Optimizing Control of a Tubular Polymerization Reactor: Comparison of Single Shooting and Full Discretization

R. Hashemi*, R. Schilling, S. Engell

*Dynamics and Operations Group, Technische Universität Dortmund, Germany, (e-mail: {Reza.Hashemi, Ricardo.Schilling, Sebastian.Engell}@bci.tu-dortmund.de).

Abstract: The goal of this contribution is to study numeric solution techniques for implementing optimizing control of polymerization processes in tubular reactors with multiple side injections of monomer along the reactor. The configuration of the reactor causes long delays between the inputs and the measurements at the reactor outlet and sharp moving fronts when the inflows are changed. Moreover, the polymerization kinetics are strongly nonlinear. The process is described by 1D partial differential equations along the reactor length. This makes the application of optimizing control based on a rigorous process model challenging. The so called weighted essentially non-oscillatory scheme (WENO) is used to discretize the spatial dimension of the plant model. This method avoids the need for a very large number of discretization points and still the model can be simulated sufficiently accurately. The resulting ode model contains 1600 states and comprises five manipulated variables. We implement the optimizing controller using two different approaches: At first single shooting with control vector parametrization is used which is simple to implement and has fewer decision variables. This is compared to full discretization scheme using orthogonal collocation on finite elements which results in a very large but very sparse and structured nonlinear programing problem. The simulation results show that both approaches have a similar performance and drive the system to a significantly more productive steady state.

Keywords: Optimizing control, Nonlinear model predictive control, Single shooting, Control vector parametrization, Full discretization, Orthogonal collocation on finite elements.

1. INTRODUCTION

With the emergence of powerful hardware and efficient numeric optimization techniques, model-based optimizing control can be used for more complex systems. In this method, nonlinear models are employed to optimize the performance of the controlled system over a moving horizon under constraints on the future inputs and the predicted system states. The application of this type of controller has been studied for several different types of systems e.g. (Idris et al. 2012). While early applications of optimizing control concerned slow processes described by pdes e.g. chromatographic columns (Toumi and Engell 2004), only recently applications to tubular reactors have been investigated e.g. (Hashemi et al. 2013) and (Lao et al. 2013). The process considered in this work, was developed in the European project F3-Factory (Buchholz 2009) and is a benchmark for the transfer of batch polymerization to continuous operation. In this work, we follow the idea of optimizing control and maximize the product throughput directly while respecting product quality constraints (Engell 2007). The long computation times of the optimal inputs is a major bottleneck in using this type of the controllers for complex systems. In this contribution we compare the performance of the optimizing control using two different numeric methods in a complex and nonlinear process which is described by eight pdes.

The rest of this paper is organized as follows: The second section describes the process model, its derivation and the numerical method employed to discretize the spatial domain of pde model of the reactor system. Nonlinear model predictive control is briefly discussed in the third section. Section four presents the formulation of the optimizing control using single shooting with control vector parametrization and simulation results follow. In section five the optimizing control is formulated using orthogonal collocation on finite elements and the simulation results are presented. Conclusions and an outlook on future work follow in the last section.

2. PROCESS MODEL AND SIMULATION

The process which is investigated in this work is the continuous production of poly acrylic acid (PAA) in a tubular reactor with multiple side injections of monomer. The reactor consists of eight tubular reactor modules connected in series with a total length of four meters. In order to ensure an efficient mixing of the reactants, the reactor is...
equipped with static mixers. The P&ID diagram of the reactor is shown in figure 1. The reactor is divided into four segments where each segment consists of two modules. The internal volume of the first two segments is 45 ml whereas the other two segments are larger and each has a volume of 130 ml. This configuration leads to a total residence time of about 2600 seconds at the nominal flow rate of 1 kg/hr. Four side injections of monomer along the reactor serve as manipulated variables to control the product quantity and product quality. The temperature of the jacket, which is assumed to be uniform along the reactor, is set via a thermostat and offers another manipulated variable. At the reactor outlet, a measurement of the viscosity is available which is used to compute the average molecular weight of the produced polymer ($M_w$). Furthermore, the residual monomer in the product is assumed to be measured using spectroscopy at the same position ($c_M$).

![Fig. 1: Flow sheet of the modular continuous polymerization plant. $M_w$: weight average molecular weight (derived from a viscosity measurement), $c_M$: residual monomer, ($u_1$, $u_2$, $u_3$ and $u_4$): side injections of monomer, $T_{rac}$: uniform jacket temperature.](image)

The energy and component balances are used to set up a rigorous model for the reactor while a perfect mixer is assumed. The free radical polymerization of acrylic acid is modeled by the terminal equations (pde) are shown in equations 1 to 8 (Hashemi et al. 2013). The system is subject to Danckwerts boundary conditions at the monomer injection points (Danckwerts 1953). The weight average molecular weight of the product results from the moments as:

$$M_w = M_{w0} \frac{\mu_2}{\mu_1} + \frac{\lambda_2}{\mu_1 + \lambda_1}.$$  

Due to the near plug flow characteristic between the inputs and the outlets, the reactor system reacts to the changes of the input flow rates showing sharp concentration fronts and long settling times.

The literature have proposed several methods to solve such pde models numerically. In the method of lines, the derivatives in all dimensions of the pde system are substituted by algebraic approximations except of one. In other words, in this approach, the pde system is discretized on a grid/mesh and at every discretization point one or more odes are derived from the original pde. Thus the pde model is converted to a system of odes. Well-established methods for solving odes can then be applied to find an approximate solution of the pde system. Usually the spatial derivatives are substituted by algebraic approximations and the temporal derivatives are kept. The intuitive and standard choice to approximate the spatial derivatives is to use finite differences. However, as we have shown it in (Hashemi et al. 2014), this method is not capable to simulate stiff systems precisely unless a very fine discretization grid is used. As an alternative, nonlinear methods can be used to approximate the spatial derivatives. The class of such nonlinear methods is called high resolution methods. The weighted essentially non-oscillatory (WENO) scheme is a type of high resolution methods which uses the numerical flux of a variable to compute the derivatives while the numerical flux is computed on a dynamic set of stencils (Bouaswaig et al. 2009). A variant of this method, called WENO-Z (Borges et al. 2008) is used in this work to compute the first order spatial derivatives that appear in the pde model of the reactor system. The second order derivatives are approximated using the WENO scheme proposed by (Liu et al. 2011) which computes the second derivatives directly. The details about the computation of the numeric flux and determination of the smoothness indicators and weights can be found in the cited references. For the reactor system under consideration here, it was shown in (Hashemi et al. 2014) that using the above mentioned WENO schemes on a discretization grid with 200 points, one can reach a similar accuracy as by using finite differences on a grid with 5000 points. Therefore in this work, the spatial domain of the pde model of the reactor system is

\[
\begin{align*}
\frac{\partial u_1}{\partial t} & = -u \frac{\partial c_1}{\partial z} + D_{c1} \frac{\partial^2 c_1}{\partial z^2} - k_{d1} c_1 \\
\frac{\partial u_2}{\partial t} & = -u \frac{\partial c_M}{\partial z} + D_{cM} \frac{\partial^2 c_M}{\partial z^2} - k_{p} k_{0} c_M \\
\frac{\partial u_3}{\partial t} & = -u \frac{\partial \lambda_0}{\partial z} + D_{\lambda_0} \frac{\partial^2 \lambda_0}{\partial z^2} + 2f_k c_{1} - 2k_{tr} \lambda_0^2 \\
\frac{\partial u_4}{\partial t} & = -u \frac{\partial \lambda_1}{\partial z} + D_{\lambda_1} \frac{\partial^2 \lambda_1}{\partial z^2} + 2f_k c_{1} + k_{p} k_{0} c_M - k_{tr} \lambda_0 \lambda_1 \\
\frac{\partial u_5}{\partial t} & = -u \frac{\partial \mu_2}{\partial z} + k_{tr} \lambda_0 \lambda_2 \\
\frac{\partial u_6}{\partial t} & = -u \frac{\partial \mu_2}{\partial z} + k_{tr} \lambda_0 \lambda_1 \\
\frac{\partial u_7}{\partial t} & = -u \frac{\partial \mu_2}{\partial z} + k_{tr} \lambda_0 \lambda_1 \\
\frac{\partial u_8}{\partial t} & = -u \frac{\partial \mu_2}{\partial z} + k_{tr} \lambda_0 \lambda_1 
\end{align*}
\]
discretized on a grid with 200 equidistant points using the WENO schemes. This results in an ode model with 1600 states which is used throughout this paper. The analytical Jacobian of the model has been computed and will be used. A typical distribution of the monomer concentration along the reactor is shown in figure 2. The jumps happen at the beginning of a segment where a side injection takes place.

![Typical monomer distribution along the reactor](image)

**Fig. 2:** Typical monomer distribution along the reactor. The jumps occur at the positions of the reactor where the side injections of the monomer take place.

### 3. NONLINEAR MODEL PREDICTIVE CONTROL

In this work, we follow the idea of online optimizing control. The goal of our controller is to maximize the reactor economics under constraints on the product quality (Engell 2007). The controller utilizes the process model explicitly to optimize the process performance over a sufficiently long horizon by adapting the flow rates and the reactor temperature. For the numerical solution of the optimal control problem, the dynamic optimization problem has to be discretized. For this purpose, sequential optimization of single shooting based upon control vector parametrization, multiple shooting or full discretization can be used. In this work we have used the first one and the last one of these methods.

The iterative application of the optimizing controller requires to reinitialize the process model with the current states of the process at each time instance which provides feedback. In reality, the states are not available but only few measurements. State estimation for this tubular reactor has been discussed in (Hashemi et al. 2014) and as the focus here is on the numerical solution techniques, we assume the availability of the state vector in the controller. For a tubular polymerization reactor, the operation is most efficient for a desired conversion if the throughput is maximized i.e. the sum of all injections of monomer into the reactor in figure 1. The end-use properties of a polymer are highly dependent on its molecular weight. Therefore the weight average molecular weight of the produced polymer, \(M_w\), is considered as a quality constraint. Moreover, due to toxicological and environmental issues, the residual monomer in the produced polymer has to be kept below a threshold. Therefore the monomer content of the product, \(c_M\), is considered as another quality constraint for the optimizing controller. A faster and smoother transient behavior, as well as a computationally less demanding optimization problem, can be achieved if the quality constraints are formulated as soft constraints. We have shown in a previous work that the hard constraints can be replaced by soft constraints if an appropriate weighting factor for the constraints in the computed cost function is considered (Hashemi et al. 2013). Formulation of the quality constraints as soft constraints may result in short term violations of the constraints from the given bounds. However, if it is assumed that the produced polymer is post processed in large vessels and will be sold in larger quantities, a considerable degree of mixing after the production occurs and the short-time violations can be tolerated.

Considering the quality constraints as soft constraints, the optimal control problem then can be formulated as follows:

\[
\begin{align*}
\min_{\mathbf{S}} & \Phi(\mathbf{x}(t_k), \mathbf{u}(t_k), N_c, N_p) \\
\Phi &= -1000\Phi_1 + 2\Phi_2 \\
\Phi_1 &= \sum_{j=k}^{j=k+N_p} (u_{1k} + u_{2k} + u_{3k} + u_{4k}) \\
\Phi_2 &= 500\Phi_{21} + 3000 \left( \frac{1}{3950} \Phi_{22} + \frac{1}{3750} \Phi_{23} \right) \\
\Phi_{21} &= \sum_{j=k}^{j=k+N_p} (\max(c_{Mj} - cM_u, 0))^2 \\
\Phi_{22} &= \sum_{j=k}^{j=k+N_p} (\max(M_{wj} - M_{wu}, 0))^2 \\
\Phi_{23} &= \sum_{j=k}^{j=k+N_p} (\min(M_{wj} - M_{wl}, 0))^2
\end{align*}
\]

where \(\mathbf{S}\) is the vector of manipulated variables. The subscripts \(l\) and \(u\) denote the lower and upper bounds of the corresponding variables. \(k\) refers to the sampling time and \(N_p\) denotes the prediction horizon. Throughout this paper a sampling time of 500 seconds is used and the inputs are considered to be piece-wise constant between the sampling times. The product throughput is maximized through the first part of the cost function (\(\Phi_1\)) and the violations of the controlled variables from the given bounds are minimized through the second part (\(\Phi_2\)). The constants in the cost function have been chosen such that a compromise between fast transition and small violations of the constraints from the bounds is achieved.

In addition to the quality constraints, the optimal control problem is subject to two physical constraints as follows:

\[
\begin{align*}
0 \leq u_1, u_2, u_3, u_4 &\leq 100 \\
340 \leq T_{jac} &\leq 443
\end{align*}
\]

The monomer side injections are modeled through the opening of the controlled valves and the constraint eq. (14a) results from the limitations on their opening. Moreover, since the reactor is operated at a pressure of 8 bars, the highest temperature of the reactor before boiling the water inside is 443 K which is imposed by the constraint eq. (14b).

The solution of this problem is a sequence of inputs (\(u^*\)) for \([t_k, t_{k+N_p}]\). As usual, the sequence is applied only for the first sampling period (\(t \in [t_k, t_{k+1}]\)) and the inputs are kept constant until the next sampling time. The optimization problem is resolved at the next sampling instance after new measurements have become available.
4. SINGLE SHOOTING WITH CONTROL VECTOR PARAMETRIZATION

4.1 Problem formulation

Our first approach to formulate the optimal control problem is to use single shooting with control vector parametrization. In this approach, only the control inputs are discretized and an external integrator is used to compute the trajectories of the states. The advantage of this method is that it is simple to implement and the small size of the resulting optimization problem in terms of degrees of freedom. As the optimization problem in this approach is solved in a sequential manner, the dynamic model of the system is implicit in the constraints. The number of decision variables depends on the length of the control horizon. Because of the high computational load, in this study, the control horizon is set to one which means the control inputs are kept constant for whole prediction horizon. Therefore the decision variables of the optimizing controller for this formulation then are: \( S = [u_1, u_2, u_3, u_4, T_{jac}] \). The jacket temperature \( (T_{jac}) \) affects only the quality parameters \( c_M \) and \( M_w \). For the nominal flow rates, the residence time of the reactor system is about 2600 seconds. Since the controller manipulates the flow rates of the monomer side injections, the controller moves can cause a shift in the states or even postpone or expedite the effect of the previous control moves. In order to counteract this behavior and to capture the dynamic evolution of the process and to assure that the constraints are met, in this work, a prediction horizon of 3500 seconds with a sampling time of 500 seconds is used. The ode model is integrated using the CVODE from the Matlab interface of Sundials (SundialsTB) with the analytic Jacobian of the model. The optimization problem is solved using the SNOPT solver from the TOMLAB package where the solver uses numerical derivatives of the cost function.

4.2 Simulation Results

In this section a simulation result of the optimal controller described in section 4.1 is presented. It is assumed that the reactor system is in a steady state and produces the requested polymer with a designed set of inputs. At \( t = 0 \) the controller is switched on and the goal is to produce the same polymer with the maximum possible throughput. The quality constraints are: \( c_M \leq 10 \) and \( 375 \leq M_w \leq 395 \). The plant is simulated using the WENO scheme, and the same model is used in the single shooting optimization, so there is no model-plant mismatch. The manipulated variables generated by the optimizing controller are shown in figure 3. As it can be seen in this figure, the controller drives the reactor to an operation point with higher temperature and increases the sum of all feeds of monomer. This is achieved by increasing the side feeds in the first three sections while the feed in the last injection point is set to zero. This makes sense, as for higher throughputs the residence time in each segment is reduced and to achieve the desired conversion, the monomer is fed mostly at the initial locations. The final throughput is 2.12 times the initial throughput. The controlled variables for this simulation are shown in figure 4. The controller is able to keep the controlled variables within the desired bounds for whole simulation time.

5. FULL DISCRETIZATION

5.1 Problem formulation

As the second method, we use the full discretization approach and parametrize the states in the time domain using the orthogonal collocation scheme. The discretization of the spatial domain is performed as in the previous case by the WENO scheme in a grid of 200 equidistant points. Orthogonal collocation is a class of direct collocation methods and approximates the states of a system by a linear combination of orthogonal polynomials. Collocation methods are generally more difficult to implement and the resulting NLP problem is much larger. However, this method does not need an additional integrator to solve the model and the resulting NLP problems are sparse and structured which can be exploited by many solvers. The set of flipped Legendre-Gauss-Radau (LGR) collocation points were chosen for the discretization in the temporal domain (Patterson et al. 2009). This set differs from other popular orthogonal collocation methods in that it collocates the last point in the time interval while not using the first point. The basis Lagrange polynomials have the following form:

\[
L_i(t) = \prod_{j=0, j \neq i}^{N} \frac{t - t_j}{t_i - t_j}, \quad i = 0, ..., N \tag{15}
\]

where \( N \) is the order of the Lagrange polynomial. The states are then approximated as follows:

\[
x(t) \approx \sum_{i=0}^{N} x(t_i)L_i(t) \tag{16}
\]
It has to be noticed that the first point in the time interval of each finite element is not collocated but it is still included as an interpolation point in the sum of Lagrange polynomials. By differentiating the series and evaluating at the $k^{th}$ collocation point, it follows:

$$\dot{x}(t_k) \approx \sum_{i=0}^{N} x(t_i) \hat{L}_i(t_k)$$

(17)

$$\dot{L}_i(t_k) = D_{ki}$$

(18)

where $D_{ki}$ is called the Radau differentiation matrix. Considering the system model as:

$$\dot{x}(t) = f(x(t), u(t))$$

(19)

the dynamics of the system model is collocated at the $N$ Legendre-Gauss-Radau collocation points. By evaluating this approximated model at the $k^{th}$ point, the following equation is obtained:

$$\sum_{i=0}^{N} D_{ki} x(t_i) = f(x(t_k), u(t_k))$$

(20)

Equation 20 is the full discretization of the system and it is included in the optimization problem as a constraint. The decision variables of this scheme include the parametrized states as well as the input variables of the reactor system.

5.2 Simulation Results

In this section, the simulation result for the optimal control problem which is solved using the collocation scheme is presented. After several simulations, it turned out that a scheme with four finite elements for the prediction horizon and fifth degree Lagrange polynomials on each finite element is a good compromise between the approximation accuracy and the size and the structure of the NLP problem. The resulting NLP problem has 32,005 decision variables and figure 5 shows its sparsity pattern.

![Fig. 5: Sparsity pattern of the NLP problem using four finite elements and fifth degree Lagrange polynomials.](image1)

In order to evaluate the performance of the fully discretized scheme, a step response of the system obtained with the full discretization scheme is compared with the case that only the spatial domain is discretized and CVODE from the SundialsTB serves as the integrator. The result is shown in figure 6.

![Fig. 6: Step responses of the controlled variables. For this simulation the inputs have been changed from $u = \{20.000, 25.000, 30.000, 35.000, 350.000\}$ to $u = \{24.687, 26.953, 32.679, 36.864, 350.603\}$.](image2)

The controller uses the same cost function. The control horizon is one and the prediction horizon is seven as before with a sampling time of 500 seconds. The plant was simulated using CVODE where the WENO scheme was used to discretize the spatial domain. The controller, however, uses the fully discretized model. In this work we have used IPOPT\(^1\) as the solver which is suited for large scale sparse nonlinear optimization problems (Wächter and Biegler 2006). For this simulation, a distribution of this solver in the OPTI toolbox\(^2\), with the MKL PARDISO as the linear solver, was used. First order derivatives are provided analytically while a quasi-Newton approach is used by IPOPT to approximate the second order derivatives. The manipulated variables are shown in figure 7. It can be seen that the controller drives the reactor gradually to an operating point with higher temperature and increases the throughput 2.13 times. The inputs are slightly oscillating more than in the first simulation but have a shorter transient. Figure 8 shows the controlled outputs for this simulation. The controller is able to keep the outputs within the bounds with some violations which result both from using soft constraints and by approximation errors. Figure 9 shows the computation times per iteration of the two schemes. The full discretization scheme results on the average, in slightly shorter computation times. In contrast, the single shooting method is numerically more stable and keeps the bounds better. The steady states reached by the two controllers are slightly different which may be due to the settings of the solvers.

![Fig. 7: Manipulated variables generated by the optimizing controller with full discretization scheme.](image3)

\(^1\) https://projects.coin-or.org/Ipopt
\(^2\) http://www.i2c2.aut.ac.nz/Wiki/OPTI/
In this work, two numeric methods to implement an optimizing controller for a continuous polymerization reactor, modeled by a set of eight pdes, were studied. The reactor system has a complex behavior and the concentrations exhibit sharp fronts along the reactor length coordinate. The WENO scheme was used to discretize the spatial domain of the pde model to avoid the need for a fine discretization scheme in this dimension. As the first approach, we implemented the optimal control problem using single shooting and control vector parametrization. In this approach only the control inputs are discretized and an integrator is used to compute the trajectories of the states. This method has the advantage that it is simple to implement and a small optimization problem results. Moreover it is numerically more stable and the tuning of the solver is easier. However, since a complex integrator must be used, computation of the Jacobian and of the Hessian of the Lagrangian becomes more complex. The use of analytical Hessians in the full discretization scheme will have hardly an influence on the computation times. However, due to the long settling time of our reactor system, a longer control horizon needs a longer prediction horizon which can increase the size of NPL problem. The simulation results show that both formulations drive the reactor to a similar steady state and increase the product throughput considerably. For both implementations, the worst-case computation times are still too long (up to 6 times). The use of analytical Hessians in the full discretization approach could contribute to a further reduction, as well as further model simplifications in both schemes. The investigation of other discretization schemes, as well as the multi-grid approaches are further planned extensions of this work.

6. CONCLUSION

REFERENCES


3 Intel(R) Core(TM) i7-2600 CPU @ 3.40GHz, 24GB RAM