Anton A. Kiss, Edwin Zondervan, Richard Lakerveld, Leyla Özkan (Eds.)
Proceedings of the 29th European Symposium on Computer Aided Process Engineering
June 16th to 19th, 2019, Eindhoven, The Netherlands.

A point estimate method-based back-off approach to robust optimization: application to pharmaceutical processes

Victor N. Emenike[1,a,b], Xiangzhong Xie[1,a,b], Ulrike Krewer[a,b] and René Schenkendorf[a,b,*]

[a]*Institute of Energy and Process Systems Engineering, TU Braunschweig, Franz-Liszt-Straße 35, 38106 Braunschweig, Germany*
[b]*Center of Pharmaceutical Engineering, Franz-Liszt-Straße 35A, 38106 Braunschweig, Germany*
r.schenkendorf@tu-braunschweig.de

Abstract

In this contribution, we propose estimating the means and variances required for calculating back-off terms by using the point estimate method (PEM) as a highly efficient sampling strategy in robust process design. As case studies, we consider an upstream pharmaceutical process which involves the synthesis of 2-hydroxy-ketones via enzyme-catalyzed carboligation and a downstream pharmaceutical process that includes the continuous crystallization of ibuprofen. We show that the proposed PEM-based back-off approach is significantly faster than conventional Monte Carlo brute-force sampling methods while maintaining robust solutions with low approximation errors. In general, the efficient PEM-sampling strategy guarantees the analysis and the robust design of complex (bio)pharmaceutical process chains.

Keywords: pharmaceutical manufacturing, robust optimization, back-off approach, enzyme catalysis, ibuprofen crystallization

1. Introduction

The pharmaceutical industry has a substantial impact on the social and economic welfare of the individual and society. For the industry to continue producing high-quality and effective drugs even in the face of economic constraints, rising population and diseases, regulatory bodies and industry leaders alike have stipulated Quality by Design (QbD) as an essential paradigm. At the heart of QbD are mathematical models which are crucial for analyzing, optimizing, monitoring and controlling pharmaceutical processes (Emenike et al., 2018a,b). These models need to be properly calibrated to ensure that they reflect the physical processes they represent (Schenkendorf et al., 2018). In calibrating these models, a crucial issue that has to be dealt with is the presence of model and parameter uncertainties. A possible way to robustify processes under uncertainty is by using the back-off approach. This approach involves tightening violated constraints and shrinking the feasible region by introducing margins called back-offs. By so doing, the worst-case realization of a given process will still be feasible despite variations in the constraints (Shi et al., 2016). Moreover, these back-offs are usually calculated offline, and thus, do not lead to additional complexity of the optimization problems. Typically, Monte Carlo simulations are used to estimate the statistical moments (i.e., means and variances) required for calculating back-offs (Shi et al.,

[1]These authors contributed equally.

2016). However, for these means and variances to be accurately estimated, numerous Monte Carlo simulations are usually required. Thus, this could lead to high computational costs especially when a single Monte Carlo simulation of the process is computationally expensive.

Alternatively, these statistical moments can be approximated efficiently by using the point estimate method (PEM). It has been shown that the PEM is a computationally efficient and relatively accurate sampling strategy for estimating statistical moments (Schenkendorf, 2014; Xie et al., 2018b). Recently, we proposed a systematic robust optimization framework that combines the elementary process functions methodology, global sensitivity analysis, and the back-off approach (Emenike et al., 2019). A key contribution in Emenike et al. (2019) was a new back-off algorithm that uses the PEM instead of Monte Carlo simulations. We showed that the proposed PEM-based back-off approach is at least 10 times faster than the conventional Monte Carlo-based back-off approach while maintaining the quality of robust solutions. Maußner and Freund (2018) used cubature rules in lieu of Monte Carlo simulations and came to similar conclusions.

In this contribution, we build upon our original work (Emenike et al., 2019) by applying the novel algorithm in the presence of correlated parameter uncertainties and show that the robustification algorithm is not limited to upstream processes but is a versatile tool for whole pharmaceutical process chains. To this end, we apply the novel algorithm to an upstream pharmaceutical process that involves the synthesis of 2-hydroxy-ketones via enzyme-catalyzed carboligation and a downstream pharmaceutical process that includes the crystallization of ibuprofen. Details of the algorithm and the results for the case studies are presented in sections 2 and 3, respectively.

2. Methodology

A major advantage of the back-off approach to dynamic optimization under uncertainty is that its formulation is of similar complexity as the nominal dynamic optimization problem. The robust optimization problem (Problem 1) with time-varying back-offs $\mathbf{b}(t)$ is given as:

$$\begin{aligned}
\underset{\mathbf{x}(\cdot),\mathbf{u}(\cdot),\mathbf{z}(\cdot)}{\text{minimize}} \quad & \Phi(\mathbf{x}(t_\mathrm{f}))) && \text{(1a)}\\
\text{subject to} \quad & \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t),\mathbf{z}(t),\mathbf{u}(t),\bar{\mathbf{p}}), \quad \forall t \in \mathscr{T}, && \text{(1b)}\\
& \mathbf{g}(\mathbf{x}(t),\mathbf{z}(t),\mathbf{u}(t),\bar{\mathbf{p}}) = \mathbf{0}, \quad \forall t \in \mathscr{T}, && \text{(1c)}\\
& \mathbf{h}(\mathbf{x}(t),\mathbf{z}(t),\mathbf{u}(t),\bar{\mathbf{p}}) + \mathbf{b}(t) \leq \mathbf{0}, \quad \forall t \in \mathscr{T}, && \text{(1d)}\\
& \mathbf{x}(t_0) = \mathbf{x}_0, && \text{(1e)}\\
& \mathbf{u}(t) \in \mathscr{U}, && \text{(1f)}
\end{aligned}$$

on the time horizon $\mathscr{T} := [t_0, t_\mathrm{f}] \subset \mathbb{R}$, where $\mathbf{x}$, $\mathbf{u}$, $\mathbf{z}$ represent states, controls, and algebraic variables, respectively. $\mathbf{g}$ and $\mathbf{h}$ represent the equality and inequality constraints, respectively. As we can see from Problem 1, the dynamic optimization with back-offs is optimized at the nominal parameter vector $\bar{\mathbf{p}}$. Here, the time-varying back-offs $\mathbf{b}(t)$ in Eq. 1d are included as margins to shrink the feasible region of the dynamic optimization problem and thus, making the optimal operating conditions robust.

As pointed out by Shi et al. (2016), the solution from Problem 1 is not guaranteed to be optimum, unless the back-offs are insensitive to the decision variables. Therefore, an iterative approach is proposed to update the back-offs with the optimal design from the last iteration. The solution from Problem 1 and back-offs are consistently improved and can be exported once the back-offs converge. We depict the details regarding the iterative algorithm in Fig. 1. Moreover, we use the PEM instead of Monte Carlo simulations used by Shi et al. (2016) to derive the statistical moments required to calculate the back-offs. The PEM utilizes a relatively small number of deterministic

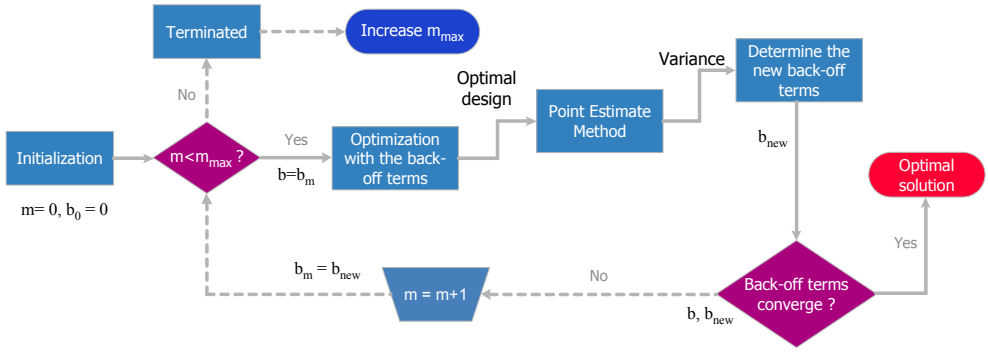


Figure 1: Computational scheme for the point estimate method-based back-off approach for robust optimization, where m is the iteration index.

samples to compute the statistical moments of system states and thus, facilitates an efficient back-off algorithm. For more details regarding the PEM, please refer to Xie et al. (2018b) and Emenike et al. (2019).

3. Case studies

3.1. Enzyme-catalyzed carboligation

First, we consider an upstream pharmaceutical process that involves an enzyme-catalyzed carboligation between propanal (A) and benzaldehyde (B) to form 2-hydroxy-ketones (BA) and benzoin (BB). Here, we aim to maximize the formation of the target product BA under the correlated parameter uncertainties specified in Xie et al. (2018a). The mechanistic model for this reaction is given as:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{z}(t), \mathbf{u}(t), \mathbf{p}) = \begin{bmatrix} \frac{u_{\mathrm{A}} \cdot C_{\mathrm{A}}^{\mathrm{in}}}{V} - \frac{C_{\mathrm{A}}}{V}(u_{\mathrm{A}} + u_{\mathrm{B}}) + r_{\mathrm{A}} \\ \frac{u_{\mathrm{B}} \cdot C_{\mathrm{B}}^{\mathrm{in}}}{V} - \frac{C_{\mathrm{B}}}{V}(u_{\mathrm{A}} + u_{\mathrm{B}}) + r_{\mathrm{B}} \\ -\frac{C_{\mathrm{BA}}}{V}(u_{\mathrm{A}} + u_{\mathrm{B}}) + r_{\mathrm{BA}} \\ -\frac{C_{\mathrm{BB}}}{V}(u_{\mathrm{A}} + u_{\mathrm{B}}) + r_{\mathrm{BB}} \\ -\frac{C_{\mathrm{E}}}{V}(u_{\mathrm{A}} + u_{\mathrm{B}}) + r_{\mathrm{E}} \\ u_{\mathrm{A}} + u_{\mathrm{B}} \end{bmatrix}, \tag{2}$$

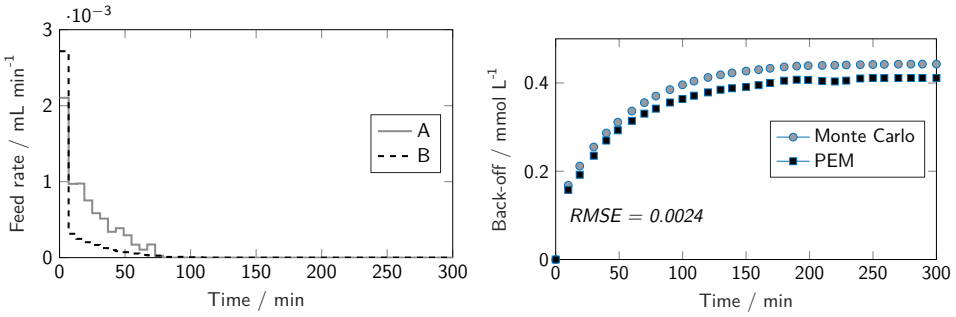
where C_i is the concentration of species i; $r_{\mathrm{A}}, r_{\mathrm{B}}, r_{\mathrm{BA}}, r_{\mathrm{BB}}$, and r_{E} are the reaction rates for A, B, BA, BB, and E, respectively; u_{A} and u_{B} are the controlled feed rates; and $C_{\mathrm{A}}^{\mathrm{in}}$ and $C_{\mathrm{B}}^{\mathrm{in}}$ are the inlet feed concentrations of A and B, respectively. For details on the model and model parameters, we refer to Emenike et al. (2019). First, forward simulations by using 10,000 Monte Carlo simulations were performed on the nominal problem to determine which constraints were violated, and it was found that only the inequality constraint bounding C_{BB} was violated. Therefore, we focus on robustifying only the C_{BB} inequality constraint as shown in Eq. (3):

$$0 \leq C_{\mathrm{BB}}(t) \leq 2.78 \text{ mM} - b(t), \quad \forall t \in \mathscr{T}, \tag{3}$$

where 2.78 mM is the solubility limit of BB, $\mathscr{T} := [t_0, t_{\mathrm{f}}] \subset \mathbb{R}$, and final time, $t_{\mathrm{f}} = 300$ min. By applying the robust optimization strategy presented in Section 2 and aiming to satisfy Eq. 3 at a probability of 99.90%, we see in Table 1 that the PEM-based back-off algorithm is able to achieve this after just one iteration of the algorithm.

Table 1: Comparison of the point estimate method-based algorithm with the Monte Carlo-based back-off algorithm for robust dynamic optimization in comparison to to the nominal case.

Scenarios	$C_{\mathrm{BA}}(t_{\mathrm{f}})$ mM	Violation probability [%]	CPU time [s]
Nominal	3.60	57.65	4
PEM-based back-off	3.48	0.16	114
Monte Carlo-based back-off	3.49	0.13	2626



(a) (b)

Figure 2: Results for the upstream pharmaceutical process: enzyme-catalyzed carboligation. (a) robust controls. (b) comparison between PEM and Monte Carlo simulations.

Moreover, this was achieved at a computation time that is 23 times faster than the conventional Monte Carlo-based back-off algorithm which was used as a benchmark for our algorithm (see Table 1). This speed-up is mainly due to the lower number of PEM sample points ($2\times 13^2+1=339$) in comparison to the 10,000 Monte Carlo sample points. It is also possible to further reduce the PEM points by using a global sensitivity analysis as shown in Emenike et al. (2019). It can be seen in Fig. 2b that the time-varying back-offs calculated by both approaches are close with a marginal root-mean-square prediction error (RMSE) of 0.0024. This low RMSE validates the accuracy of our PEM-based approach. Furthermore, we note that the probability of violation and the maximum $C_{\mathrm{BA}}(t_{\mathrm{f}})$ obtained are very close for both approaches, thus, suggesting that our PEM-based back-off approach is very accurate for the case study considered. The robust controls (see Fig 2a) lead to a maximum concentration of 3.48 mmol$\,\mathrm{L}^{-1}$ which is 3.33% lower than the nominal value. This marginal decrease shows that the novel approach is not adversely conservative while ensuring robustness. Therefore, these results demonstrate that the PEM-based back-off strategy is very efficient and useful for the enzyme-catalyzed carboligation considered in this work.

3.2. Crystallization of ibuprofen

Second, we consider the continuous crystallization of ibuprofen in a plug-flow crystallizer (PFC) as a representative downstream pharmaceutical process. The crystal size distribution (CSD) n was chosen as an important key performance indicator for QbD. A population balance equation in combination with mass balance equations in liquid and solid phase was used to predict the evolution of the CSD along the PFC. To reduce the computational complexity, we discretized the PFC model to ordinary differential equations by using the classical method of moments. The resulting moment-based model for PFC is given in Eqs. (4) to (6):

$$\frac{d\mu_0}{dz}=\frac{B}{v} \tag{4}$$

$$\frac{d\mu_l}{dz}=\frac{BL_0^l}{v}+\frac{kG\mu_{l-1}}{v}\quad l=1,\ldots,5 \tag{5}$$

Table 2: The mean value of mass-based mean crystal size d_{43} and the probability of a constraint violation (supersaturation) from the nominal design and the robust design with the PEM-based back-offs.

Scenarios	$\mathbb{E}(d_{43})$	Violation probability [%]
Nominal	109.3	48
PEM-based back-off	108.8	3

$$\text{Liquid:} \quad \frac{\partial C}{\partial z} = -\frac{k_v \rho_s}{v}(BL_0^3 + 3G\mu_2), \tag{6}$$

where z is the axis coordinate of the PFC, m; L_0 is the nuclei size, m; v is the superficial velocity of slurry along the PFC, $\mathrm{m\,s^{-1}}$; k_v and ρ_s are the shape factor and the crystal density, $\mathrm{kg\,m^{-3}}$, and C is the mass of solute per kg slurry, $\mathrm{kg\,kg^{-1}}$. μ_k is the kth moment which is used to describe the major information in CSD and defined with Eq. (7). B and G are the nucleation rate, $\#\,\mathrm{kg^{-1}\,s^{-1}}$, and the growth rate, $\mathrm{m\,s^{-1}}$, which describe the kinetics of the crystallization of ibuprofen and are determined by the degree of supersaturation S, as shown in Eqs. (8) and (9):

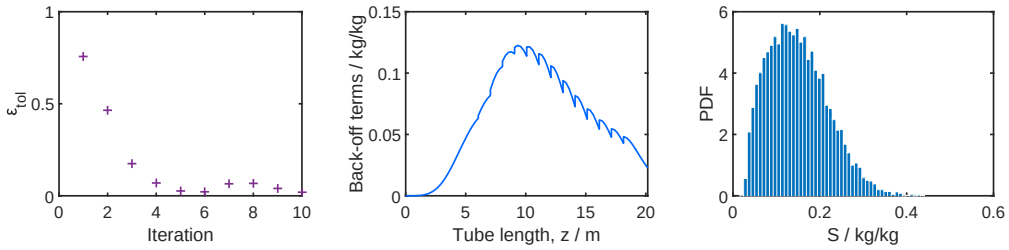
$$\mu_l = \int_0^\infty L^l n dL, \quad l = 0, \ldots, 5 \tag{7}$$

$$B = k_{b_0} S(T) \tag{8}$$

$$G = k_{g_0} \exp\left(\frac{T}{T_g}\right) S(T) \tag{9}$$

This case study aims to maximize the mass-based mean crystal size d_{43} (i.e., the ratio between μ_4 and μ_3) by manipulating the temperature along the PFC. There is an inequality constraint in the design which restricts the supersaturation of ibuprofen in the solution below the primary nucleation threshold to ensure no primary nucleation occurs. Several other inequality constraints on the yield of product and temperature gradient are also satisfied but not discussed in what follows, as they are not violated even in the presence of parameter uncertainties. Rashid (2011) estimated the values of the kinetic parameters k_{b_0}, k_{g_0}, and T_g with designed experiments and showed the estimated values are not accurate and associated with uncertainty. The parameter uncertainties are then described by Gaussian distributions and thus, are taken into account in the robust design of the PFC in this work.

For the nominal design, the parameter uncertainties are neglected. The obtained d_{43} is 109.3 μm (see Table 2). However, the inequality constraint on supersaturation is violated with a probability of 48% due to parameter uncertainties. The violation probability of inequality constraint is determined with 10,000 evaluations of the PFC model with the random samples generated from the probability distributions of kinetic parameters. The PEM-based back-off strategy introduced in Section 2 is then implemented to design a robust PFC tolerant to the parameter uncertainties, where the target violation probability of the inequality constraint is set to smaller than 1%. Results from the iterative back-off approach are depicted in Fig. 3. As we can see in Fig. 3a, the tolerance factor ε_{tol}, which represents the difference between the back-offs in the adjacent iterations, converges within 10 iterations. The resulting time-varying back-offs are plotted in Fig. 3b, in which the supersaturation is shrunk more in the middle and less on both sides of the PFC to mitigate the effect of parameter uncertainties. According to the results listed in Table 2, the d_{43} from the robust design decreases slightly when compared to the value from nominal design. The probability of constraint violation decreases to 3%. Thus, the process robustness is increased at the cost of a deteriorated performance. Although the back-off approach increases the robustness of the process significantly, the obtained violation probability 3% is still three times larger than the target value of 1%. The reason for this is that the back-offs calculated with the variances of system states are actually not accurate when their probability distributions are asymmetric, as shown in Fig. 3c.



(a) (b) (c)

Figure 3: Results for the downstream pharmaceutical process: continuous crystallization of ibuprofen. (a) Convergence plot of the back-off terms. (b) value of the time-varying back-off terms at the last iteration. (c) probability distribution of supersaturation S at location $z = 11$m of the PFC.

4. Conclusions

In this work, we proposed a PEM-based back-off approach for the robust design of upstream and downstream pharmaceutical processes. First, the approach was implemented for the design of a fed-batch reactor for enzyme-catalyzed carboligation in the presence of parameter correlations and uncertainty. The results showed that the proposed PEM-based back-off approach is significantly faster than the conventional Monte Carlo-based back-off approach while achieving high accuracy of the robust solutions. The method was also applied to the design of continuous crystallization of ibuprofen in the presence of uncertainties in the kinetic parameters. This approach also lowered the value of constraint violation and significantly improved the robustness of the process. However, the probability of a constraint violation is still three times higher than the given target value due to the asymmetric probability distribution of system states. Future work will include non-Gaussian probability distributions for robust process design.

Acknowledgements

X. Xie acknowledges funding from the "Promotionsprogramm μ-Props" by MWK Niedersachsen. V. Emenike and X. Xie acknowledge the support from the International Max Planck Research School for Advanced Methods in Process and Systems Engineering, MPI Magdeburg.

References

V. N. Emenike, R. Schenkendorf, U. Krewer, 2018a. A systematic reactor design approach for the synthesis of active pharmaceutical ingredients. European Journal of Pharmaceutics and Biopharmaceutics 126, 75–88.

V. N. Emenike, R. Schenkendorf, U. Krewer, 2018b. Model-based optimization of biopharmaceutical manufacturing in Pichia pastoris based on dynamic flux balance analysis. Computers & Chemical Engineering 118, 1–13.

V. N. Emenike, X. Xie, R. Schenkendorf, A. C. Spiess, U. Krewer, 2019. Robust dynamic optimization of enzyme-catalyzed carboligation: A point estimate-based back-off approach. Computers & Chemical Engineering 121, 232–247.

J. Maußner, H. Freund, 2018. Optimization under uncertainty in chemical engineering: Comparative evaluation of unscented transformation methods and cubature rules. Chemical Engineering Science 183, 329–345.

M. A. Rashid, 2011. Crystallization engineering of ibuprofen for pharmaceutical formulation. PhD dissertation, Queensland University of Technology.

R. Schenkendorf, 2014. A general framework for uncertainty propagation based on point estimate methods. Second European Conference of the Prognostics and Health Management Society, PHME14, Nantes, France.

R. Schenkendorf, X. Xie, M. Rehbein, S. Scholl, U. Krewer, 2018. The impact of global sensitivities and design measures in model-based optimal experimental design. Processes 6 (4), 27.

J. Shi, L. T. Biegler, I. Hamdan, J. Wassick, 2016. Optimization of grade transitions in polyethylene solution polymerization process under uncertainty. Computers and Chemical Engineering 95, 260–279.

X. Xie, R. Ohs, A. Spiess, U. Krewer, R. Schenkendorf, 2018a. Moment-independent sensitivity analysis of enzyme-catalyzed reactions with correlated model parameters. IFAC-PapersOnLine 51 (2), 753 – 758.

X. Xie, R. Schenkendorf, U. Krewer, 2018b. Toward a comprehensive and efficient robust optimization framework for (bio)chemical processes. Processes 6 (10), 183.