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

Mahshad Valipour, Luis A. Ricardez-Sandoval*

Department of Chemical Engineering, University of Waterloo, Waterloo, Canada N2L 3G1

Supplementary material

Sections

A. WTP: Model parameters and assumptions

B. WTP: *Scenarios I-III*: Open-loop estimation

C. HIPS: Model parameters and assumptions

D. AC estimation methods

* Corresponding author information: Phone: 1-519-888-4567 ext. 38667. Fax: 1-519-888-4347. E-mail: laricard@uwaterloo.ca

A. WTP: Model parameters and assumptions

This section presents relevant information for the WTP as well as the assumptions considered in this case study. Table A.1 presents the nominal steady-state value of the states, the capacity of bioreactor $\boldsymbol{V_r}$ and the area of the decanter $\boldsymbol{A_d}$. Note that the lower bounds for $\boldsymbol{q_p}$ and $\boldsymbol{q_2}$ are set to zero, whereas the upper bound for these two MVs is set tot 600.

Table A.1. Nominal values of states, design parameters and process model parameters

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)
V_r	2500.00 (m^3)
A_d	1100.00 (m^2)
q_2^l, q_2^u	0-600 (m^3/hr)
q_p^l, q_p^u	0-600 (m^3/hr)
f_k^l, f_k^u	0.01-2.5

The 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 A.1.

iii) There are two standard choices of importance function in C-PF (or in overall PF), namely the posterior and the prior[50]. 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 distribution 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 ($\hat{\boldsymbol{x}}_{k-N-1}$) 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 A.1).

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 A.1. 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.

B. 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 A.1 in section A. Table B.1 summarizes the results obtained from *Scenarios* $\mathcal{I}$ and *Scenario* $\mathcal{II}$ during open-loop operation. As shown in Table B.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* $\mathcal{I}$, 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* $\mathcal{II}$ 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 B.1. $\widehat{SSE}^{(n)}$ Open-loop operation of the WTP, *Scenarios* $\mathcal{I}$ - $\mathcal{II}$

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 B.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 B.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 B.3 compares the MHE performance obtained from the different instances considered in *Scenario* III. As presented in Table 1 in section 4.1, 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 B.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 B.1. Note however that the normalized SSE ($\widehat{SSE}^{(n)}$) reported in Table B.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 B.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 B.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

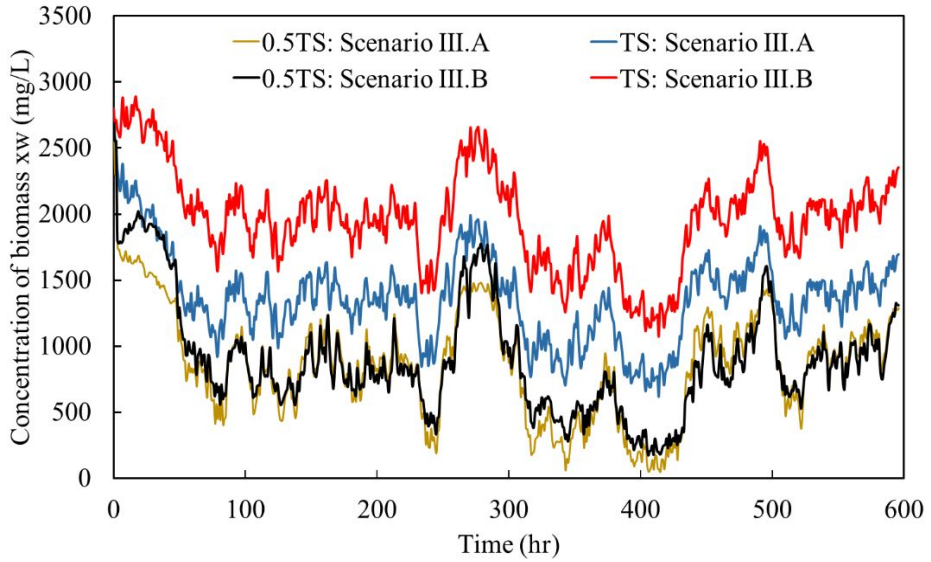

Figure B.1. Open-loop MHE estimation, *Scenario III.A and III.B*

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.

C. HIPS: Model parameters and assumptions

This section presents the information and assumptions that have been considered in the HIPS case study. Table C.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 C.1.

Table C.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_{psw}	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[31]. 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% [31,32].

iv) The weight matrix for CVs is as follows: $\boldsymbol{Q_{out}} = \boldsymbol{I}_{2\times 2}$, 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.

D. AC estimation methods

This section provides a brief overview of the EKF and PF methods.

Extended Kalman Filter (EKF): There are two main steps involved in the EKF algorithm, i.e., the predictor and corrector steps shown in Equations (26) and (27). This EKF algorithm interacts with the MHE as shown in Figure 1, i.e., past estimation results of MHE problem $\hat{\boldsymbol{x}}_{k-N-1}$ are used for initialization of AC estimator (EKF). Moreover, the EKF uses the measurement outputs of the process plant for the corrector update step. As a result, $\hat{\boldsymbol{x}}^{(\boldsymbol{EKF})}_{k-N|k-N}$ and $\boldsymbol{P}^{\boldsymbol{EKF}}{}_{k-N|k-N}$ are the mean value and covariance of the posterior distribution of the states at time *k*-*N*. The estimated terms $\hat{\boldsymbol{x}}^{(\boldsymbol{EKF})}_{k-N|k-N}$ and $\boldsymbol{P}_{\boldsymbol{EKF}k-N|K-N}$ represent the approximation of the AC parameters, i.e. $\overline{\boldsymbol{x}}_{k-N}$ and $\boldsymbol{P}_{k-N}$, are they are used as inputs to the MHE estimation at time step *k*.

Particle Filtering method (PF): In this method, a set of particles is used to estimate the conditional posterior density function for states at time *k*-*N* thus providing an approximation of AC. These particles are selected by taking the advantage of Sequential Monte Carlo sampling. In fact, the PF method assumes that the weighted sum of a sufficiently large N_P samples approximates the true posterior distribution function as follows[36,37]:

$$p\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\boldsymbol{Y}_{k-N}\right)\approx\frac{1}{N_P}\sum_{i=1}^{N_P}\delta\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}-\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}_{k-N}\right)\equiv\hat{p}\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)\qquad\text{(D1)}$$

where $\delta(.)$ is the Dirac delta function, $\mathbf{Y}_{k-N}:=\{\boldsymbol{y}_0,\boldsymbol{y}_1,...,\boldsymbol{y}_{k-N}\}$ that includes all the past measurements from the initial time instance $\boldsymbol{y}_0$ to time *k*-*N* ($\boldsymbol{y}_{k-N}$), and the samples $\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}_{k-N}$ are independently distributed particles taken from $p\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)$. Based on the approximation in equation (D1) the expected value of a given nonlinear function $f(.)$ is[36,37]:

$$\mathbb{E}\left[f\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\right)\right] \approx \int f\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\right)\hat{p}\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)d\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N} = \frac{1}{N_P}\sum_{i=1}^{N_P} f(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}{}_{k-N}) \quad \text{(D2)}$$

Evaluation of the expected value based on Equation (D2) needs sampling from a posterior distribution; however, usually the shape or type of the distribution is unknown. As an alternative, a completely defined trial or importance distribution is considered to obtain the samples. At each sampling time, particles are taken from the importance density $\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)$. Therefore, the expectation approximation can be determined by modifying the previous equation as follows[36,37]:

$$\begin{aligned}\mathbb{E}\left[f\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\right)\right] &= \int f\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\right)\frac{p\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)}{\pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)}\pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)d\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N} \\ &\approx \sum_{i=1}^{N_P} f(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}{}_{k-N})w^i_{k-N}(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}{}_{k-N})\end{aligned} \quad \text{(D3)}$$

$$\boldsymbol{wgt}^i_{k-N}(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}{}_{k-N}) = \frac{p\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)}{\pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^i}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right)} \quad \text{(D4)}$$

where $\boldsymbol{wgt}^i_{k-N}$ is the weight function for each of these samples. Based on the above, the recursive form of the chosen importance distribution can be described as follows[36,37]:

$$\begin{aligned}&\pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\mathbf{Y}_{k-N}\right) = \pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N}\middle|\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N-1},\boldsymbol{y}_{k-N}\right)\pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_{k-N-1}\middle|\mathbf{Y}_{k-N-1}\right) \\ &\pi(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_0) \sim p(\hat{\boldsymbol{x}}^{\boldsymbol{PF}}{}_0)\end{aligned} \quad \text{(D5)}$$

As the result, the weighted function can be reformulated as follows[36,37]:

$$\boldsymbol{wgt}^{i}_{k-N}\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N}\right) = \boldsymbol{wgt}^{i}_{k-N-1}(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N-1})\frac{p(\boldsymbol{y}_{k-N}|\,\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N})p(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N}|\,\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N-}}{\pi\left(\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N}\middle|\,\hat{\boldsymbol{X}}_{k-N-1},\boldsymbol{y}_{k-N}\right)} \quad \text{(D6)}$$

Note that $\hat{\boldsymbol{X}}_{k-N-1}$is a matrix of all the estimations from time instant zero to time *k*-*N*-1, $\hat{\boldsymbol{x}}^{\boldsymbol{PF}^{i}}{}_{k-N-1}$ denotes the samples around the past estimation results of MHE problem, i.e., $\hat{\boldsymbol{x}}_{k-N-1}$. Based on the equations (D5), (D6), (2), the AC expression for PF can be approximated as follows:

$$\varphi_{k-N} \approx \sum_{i=1}^{N_P} \boldsymbol{wp}^{i}_{k-N}\left\|\boldsymbol{x}_{k-N} - \overline{\boldsymbol{x}^{\boldsymbol{PF}}}^{i}{}_{k-N}\right\|^{2}_{[\overline{\boldsymbol{P}}^{i}_{k-N|k-N}]^{-1}} \quad \text{(D7)}$$

here $\overline{\boldsymbol{x}^{\boldsymbol{PF}}}^{i}{}_{k-N}$ and $\overline{\boldsymbol{P}}^{i}_{k-N|k-N}$ are the mean and the covariance associated with the posterior distribution of i^{th} particle at time *k*-*N*. The chosen importance distribution is a key aspect of the PF approach since it can significantly reduce the number of required particles to approximate the posterior distribution. Particularly, in the presence of bounds on the state variables it is essential to choose a consistent importance distribution with respect to these bounds. In fact, the key factor is that the particles should be drawn from the bounded distributions, which makes Constrained-PF (C-PF) an adequate constrained estimator for approximating AC parameters for constrained nonlinear systems.