *64* (21), 5499.

(104) Arbuckle, J. L. Full Information Estimation in the Presence of Incomplete Data. In *Advanced structural equation modeling: Issues and techniques*; G. A. Marcoulides, R. E. S., Ed.; 1996; pp 243–277.

(105) Rawlings, J. B. Moving Horizon Estimation. *Encycl. Syst. Control* **2013**, 1.

(106) Rao, C. V; Rawlings, J. B.; Mayne, D. Q. Constrained State Estimation for Nonlinear Discrete-Time Systems : Stability and Moving Horizon Approximations. *IEEE Trans. Automat. Contr.* **2003**, *48* (2), 246.

(107) Al-matouq, A. A.; Vincent, T. L. Multiple Window Moving Horizon Estimation ☆. *Automatica* **2015**, *53*, 264.

(108) Rawlings, J.B. and Mayne, D. Q. Model Predictive Control: Theory and Design. *Madison, Wisconsin Nob Hill Pub.* **2009**, *1912*, 3430.

(109) Rawlings, J. B.; Ji, L. Optimization-Based State Estimation : Current Status and Some New Results &. *J. Process Control* **2012**, *22* (8), 1439.

(110) Wynn, A.; Vukov, M.; Diehl, M. Convergence Guarantees for Moving Horizon Estimation Based on the Real-Time Iteration Scheme. *IEEE Trans. Automat. Contr.* **2014**, *59* (8), 2215.

(111) Ellis, M.; Zhang, J.; Liu, J.; Christofides, P. D. Systems & Control Letters Robust Moving Horizon Estimation Based Output Feedback Economic Model Predictive Control. *Syst. Control Lett.* **2014**, *68*, 101.

(112) Müller, M. A. Nonlinear Moving Horizon Estimation in the Presence of Bounded Disturbances. *Automatica* **2017**, *79*, 306.

(113) Ji, L.; Rawlings, J. B.; Hu, W.; Wynn, A.; Diehl, M. Robust Stability of Moving Horizon Estimation under Bounded Disturbances. *IEEE Trans. Automat. Contr.* **2015**, *61* (11), 3509.

(114) Liu, J. Moving Horizon State Estimation for Nonlinear Systems with Bounded Uncertainties. *Chem. Eng. Sci.* **2013**, *93*, 376.

(115) Alessandri, A.; Awawdeh, M. Moving-Horizon Estimation with Guaranteed Robustness for Discrete-Time Linear Systems and Measurements Subject to Outliers. *Automatica* **2016**, *67*, 85.

(116) Christian, G. To Linear Moving Horizon to Linear Moving Horizon. *IFAC-PapersOnLine* **2018**, *51* (20), 549.

(117) Qu, C. C.; Hahn, J. Computation of Arrival Cost for Moving Horizon Estimation via Unscented Kalman Filtering. *J. Process Control* **2009**, *19* (2), 358.

(118) Gharbi, M.; Ebenbauer, C. A Proximity Approach to Linear Moving Horizon Estimation. *IFAC-PapersOnLine* **2018**, *51* (20), 549.

(119) Tejeda-iglesias, M.; Lappas, N. H.; Gounaris, C. E.; Ricardez-sandoval, L. Explicit Model

Predictive Controller under Uncertainty : An Adjustable Robust Optimization Approach. *J. Process Control* **2019**, *84*, 115.

(120) Garriga, J. L.; Soroush, M. Model Predictive Control Tuning Methods: A Review. *Ind. Eng. Chem. Res.* **2010**, *49* (8), 3505.

(121) Rao, C. V. Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems. (Doctoral dissertation) University of Wisconsin--Madisonrao 2000.

(122) Zimmer, G. State Observation by On-Line Minimization. *Int. J. Control* **1994**, *60* (4), 595.

(123) Wilson, D. I.; Agarwal, M.; Rippin, D. W. T. Experiences Implementing the Extended Kalman Filter on an Industrial Batch Reactor. *Comput. Chem. Eng.* **1998**, *22* (11), 1653.

(124) Ungarala, S. Computing Arrival Cost Parameters in Moving Horizon Estimation Using Sampling Based Filters. *J. Process Control* **2009**, *19* (9), 1576.

(125) Rigatos, G. G. Particle Filtering for State Estimation in Nonlinear Industrial Systems. *IEEE Trans. Instrum. Meas.* **2009**, *58* (11), 3885.

(126) Lang, L.; Chen, W.; Bakshi, B. R.; Goel, P. K.; Ungarala, S. Bayesian Estimation via Sequential Monte Carlo Sampling—Constrained Dynamic Systems. *Automatica* **2007**, *43* (9), 1615.

(127) Arulampalam, M. S.; Maskell, S.; Gordon, N.; Clapp, T. A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE Trans. signal Process.* **2002**, *50* (2), 174.

(128) Kühl, P.; Diehl, M.; Kraus, T.; Schlöder, J. P.; Bock, H. G. A Real-Time Algorithm for Moving Horizon State and Parameter Estimation. *Comput. Chem. Eng.* **2011**, *35* (1), 71.

(129) Lao, L.; Ellis, M.; Durand, H.; Christofides, P. D. Real-time Preventive Sensor Maintenance Using Robust Moving Horizon Estimation and Economic Model Predictive Control. *AIChE J.* **2015**, *61* (10), 3374.

(130) Bae, H.; Oh, J.-H. Humanoid State Estimation Using a Moving Horizon Estimator. *Adv. Robot.* **2017**, *31* (13), 695.

(131) Yin, L.; Gao, H. Moving Horizon Estimation for ARMAX Processes with Additive Output Noise. *J. Franklin Inst.* **2019**, *356* (4), 2090.

(132) Arora, N.; Biegler, L. T. Redescending Estimators for Data Reconciliation and Parameter Estimation. *Comput. Chem. Eng.* **2001**, *25* (11–12), 1585.

(133) Albuquerque, J. S.; Biegler, L. T. Data Reconciliation and Gross-error Detection for Dynamic Systems. *AIChE J.* **1996**, *42* (10), 2841.

(134) Chen, X.; Pike, R. W.; Hertwig, T. A.; Hopper, J. R. Optimal Implementation of On-Line Optimization. *Comput. Chem. Eng.* **1998**, *22*, S435.

(135) Fuente, M. J.; Gutierrez, G.; Gomez, E.; Sarabia, D.; de Prada, C. Gross Error Management in Data Reconciliation. *IFAC-PapersOnLine* **2015**, *48* (8), 623.

(136) Johnston, L. P. M.; Kramer, M. A. Maximum Likelihood Data Rectification: Steady-state Systems. *AIChE J.* **1995**, *41* (11), 2415.

(137) Liu, J.; Su, Q.; Moreno, M.; Laird, C.; Nagy, Z.; Reklaitis, G. Robust State Estimation of Feeding–blending Systems in Continuous Pharmaceutical Manufacturing. *Chem. Eng. Res. Des.* **2018**, *134*, 140.

(138) Nicholson, B.; Lopez-Negrete, R.; Biegler, L. T. On-Line State Estimation of Nonlinear Dynamic Systems with Gross Errors. *Comput. Chem. Eng.* **2014**, *70*, 149.

(139) Ellis, M.; Zhang, J.; Liu, J.; Christofides, P. D. Robust Moving Horizon Estimation Based Output Feedback Economic Model Predictive Control. *Syst. Control Lett.* **2014**, *68*, 101.

(140) Zhang, J.; Liu, J. Lyapunov-based MPC with Robust Moving Horizon Estimation and Its Triggered Implementation. *AIChE J.* **2013**, *59* (11), 4273.

(141) Sahraei, M. H. Reduced Order Modeling and Scale-up of an Entrained Flow Gasifier. *Ph.D Diss. Chem. Eng., Univ. Waterloo* **2017**.

(142) Hossein Sahraei, M.; Duchesne, M. A.; Yandon, R.; Majeski, A.; Hughes, R. W.; Ricardez-Sandoval, L. A. Reduced Order Modeling of a Short-Residence Time Gasifier. *Fuel* **2015**, *161*, 222.

(143) Talley, P. Capitalizing on the Emerging Coal Gasification Market. *Aspen Technol. Webinar* **2007**.

(144) Monaghan, R. F. D.; Ghoniem, A. F. A Dynamic Reduced Order Model for Simulating Entrained Flow Gasifiers: Part I: Model Development and Description. *Fuel* **2012**, *91* (1), 61.

(145) Monaghan, R. F. D.; Ghoniem, A. F. A Dynamic Reduced Order Model for Simulating Entrained Flow Gasifiers. Part II: Model Validation and Sensitivity Analysis. *Fuel* **2012**, *94*, 280.

(146) Gazzani, M.; Manzolini, G.; MacChi, E.; Ghoniem, A. F. Reduced Order Modeling of the Shell-Prenflo Entrained Flow Gasifier. *Fuel* **2013**, *104*, 822.

(147) Li, C.; Dai, Z.; Sun, Z.; Wang, F. Modeling of an Opposed Multiburner Gasifier with a Reduced-Order Model. *Ind. Eng. Chem. Res.* **2013**, *52* (16), 5825.

(148) Lee, H. H.; Lee, J. C.; Joo, Y. J.; Oh, M.; Lee, C. H. Dynamic Modeling of Shell Entrained Flow

Gasifier in an Integrated Gasification Combined Cycle Process. *Appl. Energy* **2014**, *131*, 425.

(149) Bishop, G.; Welch, G. An Introduction to the Kalman Filter. *Proc SIGGRAPH, Course* **2001**, *8* (27599–23175), 41.

(150) Xu, L.; Xu, X.; Ren, X.; Guo, Y.; Feng, Y.; Yang, X. Stability Evaluation of High-Order Splitting Method for Incompressible Flow Based on Discontinuous Velocity and Continuous Pressure. *Adv. Mech. Eng.* **2019**, *11* (10), 1687814019855586.

(151) Paige, C. Properties of Numerical Algorithms Related to Computing Controllability. *IEEE Trans. Automat. Contr.* **1981**, *26* (1), 130.

(152) Oppenheim, A. V; Willsky, A. S.; Nawab, S. H.; Hernández, G. M. *Signals & Systems*; Pearson Educación, 1997.

(153) Henderson, C. Increasing the Flexibility of Coal-Fired Power Plants. *IEA Clean Coal Cent.* **2014**, *15*.

(154) Wang, H.; Ricardez-Sandoval, L. A. Dynamic Optimization of a Pilot-Scale Entrained-Flow Gasifier Using Artificial Recurrent Neural Networks. *Fuel* **2020**, *272*, 117731.

(155) Seliger, B.; Hanke-Rauschenbach, R.; Hannemann, F.; Sundmacher, K. Modelling and Dynamics of an Air Separation Rectification Column as Part of an IGCC Power Plant. *Sep. Purif. Technol.* **2006**, *49* (2), 136.

(156) Williams, T. J.; Otto, R. E. A Generalized Chemical Processing Model for the Investigation of Computer Control. *Trans. Am. Inst. Electr. Eng. Part I Commun. Electron.* **1960**, *79* (5), 458.

(157) Amrit, R.; Rawlings, J. B.; Biegler, L. T. Optimizing Process Economics Online Using Model Predictive Control. *Comput. Chem. Eng.* **2013**, *58*, 334.

(158) Bahakim, S. S.; Ricardez-sandoval, L. A. Simultaneous Design and MPC-Based Control for Dynamic Systems under Uncertainty : A Stochastic Approach. *Comput. Chem. Eng.* **2014**, *63*, 66.

(159) Vega, P.; Lamanna, R.; Revollar, S.; Francisco, M. Integrated Design and Control of Chemical Processes–Part II: An Illustrative Example. *Comput. Chem. Eng.* **2014**, *71*, 618.

(160) Roberts, S. J.; Husmeier, D.; Rezek, I.; Penny, W. Bayesian Approaches to Gaussian Mixture Modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20* (11), 1133.

(161) Fan, R.; Huang, R.; Diao, R. Gaussian Mixture Model-Based Ensemble Kalman Filter for Machine Parameter Calibration. *IEEE Trans. Energy Convers.* **2018**, *33* (3), 1597.

(162) Li, R.; Jan, N. M.; Huang, B.; Prasad, V. Constrained Multimodal Ensemble Kalman Filter Based on Kullback–Leibler (KL) Divergence. *J. Process Control* **2019**, *79*, 16.

(163) Yao, Y.; Chen, T.; Gao, F. Multivariate Statistical Monitoring of Two-Dimensional Dynamic Batch Processes Utilizing Non-Gaussian Information. *J. Process Control* **2010**, *20* (10), 1188.

(164) Bilmes, J. A. A Gentle Tutorial of the EM Algorithm and Its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. *Int. Comput. Sci. Inst.* **1998**, *4* (510), 126.

(165) Lee, G.; Scott, C. EM Algorithms for Multivariate Gaussian Mixture Models with Truncated and Censored Data. *Comput. Stat. Data Anal.* **2012**, *56* (9), 2816.

(166) Keshavarz, M.; Huang, B. Bayesian and Expectation Maximization Methods for Multivariate Change Point Detection. *Comput. Chem. Eng.* **2014**, *60*, 339.

(167) Biernacki, C.; Celeux, G.; Govaert, G. Choosing Starting Values for the EM Algorithm for Getting the Highest Likelihood in Multivariate Gaussian Mixture Models. *Comput. Stat. Data Anal.* **2003**, *41* (3–4), 561.

(168) Varshney, D.; Bhushan, M.; Patwardhan, S. C. State and Parameter Estimation Using Extended Kitanidis Kalman Filter. *J. Process Control* **2019**, *76*, 98.

(169) Roberts, P. D. An Algorithm for Steady-State System Optimization and Parameter Estimation. *Int. J. Syst. Sci.* **1979**, *10* (7), 719.

(170) Matias, J. O. A.; Le Roux, G. A. C. Real-Time Optimization with Persistent Parameter Adaptation Using Online Parameter Estimation. *J. Process Control* **2018**, *68*, 195.

(171) Marchetti, A. G. A New Dual Modifier-Adaptation Approach for Iterative Process Optimization with Inaccurate Models. *Comput. Chem. Eng.* **2013**, *59*, 89.

(172) Revollar, S.; Vega, P.; Vilanova, R.; Francisco, M. Optimal Control of Wastewater Treatment Plants Using Economic-Oriented Model Predictive Dynamic Strategies. *Appl. Sci.* **2017**, *7* (8), 813.

(173) Cristea, S.; de Prada, C.; Sarabia, D.; Gutiérrez, G. Aeration Control of a Wastewater Treatment Plant Using Hybrid NMPC. *Comput. Chem. Eng.* **2011**, *35* (4), 638.

(174) 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* (2), 143.

(175) Vachhani, P.; Rengaswamy, R.; Gangwal, V.; Narasimhan, S. Recursive Estimation in Constrained Nonlinear Dynamical Systems. *AIChE J.* **2005**, *51* (3), 946.

(176) Zambare, N.; Soroush, M.; Grady, M. C. Real-time Multirate State Estimation in a Pilot-scale Polymerization Reactor. *AIChE J.* **2002**, *48* (5), 1022.

(177) Calderón, C.; Alzate, A.; Gómez, L.; Alvarez, H. State Controllabillity Analysis and Re-Design for a Wastewater Treatment Plant. In *2012 20th Mediterranean Conference on Control & Automation (MED)*; IEEE, 2012; pp 776–781.

(178) Francisco, M.; Skogestad, S.; Vega, P. Model Predictive Control for the Self-Optimized Operation in Wastewater Treatment Plants: Analysis of Dynamic Issues. *Comput. Chem. Eng.* **2015**, *82*, 259.

(179) Han, H.; Qiao, J. Nonlinear Model-Predictive Control for Industrial Processes: An Application to Wastewater Treatment Process. *IEEE Trans. Ind. Electron.* **2013**, *61* (4), 1970.

(180) Busch, J.; Elixmann, D.; Kühl, P.; Gerkens, C.; Schlöder, J. P.; Bock, H. G.; Marquardt, W. State Estimation for Large-Scale Wastewater Treatment Plants. *Water Res.* **2013**, *47* (13), 4774.

(181) Flores-Tlacuahuac, A.; Biegler, L. T.; Saldívar-Guerra, E. Dynamic Optimization of HIPS Open-Loop Unstable Polymerization Reactors. *Ind. Eng. Chem. Res.* **2005**, *44* (8), 2659.

(182) Flores-Tlacuahuac, A.; Biegler, L. T. Integrated Control and Process Design during Optimal Polymer Grade Transition Operations. *Comput. Chem. Eng.* **2008**, *32* (11), 2823.

(183) Patil, B. P.; Maia, E.; Ricardez-Sandoval, L. A. Integration of Scheduling, Design, and Control of Multiproduct Chemical Processes under Uncertainty. *AIChE J.* **2015**, *61* (8), 2456.

(184) Moore, T.; Stouch, D. A Generalized Extended Kalman Filter Implementation for the Robot Operating System. In *Intelligent Autonomous Systems 13*; Menegatti, E., Michael, N., Berns, K., Yamaguchi, H., Eds.; Springer International Publishing: Cham, 2016; pp 335–348.

(185) Djuric, P. M.; Kotecha, J. H.; Zhang, J.; Huang, Y.; Ghirmai, T.; Bugallo, M. F.; Miguez, J. Particle Filtering. *IEEE Signal Process. Mag.* **2003**, *20* (5), 19.

(186) Doucet, A.; Johansen, A. M. A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later. *Handb. nonlinear Filter.* **2009**, *12* (656–704), 3.

(187) Segovia, P.; Rajaoarisoa, L.; Nejjari, F.; Duviella, E.; Puig, V. Model Predictive Control and Moving Horizon Estimation for Water Level Regulation in Inland Waterways. *J. Process Control* **2019**, *76*, 1.

(188) Christofides, P. D.; Scattolini, R.; de la Pena, D. M.; Liu, J. Distributed Model Predictive Control: A Tutorial Review and Future Research Directions. *Comput. Chem. Eng.* **2013**, *51*, 21.

(189) Hartley, E. N.; Trodden, P. A.; Richards, A. G.; Maciejowski, J. M. Model Predictive Control System Design and Implementation for Spacecraft Rendezvous. *Control Eng. Pract.* **2012**, *20* (7), 695.

(190) Srinivasan, B.; Palanki, S.; Bonvin, D. Dynamic Optimization of Batch Processes: I. Characterization of the Nominal Solution. *Comput. Chem. Eng.* **2003**, *27* (1), 1.

(191) Biegler, L. T.; Cervantes, A. M.; Wächter, A. Advances in Simultaneous Strategies for Dynamic Process Optimization. *Chem. Eng. Sci.* **2002**, *57* (4), 575.

(192) Iratni, A.; Chang, N.-B. Advances in Control Technologies for Wastewater Treatment Processes: Status, Challenges, and Perspectives. *IEEE/CAA J. Autom. Sin.* **2019**, *6* (2), 337.

(193) Assassa, F.; Marquardt, W. Dynamic Optimization Using Adaptive Direct Multiple Shooting. *Comput. Chem. Eng.* **2014**, *60*, 242.

(194) Lopez Negrete de la Fuente, R. Nonlinear Programming Sensitivity Based Methods for Constrained State Estimation, Carnegie Mellon University, 2011.

(195) Sun, Y.; El-Farra, N. H. Quasi-Decentralized Model-Based Networked Control of Process Systems. *Comput. Chem. Eng.* **2008**, *32* (9), 2016.

(196) Mhaskar, P.; Liu, J.; Christofides, P. D. *Fault-Tolerant Process Control: Methods and Applications*; Springer Science & Business Media, 2012.

(197) Wu, Z.; Luo, J.; Rincon, D.; Christofides, P. D. Machine Learning-Based Predictive Control Using Noisy Data: Evaluating Performance and Robustness via a Large-Scale Process Simulator. *Chem. Eng. Res. Des.* **2021**, *168*, 275.

(198) Wu, Z.; Rincon, D.; Luo, J.; Christofides, P. D. Machine Learning Modeling and Predictive Control of Nonlinear Processes Using Noisy Data. *AIChE J.* **2021**, *67* (4), e17164.

(199) Ellis, M.; Christofides, P. D. Real-time Economic Model Predictive Control of Nonlinear Process Systems. *AIChE J.* **2015**, *61* (2), 555.

(200) Ricardez-Sandoval, L. A.; Budman, H. M.; Douglas, P. L. Application of Robust Control Tools to the Simultaneous Design and Control of Dynamic Systems. *Ind. Eng. Chem. Res.* **2009**, *48* (2), 801.

(201) Sandoval, L. A. R.; Budman, H. M.; Douglas, P. L. Simultaneous Design and Control of Processes under Uncertainty: A Robust Modelling Approach. *J. Process Control* **2008**, *18* (7–8), 735.

(202) Andrés-Martínez, O.; Ricardez-Sandoval, L. A. A Switched System Formulation for Optimal

Integration of Scheduling and Control in Multi-Product Continuous Processes. *J. Process Control* **2021**, *106*, 94.

(203) Rangel-Martinez, D.; Nigam, K. D. P.; Ricardez-Sandoval, L. A. Machine Learning on Sustainable Energy: A Review and Outlook on Renewable Energy Systems, Catalysis, Smart Grid and Energy Storage. *Chem. Eng. Res. Des.* **2021**.

(204) Liu, Y.; Guo, B.; Zou, X.; Li, Y.; Shi, S. Machine Learning Assisted Materials Design and Discovery for Rechargeable Batteries. *Energy Storage Mater.* **2020**.

(205) Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.-L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine Learning Methods for Solar Radiation Forecasting: A Review. *Renew. Energy* **2017**, *105*, 569.

(206) Guan, Y.; Chaffart, D.; Liu, G.; Tan, Z.; Zhang, D.; Wang, Y.; Li, J.; Ricardez-Sandoval, L. Machine Learning in Solid Heterogeneous Catalysis: Recent Developments, Challenges and Perspectives. *Chem. Eng. Sci.* **2021**, 117224.

(207) Rafiei, M.; Ricardez-Sandoval, L. A. Stochastic Back-off Approach for Integration of Design and Control under Uncertainty. *Ind. Eng. Chem. Res.* **2018**, *57* (12), 4351.

(208) Rafiei, M.; Ricardez-Sandoval, L. A. A Trust-region Framework for Integration of Design and Control. *AIChE J.* **2020**, *66* (5), e16922.

(209) Diangelakis, N. A.; Burnak, B.; Katz, J.; Pistikopoulos, E. N. Process Design and Control Optimization: A Simultaneous Approach by Multi-parametric Programming. *AIChE J.* **2017**, *63* (11), 4827.

(210) Flores-Tlacuahuac, A.; Biegler, L. T. Simultaneous Mixed-Integer Dynamic Optimization for Integrated Design and Control. *Comput. Chem. Eng.* **2007**, *31* (5–6), 588.

(211) Kookos, I. K.; Perkins, J. D. An Algorithm for Simultaneous Process Design and Control. *Ind. Eng. Chem. Res.* **2001**, *40* (19), 4079.

(212) Dinov, I. D. Expectation Maximization and Mixture Modeling Tutorial. **2008**.

(213) Huang, Y.; Djuric, P. M. A Hybrid Importance Function for Particle Filtering. *IEEE Signal Process. Lett.* **2004**, *11* (3), 404.

(214) Meurer, A.; Smith, C. P.; Paprocki, M.; Čertík, O.; Kirpichev, S. B.; Rocklin, M.; Kumar, Am.; Ivanov, S.; Moore, J. K.; Singh, S. SymPy: Symbolic Computing in Python. *PeerJ Comput. Sci.* **2017**, *3*, e103.

Appendix A
Expectation-maximization algorithm

This appendix provides a brief discussion on the Expectation-Maximization (EM) algorithm. Consider a d-dimension Gaussian mixture model that approximates the vector-valued non-Gaussian variable ($\boldsymbol{\varpi}$), as presented in Equations (4-11)-(4-12). Let $z = (z_1, \dots, z_{ng})$ represent a vector of ng binary indicator variables to identity of the Gaussian components of the Gaussian mixture model. The first step in EM algorithm is to draw NS number of samples from the original non-Gaussian multivariate distribution of the vector-variable $\boldsymbol{\varpi}$. The possibility that a sample $\boldsymbol{\varpi}_s$ follow the i^{th} Gaussian component in the mixture is mw_{is}; where mw is referred to as the *membership weight*. An adequate initial guess for the Gaussian components' parameters (i.e., $\boldsymbol{\eta}^i$, $\boldsymbol{M}^i$, and α^i for each i Gaussian component) should be provided to the EM algorithm to begin the parameter estimation. Then, *E-step* and *M-step* are performed in a recursive fashion to estimate the mean value, covariance, and weight corresponded to each Gaussian component. EM algorithm converges once the log-likelihood does not change significantly from one iteration to the next one. The required computations in E-step and M-step are presented next.

Expectation step (E-step): Equation (A-1) aims to evaluate mw_{is} associated with the sample $\boldsymbol{\varpi}_s$ in cluster i. the membership weight is required to be determined for all the samples and all the Gaussian components in the mixture, as follows[164,212]:

$$mw_{is} = \frac{p_i\left(\boldsymbol{\varpi}_s \middle| z_i; \left\{\boldsymbol{\eta}^i, \boldsymbol{M}^i\right\}\right) \alpha^i}{\sum_{m=1}^{ng} p_m\left(\boldsymbol{\varpi}_s \middle| z_m; \left\{\boldsymbol{\eta}^m, \boldsymbol{M}^m\right\}\right) \alpha^m}; \quad \forall\, i = 1, \dots, ng\ ; \ \ \forall\, s = 1, \dots, NS$$
$$\sum_{i=1}^{ng} mw_{is} = 1; \qquad \forall\, i = 1, \dots, ng\ ; \ \ \forall\, s = 1, \dots, NS \tag{A-1}$$

where the probability density function of $\boldsymbol{\varpi}_s$ given the i^{th} Gaussian component is[164,212]:

$$p_i\left(\boldsymbol{\varpi}_s \middle| z_i; \left\{\boldsymbol{\eta}^i, \boldsymbol{M}^i\right\}\right) = \frac{1}{(2\pi)^{d/2} |\boldsymbol{M}^i|^{1/2}} e^{-\frac{1}{2}\left(\boldsymbol{\varpi}_s - \boldsymbol{\eta}^i\right)^T \boldsymbol{M}^{i^{-1}} \left(\boldsymbol{\varpi}_s - \boldsymbol{\eta}^i\right)} \tag{A-2}$$

Maximization step (M-step): The membership weights provided by Equation (A-1) are used to re-approximate the corresponding weight, mean value, and covariance of each Gaussian component in the Gaussian mixture model. To pursue this goal, Equations (A-3), (A-4), and (A-5) are required to be evaluated as ordered[164,212].

$$\alpha^{i^{(new)}} = \frac{\sum_{s=1}^{NS} mw_{is}}{NS} \tag{A-3}$$

$$\boldsymbol{\eta}^{i^{(new)}} = \left(\frac{1}{\sum_{s=1}^{NS} mw_{is}}\right) \sum_{s=1}^{NS} (mw_{is}.\boldsymbol{\varpi}_s)\,; \tag{A-4}$$

$$\boldsymbol{M}^{i^{(new)}} = \left(\frac{1}{\sum_{s=1}^{NS} mw_{is}}\right) \sum_{s=1}^{NS} mw_{is} \left(\boldsymbol{\varpi}_s - \boldsymbol{\eta}^{i^{(new)}}\right)\left(\boldsymbol{\varpi}_s - \boldsymbol{\eta}^{i^{(new)}}\right)^T\,; \tag{A-5}$$

$\forall\ s = 1, \ldots, NS$

Appendix B

Nominal steady-state value for different case studies

This appendix provides important information of the Williams-Otto reactor, WTP, and the CSTRs-Separator network. Table B-1 presents the nominal steady-state value of the states in the Williams-Otto reactor.

Table B-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

Likewise, Table B-2 presents the nominal steady-state value of the states, the capacity of bioreactor V_r and the area of the decanter A_d. Note that when performing the closed-loop for WTP, the lower bounds for q_p and q_2 are set to zero, whereas their upper bound is set to 600.

Table B-2: Nominal values of states for WTP

Process Variables	Base case value
x_w	280.18 (mg/L)
s_w	100.02 (mg/L)
x_d	102.03 (mg/L)
x_b	951.73 (mg/L)
x_r	5975.82 (mg/L)
c_w	0.08 (mg/L)
c_w	0.08 (mg/L)
V_r	2500.00 (m^3)
A_d	1100.00 (m^2)

Likewise, Table B-3 presents the nominal steady-state value of the states in the CSTRs-Separator network.

Table B-3: Nominal values of states for the CSTRs-Separator network

Process Variables	Base case value
T_1	380 (K)

C_{A_1}	2.67 (kmol/m^3)
C_{B_1}	2.15 (kmol/m^3)
T_2	380 (K)
C_{A_2}	2.42 (kmol/m^3)
C_{B_2}	2.06 (kmol/m^3)
T_3	390 (K)
C_{A_3}	1.85 (kmol/m^3)
C_{B_3}	2.15 (kmol/m^3)

Appendix C

Supplementary material for WTP study in section 6.2.1

In addition to the information presented in Table B-2, WTP case study considers the following assumptions:

i) The observability for the states s_w, x_d, and c_w was confirmed by checking that the observability matrix is full rank (not shown for brevity).

ii) The length of estimation (N), prediction (L) and control (C) horizon were set to 3, 10, 5, respectively, and the time interval is 1 hour. *N* has to be small enough such that the effect of the AC method in the MHE estimation can be considered. Preliminary tests involving multiple combinations between L and C were performed to obtain the appropriate prediction and control horizons in the NMPC framework. The corresponding horizons considered in this work presented acceptable closed-loop performance at the nominal conditions reported in Table B-2.

iii) There are two standard choices of importance function in C-PF (or in overall PF), namely the posterior and the prior[213]. In the current study, the prior importance function has been considered, which is the most frequently used importance function due to its simplicity. This probability density function has a zero mean bounded distribution that is assigned to the process uncertainties ($\boldsymbol{w}_j$). That is, $\boldsymbol{wp}_j \in \mathbb{R}^{NP \times n_x}$ are the samples that are drawn randomly from the probability distribution assigned to the process uncertainties at j^{th} time interval ($\boldsymbol{w}_j$). The samples are selected using Monte Carlo sampling techniques. These samples are imposed on the estimated states calculated from MHE at the time step $k-N-1$ ($\widehat{\boldsymbol{x}}_{k-N-1}$) to generate the NP particles used as the prior estimation of the states in the C-PF estimation. Moreover, it is critical to select a large enough number of particles to represent an acceptable approximation of the actual distribution of the states of the system while keeping reasonable computational costs. Preliminary simulations showed that 100 particles provided an acceptable representation of the states in the C-PF method at reasonable computational costs.

iv) Both the MHE and NMPC are subject to bounds on the MVs (see Table B-2).

v) Step changes in the set-point for the biomass concentration x_w were considered. Note that for the case of symmetric bounded process uncertainties, set points of CVs have not been considered the same as those used for non-symmetric bounded process uncertainties. This is because process uncertainties are identified as model structural errors. Due to the non-symmetric process

uncertainty tested in this work, the plant may not operate around the nominal operating point depicted in Table B-2. Thus, in order to make the WTP dynamically operable in closed-loop, the set points of CVs were modified for the case of non-symmetric bounded uncertainty.

vi) The weight matrix for the CVs is as follows: $\boldsymbol{Q}_{\boldsymbol{out}}^{T} = diag([1,200,1])$. The diagonal elements are the weights for x_w and s_w and c_w, respectively. These weights have been tuned based on preliminary tests. Note that x_w and is at least two orders of magnitude higher than s_w; hence, the larger weight on s_w.

vii) Weights on the MVs were not considered in the NMPC framework to simplify the analysis.

WTP: Scenarios I-III: Open-loop estimation

For the open-loop operation, all the input variables remained constant and equal to their nominal values reported in Table B-2 in section B. Table C-1 summarizes the results obtained from *Scenarios I* and *Scenario II* during open-loop operation. As shown in Table C-1, the highest $\widehat{SSE}^{(n)}$ corresponds to MHE results using 0.5TS whereas the lowest $\widehat{SSE}^{(n)}$ is that obtained from using TS as the AC method. In *Scenario I*, using C-PF as the AC estimator slightly reduces the estimation error in MHE when compared to an MHE estimation associated with EKF as the AC estimator (~1% reduction). *Scenario II* considers non-symmetric bounded distribution; hence, it was expected that EKF may not perform well since the process uncertainty significantly deviates from a Gaussian distribution. Nevertheless, the results obtained for this scenario using the different approximation methods showed no significant differences in performance. This is mostly because the process uncertainty has zero mean value with narrower bounds; hence, process uncertainties may not significantly impact open-loop operation.

Table C-1: $\widehat{SSE}^{(n)}$ Open-loop operation of the WTP, *Scenarios I-II*

AC estimation method (*n*)	*Scenario I*	*Scenario II*
TS	0.0165	0.0208
0.5TS	1.0000	1.0000
EKFexpc	0.0409	0.0216
C-PFexpc	0.0306	0.0228
EKF	0.0382	0.0217
C-PF	0.0305	0.0238

Table C-2 presents the averaged CPU time needed to execute *Scenario I* and *Scenario II* at each time interval, respectively. As shown in this table, the required CPU time for *Scenario I* in the case of using EKF is almost 24% smaller than that using C-PF. Similarly, for *Scenario*

II, using EKF as the AC estimator reduces the required CPU time by almost 20% in comparison to the case of using C-PF. Since minimizing computational expenses is a critical element in an online estimation and control framework, the performance improvement achieved by using C-PF (measured in this work as a function of $\widehat{SSE}^{(n)}$) is not significant enough to make the computational efforts worthwhile. Note that the averaged CPU times reported for the cases of TS and 0.5TS can be used as benchmarks to determine the additional computational expenses required to apply an AC estimation method, i.e., no additional calculations are needed for those two AC estimation methods. For instance, compared to the benchmark, an additional 1.71s and 1.78s are required on average to estimate the AC parameters using EKFexpc for *Scenario I* and *Scenario II*, respectively.

Table C-2: Averaged CPU time (s), *Scenarios I* and *II*

AC approximation approach	*Scenario I*	*Scenario II*
TS	1.85	1.86
0.5TS	1.88	1.87
EKFexpc	3.56	3.63
C-PFexpc	4.41	4.36
EKF	3.78	3.77
C-PF	4.37	4.42

Table C-3 compares the MHE performance obtained from the different instances considered in *Scenario III*. As presented in Table 6-1 in section 6.2.1, this scenario aims to perform open-loop MHE estimation using smaller and larger plant designs while considering both symmetric and non-symmetric bounded process uncertainties. A comparison between $\widehat{SSE}^{(n)}$ for *Scenario III.A* and *Scenario III.B* shows the effect of different plant designs under symmetric bounded distributions. As shown in Table C-3, the $\widehat{SSE}^{(n)}$ reported for TS, EKFexpc, C-PF-excpc, EKF, and C-PF in *Scenario III.B* are larger than those obtained for *Scenario III.A*. For a fixed estimation horizon, a higher plant capacity (*Scenario III.B*) is expected to be more dependent to the AC since the response time of the process is larger in comparison with a smaller plant design (*Scenario III.A*). Therefore, a more significant loss in open-loop performance is expected in the absence of an acceptable AC approximation method (e.g., 0.5TS) for the case of a larger plant design, as shown in Figure C-1. Note however that the normalized SSE ($\widehat{SSE}^{(n)}$) reported in Table C-3 are smaller for larger plant sizes (*Scenario III.B*) than those

obtained for a smaller plant size (*Scenario III.A*). This is because of the normalization performed using the SSE reported for 0.5TS, which resulted in the worst estimation method for both scenarios. As a result, the need of using an adequate AC approximation method becomes more important as the plant capacity increases (i.e., *Scenario III.B*). Moreover, the effect of plant design in the case of non-symmetric bounded uncertainties, i.e., *Scenarios III.C* and *III.D*, is depicted in Table C-3. This table shows that, regardless of the plant capacity, the estimation during open-loop operation remains fairly the same for all the AC approximation methods.

Table C- 3: $\widehat{SSE}^{(n)}$ Open-loop MHE estimation for the WTP, *Scenario III*

AC approximation approach		Scenarios		
	III.A	III.B	III.C	III.D
TS	0.0301	0.0109	0.0248	0.0217
0.5TS	1.0000	1.0000	1.0000	1.0000
EKFexpc	0.0715	0.0210	0.0261	0.0256
C-PFexpc	0.0558	0.0223	0.0260	0.0249
EKF	0.0693	0.0213	0.0295	0.0258
C-PF	0.0559	0.0222	0.0262	0.0256

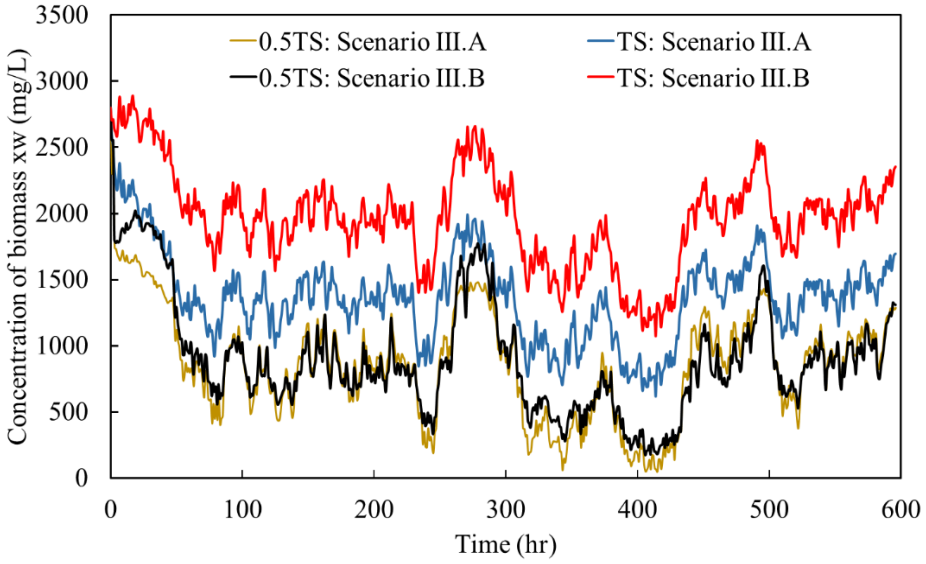


Figure C-1: Open-loop MHE estimation, *Scenario III.A and III.B*

Based on the above, both EKF and C-PF returned similar performances in terms of SSE. However, EKF is capable of providing an accurate AC approximation during open-loop operation at lower computational costs.

Appendix D

Supplementary material for HIPS study in section 6.2.2

This section presents the information and assumptions that have been considered in the HIPS case study. Table D-1 shows the main parameters of HIPS process that has been adopted in the present work for this case study. The manipulated variables (MVs) considered in NMPC framework for this process are the outlet flowrate (Q) and the cooling water flowrate (Q_{cw}). The lower and upper bounds for these two MVs are presented in Table D-1.

Table D-1: Nominal values of states at steady-state, and model parameters for HIPS

Process Variable	Scaled case value (**SCV**)	Base case value	
C_i	0.61	6.14e-05 (mol/L)	
C_m	1.00	6.07 (mol/L)	
C_b	1.00	1.05 (mol/L)	
C_r	2.51e-11	2.51e-11 (mol/L)	
C_{br}	2.24e-12	2.24e-12 (mol/L)	
T	0.78	389.31 (K)	
T_j	0.64	320.52 (K)	
μ_r^0	4.55	4.55e-08 (mol/L)	
μ_b^0	1.97e-09	1.97e-09 (mol/L)	
Q^l, Q^u	0-1	0-1.14 (L/s)	
Q_{cw}^l, Q_{cw}^u	0-1	0-1 (L/s)	
Process Parameters	Value	Process Parameters	Value
C_i^f	0.9815 (mol/L)	Q_i^s	0.0015 (L/s)
C_m^f	8.63 (mol/L)	Q_{cw}^s	1 (L/s)
C_b^f	1.0548 (mol/L)	Q^s	1.1412 (L/s)
T^f	333 (K)	U	80 (J/(s.K.m^2))
T_j^f	350 (K)	A_H	19.5 (m^2)
Q_i	0.0015 (L/s)	V	94.50 (L)
C_i^s	0.0001 (mol/L)	ρ_s	0.9150 (Kg/L)
C_m^s	6.0723 (mol/L)	C_{ps}	1647.265 (J/kg.K)
C_b^s	1.0545 (mol/L)	e_f	0.57
C_r^s	1 (mol/L)	ρ_{cw}	1 (Kg/L)
C_{br}^s	1 (mol/L)	C_{ps_w}	4045.7048 (J/kg.K)
T^s	500 (K)	V_c	2.000 (L)
$\mu_r^{0^s}$	1 (mol/L)	ΔH_r	69919.56 (J/mol)
$\mu_b^{0^s}$	1 (mol/L)	R	1.9858 cal/(mol.K)

The underlying assumptions considered for the HIPS case study are as follows:

i) The observability matrix for the states C_b, C_r, C_{br}, T and μ_b^0 is a full rank matrix that confirms the observability of the system (not shown for brevity).

ii) The length of estimation (*N*), prediction (*L*) and control (*C*) horizon are set to 10. These tuning parameters were obtained from preliminary simulations. The time interval is set to 1 s.

iii) The HIPS process requires to produce a variety of grades of polystyrene based on consumer demands[182]. Hence, three step changes in the set-point of the monomer concentration C_m were considered to develop the closed-loop framework for three different grades of the product where the conversion of monomer is 20%, 25%, and 30% [182,183].

The weight matrix for CVs is as follows: $\boldsymbol{Q_{out}} = \boldsymbol{I}_{2\times2}$, i.e., the weights are the same for both CVs (i.e., C_m and T). No weights on the MVs are considered in the NMPC framework.

Appendix E
Jacobian and Sensitivity Matrices

The analytical evaluation of the Jacobian matrix for each of the case studies considered in this thesis has been performed using the symbolic toolbox in SymPy library of Python[214]. Note that the analytical evaluation of the Jacobian matrix allows to update this matrix based on the current numerical value of the state variables of the system; this provides the possibility to check the linear observability of the system operating around a nominal operating point. As a reminder, the linear observability matrix for each process can be evaluated based on the corresponding Jacobian matrix $\boldsymbol{A}$ and the sensitivity matrix $\boldsymbol{H}$ of the system, as shown in Equation (3-9). The matrices for most of the case studies presented in this thesis are presented next.

Gas-phase Reactor

The Jacobian and sensitivity matrices for the gas-phase reactor are as follows:

$$\boldsymbol{A} = \begin{bmatrix} [-64 * p_A, & 32 * p_A], \\ [0, & 0] \end{bmatrix}$$

$$\boldsymbol{H} = [1 \quad 1]$$

where the first and second row in $\boldsymbol{A}$ represent the derivatives of the first and second equations with respect to the states, respectively. Note that the first and second column represent the derivatives with respect to p_A and p_B, respectively.

Wastewater Treatment plant (WTP)

The Jacobian matrix for WTP is a 6x6 matrix. The first to sixth element at each row represents the partial derivatives of the one of the differential equations in WTP (see Equation (4-29)) with respect to x_w, s_w, x_d, x_b, x_r, and c_w, respectively. The analytical Jacobian matrix determined for WTP based on its state variables is as follows:

$$\boldsymbol{A} = [[0.10849152 * s_w/(s_w + 300.0) - 0.12550833 - 0.0001 * x_w/s_w, 32.547456 * x_w/(s_w + 300.0)^2 + 5.0e-5 * {x_w}^2/{s_w}^2, 0, 0, 0.0503750000000000, 0],$$

$$\left[-0.1824 * \frac{s_w}{s_w + 300.0} + 2.6666e-5 + 2.0e-5 * \frac{x_w}{sw}, -54.72 * \frac{x_w}{(s_w + 300.0)^2} - 0.075 - 1.0e-5 * \frac{{x_w}^2}{{s_w}^2}, 0, 0, 0, 0\right],$$

$$[0, 0, 0.0012399051105 * x_d * \exp(-0.00078567 * x_d) - 0.0725 - 1.57815 * \exp(-0.00078567 * x_d), 0.0725000000000000, 0, 0],$$
$$[0.250750000000000, 0, -0.002479810221 * x_d * \exp(-0.00078567 * x_d) + 3.1563 * \exp(-0.00078567 * x_d), 0.002479810221 * x_b * \exp(-0.00078567 * x_b) - 0.25075 - 3.1563 * \exp(-0.00078567 * x_b), 0, 0],$$
$$[0, 0, 0, -0.002479810221 * x_b * \exp(-0.00078567 * x_b) + 0.10575 + 3.1563 * \exp(-0.00078567 * x_b), -0.105750000000000, 0],$$
$$[-1.824e-5 * s_w/(s_w + 300.0), -0.005472 * x_w/(s_w + 300.0) * 2, 0, 0, 0, -0.265375000000000]]$$

The sensitivity matrix $\boldsymbol{H}$ for this system is:

$$\boldsymbol{H} = \begin{bmatrix} [0 & 1 & 0 & 0 & 0 & 0\,], \\ [0 & 0 & 1 & 0 & 0 & 0\,], \\ [0 & 0 & 0 & 0 & 0 & 1\,] \end{bmatrix}$$

Network of continuous-stirred tank reactors (CSTRs) and separator

The Jacobian matrix for this process is a 9x9 matrix. The first to ninth element at each row is the first derivative of the one of the differential equations (see Equation (8-9)) with respect to T_1, C_{A_1}, C_{B_1}, T_2, C_{A_2}, C_{B_2}, T_3, C_{A_3}, and C_{B_3}, respectively. With this in mind, the Jacobian matrix and the sensitivity matrix for this system are as follows:

$$\boldsymbol{A} = [[547878871188951.0 * {C_{A_1}}^2 * \exp(-8539.81236468607/T_1)/T_1^2 - 1.4, 128311688311.688 * C_{A_1} * \exp(-8539.81236468607/T_1), 0, 0, 0, 0, 0.400000000000000, 0, 0],$$
$$\left[-16225643492903.5 * {C_{A_1}}^2 * \exp(-8539.81236468607/T_1)/T_1^2, -3800000000.0 * C_{A_1} * \exp(-8539.81236468607/T_1) - 1.4, 0, 0, 0, 0, 0, -43200.0 * C_{A_3}/\left(36.0 * C_{A_3} - 3.6 * CB3 + 1000.0\right)^2 + 1200.0/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right), 4320.0 * C_{A_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2\right],$$

$$\Big[16225643492903.5 * {C_{A_1}}^2 * \exp(-8539.81236468607/T_1)/T_1^2, 3800000000.0 * C_{A_1}$$
$$* \exp(-8539.81236468607/T_1), -1.40000000000000, 0, 0, 0, 0, -11520.0$$
$$* C_{B_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right), 1152.0$$
$$* C_{B_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2$$
$$+ 320.0/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)\Big],$$

$$\big[1.40000000000000, 0, 0, 547878871188951.0 * {C_{A_2}}^2 * \exp(-8539.81236468607/T_2)/T_2^2$$
$$- 2.4, 128311688311.688 * C_{A_2} * \exp(-8539.81236468607/T_2), 0, 0, 0, 0\big],$$

$$\big[0, 1.40000000000000, 0, -16225643492903.5 * {C_{A_2}}^2$$
$$* (\exp(-8539.81236468607/T_2))/T_2^2, -3800000000.0 * C_{A_2}$$
$$* \exp(-8539.81236468607/T_2) - 2.4, 0, 0, 0, 0\big],$$

$$\big[0, 0, 1.40000000000000, 16225643492903.5 * {C_{A_2}}^2$$
$$* \exp(-8539.81236468607/T_2)/T_2^2, 3800000000.0 * C_{A_2}$$
$$* \exp(-8539.81236468607/T_2), -2.40000000000000, 0, 0, 0\big],$$

$$\Big[0, 0, 0, 4.00000000000000, 0, 0, -4.00000000000000, 12529870.1298701$$
$$* C_{A_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2 + 3341298.7012987$$
$$* C_{B_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2 + 4176.62337662338$$
$$* \left(-450.0 * C_{A_3} - 450.0 * C_{B_3} + 25000.0\right)/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2$$
$$- 295844.155844156/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right), -1252987.01298701$$
$$* C_{A_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2 - 334129.87012987$$
$$* C_{B_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2 - 417.662337662338$$
$$* \left(-450.0 * C_{A_3} - 450.0 * C_{B_3} + 25000.0\right)/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2$$
$$- 40606.0606060606/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)\Big],$$

$$\Big[0, 0, 0, 0, 4.00000000000000, 0, 0, 72000.0 * C_{A_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2$$
$$- 3.33333333333333 - 2000.0/\left(36.0 * C_{A_3} - 3.6 * CB3 + 1000.0\right), -7200.0$$
$$* C_{A_3}/\left(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0\right)^2\Big],$$

$$[0, 0, 0, 0, 0, 4.00000000000000, 0, 19200.0 * CB3/(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0)^2, -1920.0 * C_{B_3}/(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0)^2 - 3.33333333333333 - 533.333333333333/(36.0 * C_{A_3} - 3.6 * C_{B_3} + 1000.0)]]$$

And the $\boldsymbol{H}$ is:

$$\boldsymbol{H} = \begin{bmatrix} [1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\ ], \\ [0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\ ], \\ [0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\ ] \end{bmatrix}$$