ADCHEM 2006

International Symposium on Advanced Control of Chemical Processes
Gramado, Brazil – April 2-5, 2006

OPTIMAL GRADE TRANSITION IN POLYMERIZATION REACTORS:

A COMPARATIVE CASE STUDY

Nitin Padhiyar, Sharad Bhartiya[1], Ravindra D. Gudi

Department of Chemical Engineering, Indian Institute of Technology,
Bombay, Mumbai-400 076, India

A study of grade transitions as encountered in polymerization reactors is presented. The results underscore the need for global optimization algorithms to fully realize the benefits of grade transition that are necessarily non-convex. For comparison purposes we use a non-gradient, parallel search, stochastic method, namely differential evolution (DE). Our simulations indicate that while the DE solution is highly dependent on the algorithm parameters and mutation strategy, the SQP solution depends on the initial guess value and consistently provides faster convergence. Finally, we also explore the issue of evaluating the optimal grade changeover time. All of the above issues have been demonstrated for the grade transition of polymethyl methacrylate (PMMA) in a non-isothermal CSTR.

Keywords: gradient optimisation; genetic algorithm, nonlinear programming, dynamic programming, off-line programming, process model, time optimal control

1. INTRODUCTION

In the contemporary synthetic polymer-manufacturing environment, flexible operation holds the key towards processing of diverse product orders and maintaining profitable operation in the face of fluctuating market conditions. This has led to a shift from single polymer grade production in large continuous reactors to high quality, low volume, multiple grade manufacturing. Scheduling of various grades in continuous reactors necessitates grade transition, typically characterized by significant perturbations in the process operating conditions. Despite this, the polymer industry routinely performs grade transitions. In fact, as many as 30 to 40 different grades may be produced in a polyolefin plant (Chatzidoukas, et al., 2003). The ultimate incentive for such process upsets lies in transition from one grade to another in a safe and optimal fashion.

Complex kinetics as well as energy effects usually give rise to complex dynamic behavior. Such complex dynamics typically make grade transition a non-convex optimization problem having multiple optima.

Recipes of manipulated variables to achieve optimal grade transition are typically calculated using a dynamic model in conjunction with optimization. The optimization reflects minimization of off-specification material production, which is an essential ingredient of any profitability analysis. In literature, the dynamic optimization problem is converted to a standard NonLinear Program (NLP) using Control Vector Parameterisation (CVP) and then solved using various gradient-based methods, for which efficient solvers exist. (McAuley and MacGregor, 1992; Takeda and Ray, 1999; Seki, *et al.,* 2001). In each of these works, the resulting nonlinear program was solved using CVP followed by a nonlinear, gradient-based optimization method such as Sequential Quadratic Programming (SQP). Moreover, the gradients of objective function were computed using finite difference. Wang et al (2000) used sensitivity of the ODEs to determine gradient information.

It is likely that the nonlinear behavior, which is further accentuated during grade transition often leads to a non-convex optimization problems characterized by multiple optima. Conventional gradient-based nonlinear optimization methods such as SQP may at best lead to only a local optimum. In

[1] Correspondence should be addressed to Sharad Bhartiya at bhartiya@che.iitb.ac.in

such a case, the benefits of grade transition may not be fully realized. It is therefore important to analyze the structure of the nonlinear program encountered in grade transition control and quantify the impact of the optimization technique on the quality of the solution. To study this aspect, we compare the SQP optimization technique with a stochastic, direct parallel search method, namely Differential Evolution (DE), which is widely recognized for its ability to provide global solutions to optimization problems.

In this work, we use a model representing the non-isothermal polymerisation system for production of polymethyl methacrylate (PMMA). The model has been presented in Section 2. Our results in Section 3 present two case studies. Case Study I investigates presence of multiple optima in the PMMA grade transition problem. Using a fixed grade changeover time, we find that the gradient-based methods such as SQP are likely to yield only a locally optimal solution, whereas the DE method typically provides a superior optimal solution in terms of the objective function value. In Case Study II, we explore the benefits of optimal grade changeover time relative to fixed transition times used in Case Study I. Finally, the work is concluded in Section 4.

2. GRADE TRANSITION IN PMMA PRODUCTION

As a case study, we choose grade transitions encountered in a nonisothermal free-radical polymerization of methyl methacrylate (MMA) using azobisisobutyronitrile (AIBN) as initiator and toluene as solvent (Daoutidis *et al.*, 1990). The same non-isothermal version of this model has been adapted as a test bed in our study. The model equations consist of molar balances for the monomer and initiator, reactor temperature, jacket temperature, and the zeroth and first moment of dead chains and have been reproduced below,

$$\frac{dC_m}{dt} = -\left[Z_p \exp\left(\frac{-E_p}{RT}\right) + Z_{fm} \exp\left(\frac{-E_{fm}}{RT}\right)\right] C_m P_0(C_I, T) + \frac{F(C_{m_{in}} - C_m)}{V} \quad (1a)$$

$$\frac{dC_I}{dt} = -Z_I \exp\left(\frac{-E_I}{RT}\right) C_I + \frac{(F_I C_{I_{in}} - F C_I)}{V} \quad (1b)$$

$$\frac{dT}{dt} = Z_p \exp\left(\frac{-E_p}{RT}\right) Cm \frac{\left(-\Delta H_P\right)}{\rho c_p} P_0(C_I, T) - \frac{UA}{\rho C_p V}\left(T - T_j\right) + \frac{F}{V}\left(T_{in} - T\right) \quad (1c)$$

$$\frac{dD_0}{dt} = \left\{0.5 Z_{T_c} \exp\left(\frac{-E_{T_c}}{RT}\right) + Z_{T_d} \exp\left(\frac{-E_{T_d}}{RT}\right)\right\}\left[P_0(C_I, T)\right]^2 + Z_{f_m} \exp\left(\frac{-E_{f_m}}{RT}\right) C_m P_0(C_I, T) - \frac{FD_0}{V} \quad (1d)$$

$$\frac{dD_1}{dt} = M_m\left[(Z_p \exp\left(\frac{-E_p}{RT}\right) + (Z_{f_m} \exp\left(\frac{-E_{f_m}}{RT}\right)\right] C_m P_0(C_I, T) - \frac{FD_1}{V} \quad (1e)$$

$$\frac{dT_j}{dt} = \frac{F_{cw}}{V_o}\left(T_{w_o} - T_j\right) + \frac{UA}{\rho_w C_w V_o}\left(T - T_j\right) \quad (1f)$$

$$MW(t) = \frac{D_1(t)}{D_0(t)} \quad (1h)$$

where

$$P_0(C_I, T) = \left[\frac{2 f^* C_I Z_I \exp\left(\frac{-E_I}{RT}\right)}{Z_{T_d} \exp\left(\frac{-E_{T_d}}{RT}\right) + Z_{T_c} \exp\left(\frac{-E_{T_c}}{RT}\right)}\right]^{0.5} \quad (1g)$$

where C_m and C_I represent the molar concentrations of monomer and initiator, respectively. D_0, and D_1 represent the zeroth and first moments of the molecular weight distribution of the dead chains, respectively, and T, and T_j represent reactor and jacket temperatures, respectively. The process parameters values have been taken from Daoutidis *et al.* (1990). We consider a single grade change in the number average molecular weight MW_n of PMMA from 25,000 to 35,000 by manipulating initiator (F_I) and coolant (F_{cw}) flow rates. The steady-state operating conditions for the two grades were evaluated using the model in Equation 1 and are listed in Table 1. The SQP algorithm was implemented using the *fmincon* function in MATLAB's optimization toolbox, whereas the DE algorithm was coded in MATLAB.

Table 1. Steady state operating conditions for the initial and final grades

Grade indices	**Initial Grade**	**Final Grade**
Molecular weight	25000	35000
$\mathbf{F_{cw}}$, $\boldsymbol{m^3/hr}$	3.2636	9.2944×10^{-1}
F_I, m^3/hr	1.6883×10^{-2}	6.596×10^{-3}

3. RESULTS AND DISCUSSIONS

3.1. CASE I. Grade transition in PMMA shows multiple optima

We consider a grade transition in PMMA product from a grade characterized by molecular weight of 25,000 to a new grade with molecular weight of 35,000. The quantity of off-specification material produced and the grade changeover time may be represented by the following objective function,

$$J_1 = \int_{t_0}^{t_f} \left\| MW(t) - MW_{sp} \right\|_P^2 \, dt \quad (2a)$$

where $\left\| \cdot \right\|_P$ denotes the deviations form the target,

MW_{sp}, suitably weightedby P. Minimizing the objective function (Equation 2a) along with the process model (Equation 1) and following constraints (Equations 2b-2f) represent the optimization problem

$$MW_{\min} \le \max\left(MW(t)\right) \le MW_{\max} \quad (2b)$$

$$F_{cw,\,\min} \le F_{cw}(t) \le F_{cw,\,\max} \quad (2c)$$

$$F_{I,\min} \le F_I(t) \le F_{I,\max} \quad (2d)$$

$$\Delta t_{\min} \le \Delta t_{ij} \le \Delta t_{\max}\ ,\ \ i = 1,2\ ;\ \ j = 1,2,...,na \quad (2e)$$

$$\sum_{j=1}^{na} \Delta t_{ij} = t_f \quad (2f)$$

where Δt_{ij} represents the j^{th} switching interval of the i^{th} manipulated variable and na represents the number of times that the coolant flow rate (i=1) and the initiator flow rate (i=2) are switched during the transition horizon. The objective function penalizes squared deviations from the molecular weight of the new grade over the transition horizon, $t_f - t_o$. Equation 2b represents output constraints, i.e. the transients on the molecular weight, whereas Equations 2c,d represent constraints on the inputs. Equation 2e ensures that the switching interval is of an adequate size. Since in this case study, the grade changeover time is known a priori, we used the equality constraint Equation 2f to reduce one degree of freedom.

Equation 2 may be solved with a standard NLP solver through use of CVP, where manipulated variables are parameterised and approximated by a series of trial functions γ. Thus, for the i^{th} manipulated variable,

$$u_i(t) = \sum_{j=1}^{na} a_{ij}\gamma_{ij}(t - t_{ij}) \quad (3)$$

where t_{ij} is the j^{th} switching time of the i^{th} manipulated variable, na as explained above is the number of switching intervals, and a_{ij} represents the amplitude of the i^{th} manipulated variable at the switching time t_{ij}. We make use of a ramp trial function as shown in Fig. 1 since a ramp can approximate the optimal profile more closely relative to a zero order trial function. Gradient-based SQP further requires the gradient of the objective function and constraints. In this work we used sensitivity of the ODEs to evaluate the gradient of the objective function as follows (Leis and Kramer, 1988),

$$\dot{\mathbf{S}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{S} + \frac{\partial \mathbf{f}}{\partial \mathbf{p}} \quad (4a)$$

where the (i,j) component of the sensitivity matrix is defined as follow,

$$S_{ij} = \frac{\partial x_i}{\partial p_j}\ \ ,\ \ S_{ij}(0) = 0\ \ ;i = 1,...,n; j = 1,2,...,m \quad (4b)$$

Where n and m represent the number of states, **x,** and the decision variables, **p**, respectively.

In the current section, the transition horizon is fixed at 4 hours. The constraint in Equation 2f ensures that the sum of the switching intervals equals the transition horizon. The constraint values are summarized in Table 2. All simulations presented in the current Case study use three optimized switching intervals. Thus the decision variables include two amplitude and two switching times for each manipulated variable, the coolant flowrate and initiator, thereby making a total of eight variables.

Table 2. Bound constraints for variables associated with the NLP of Equation 7.

	Minimum	Maximum
MW	24,800	35,200
Fcw, m^3/h	0.3	6
F_I, m^3/h	0.0007	0.1
Δt, h	0.016667	3.96
t_f, h	0.0501	4

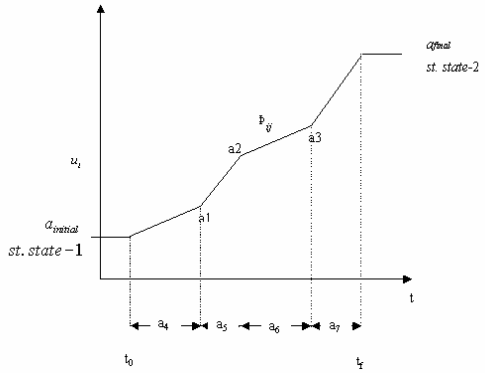

Fig. 1. Ramp trial function for parameterization of the manipulated variables

It is well known that gradient-based optimization methods such as SQP fail to provide the global optimum if the optimization problem has multiple optima. Dependence of the optimal solution on the choice of the initial guess value in a gradient-based algorithm such as SQP constitutes an empirical verification of the presence of multiple optima. Table 3 documents the values of the objective function J_1 for 10 different optimization runs using SQP each with a distinct initial guess of the decision variables. Simulation No. 5, 6, 7, and 8 yielded solutions that represent identical points in the decision space. The remaining simulations resulted in distinct solutions. Thus, the results indicate the presence of several optima. The tolerances used by the integrator and the optimizer in obtaining these solutions have been reported in Table 4. We have verified by varying the tolerance values that the difference between any two different optima is not due to a particular tolerance value used by the SQP solver. We also observed that use of perturbation methods to evaluate the gradient adversely affected the quality of the optimal solution in each of the 10 cases. For example, in simulation No. 10 in Table 3, use of finite difference to calculate gradients resulted in a minimum objective function value of 2.2533 x 10^7 as opposed to 1.1457 x 10^7 obtained when using sensitivity information (see Equation 4). Furthermore, solutions based on finite differences for

calculation of gradients were strongly dependent on the particular set of tolerance values used by the optimizer. The number of function evaluations required during optimization for each of the 10 simulations is also reported in Table 3. It has been widely claimed that the DE algorithm is a promising candidate for obtaining the global optimum in such cases. However, the quality of the solution and the computational expense is determined by the choice of its parameters such as the population size, the maximum number of generations as well as the mutation and crossover factors. To compare the relative performances of DE and SQP for grade transition from a steady molecular weight of 25,000 to 35,000, 10 optimization runs were performed using DE and the results documented in Table 3. In case of all DE based solutions, we utilized a population size of 40 members and the maximum number of generations was set to 200. Eight out of the 10 runs using DE were distinct from each other based on the choice of the members of the initial population. The mutation and crossover factors were decided by trial and error and were selected as 0.5 and 0.75, respectively. However, a subsequent study of the sensitivity of the optimal solutions using DE to the algorithm parameters shows that the value of 0.75 for both the factors is optimal. As noted from Table 3 we arrived at several DE solutions at the end of 200 generations. As can be seen from Table 3, the best solution using either DE or SQP corresponds to an objective function value of 1.0698 x 10^7. The subsequent six best solutions (simulation No. 2-7) obtained using the DE algorithm were very close to this best solution. The top seven SQP solutions showed a larger spread from this best solution. However, the last three simulation results using DE were worse than the worst SQP solution. This is due to the following mutation strategy used in this work

$$\hat{\mathbf{u}}_i^{g+1} = \mathbf{w}_b^g + F_m(\mathbf{w}_{r_1}^g - \mathbf{w}_{r_2}^g), \qquad i = 1,...N_p \qquad (5)$$

Since the mutated members represented perturbations around the best member of a given generation (based on elitism), it is likely that the members of the new generation were trapped in the vicinity of a local minimum. To verify this notion, we used a different mutation strategy for simulation No 10 as follows,

$$\hat{\mathbf{u}}_i^{g+1} = \mathbf{w}_{r_3}^g + F_m(\mathbf{w}_{r_1}^g - \mathbf{w}_{r_2}^g), \qquad i = 1,...N_p \qquad (6)$$

Since vectors $\mathbf{w}_{r_i}^g$ represent randomly selected and distinct members of generation *g*, the mutated members using this strategy enhance the diversity of the new generation relative to the mutation strategy used in Equation 5 where the mutated members represented perturbations around the best member of generation *g*. This mutation strategy was found to give substantial improvement to the solution when used alone; however when further accompanied with an increase in the number of generations from 200 to 300, the results for simulation No. 10 yielded a minimum objective function value of 1.10253 x 10^7.

We have used the number of objective function evaluations needed by the two methods as a measure of the computational expense. Although the best individual solutions yielded by SQP and DE have identical objective function values of J_1^*=1.0698 x 10^7 and represent an identical point in the decision space, the SQP algorithm is more computationally efficient by an average factor of 4.9. The number of function evaluations needed to obtain the best individual solution of SQP and DE were 1669 and 6331, respectively. All the 10 distinct solutions resulting from the different initial conditions used in SQP satisfy the KKT conditions for optimality and therefore confirm that grade transition in PMMA production exhibit multiple local optimal solutions. Figs. 2 and 3 show the transition profile of the molecular weight and the manipulated variables, respectively, using the best DE (solid line) and the best SQP (dash-dotted line) solutions. Since the two strategies yielded identical solutions, the lines are indistinguishable. We note that although we used a fixed changeover time of four hours, the integration necessary in computation of the objective function in Equation 9a is performed over six hours. The extra two hours enables penalizing deviation of the steady state output from the new grade properties. This feature of integration has been maintained in all subsequent results. The molecular weight profiles obtained using the best DE and best SQP solutions show identical rise times of 0.28 hr, and settle to within ± 100 Kg/Kmol limit of the steady state value in 1.37 hr. Fig. 2 also shows that the transition profile achieved by changing the two flow rates to their respective new grade steady state values in a step-like manner does not provide the optimal transition trajectory (see dotted line). Fig. 3 shows that the input recipes generated by the two methods make aggressive moves at the start of transition. However the lower bound on the size of switching intervals ensures that at least one minute elapses between two successive moves (see Table 2 for constraints values).

3.2. Grade changeover time as a decision variable

All simulations presented until now optimized switching time intervals but used a fixed time horizon of 4 hr. However, it is interesting to note that the defacto changeover times were less than the time horizon of 4 hr. For example, the best DE solution in Case Study I shows that the grade transition was accomplished in only 1.37 hr. (see Fig 2). One could also explicitly optimize the transition horizon by relaxing the equality constraint equation 2f and

Table 3. Summary of the objective function values and CPU times needed to perform 10 simulation runs for each of the two algorithms, namely SQP and DE

Sim No	SQP: Gradients using sensitivity equations		DE: Gradient free	
	J^* x 10^{-7}	No of function evaluations	J^* x 10^{-7},	No of function evaluat-ions
1	1.0698	1669	1.0698	6331
2	1.0818	1615	1.0708	6841
3	1.0991	1432	1.0945	7066
4	1.1011	2028	1.1016	6040
5	1.1313	586	1.1031	5804
6	1.1313	816	1.1045	6895
7	1.1313	881	1.1045	7109
8	1.1313	1303	1.1917	7597
9	1.1347	2056	1.2268	7373
10	1.1457	1437	1.2268	7373

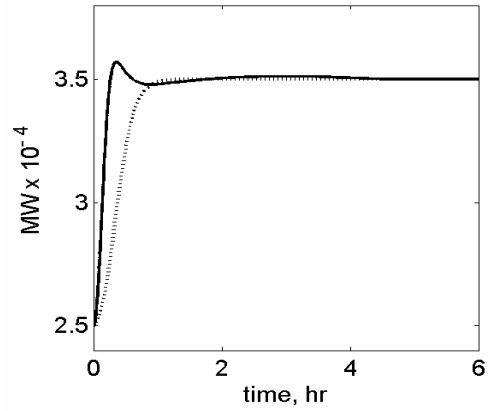

Fig. 2. CASE I: Optimal transition trajectory of molecular weight of PMMA, using DE (solid line) and SQP (dash dotted line) algorithms for solution of Equation 2. The dotted line represents the transition if the manipulated inputs corresponding to the new grade were implemented in a single step (non-optimal).A fixed time horizon of 4 hr is assumed. P = 1.

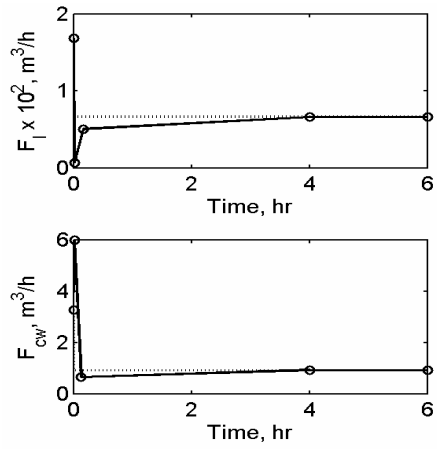

Fig. 3. CASE I. Optimal transition trajectory of manipulated inputs, namely initiator and coolant flow rate, using DE (solid line) and SQP (dash dotted line) algorithms.

adding t_f as an extra degree of freedom. The current case is similar to Case Study I in all respects except that in Equation 2f t_f is considered as a decision variable. Thus, we treat the *na*-1 switching intervals and t_f as the time related decision variables. As with the Case Study I, ten optimization runs starting with different initial guesses for SQP and 10 different initial populations for DE were performed. Figs. 4 and 5 show the molecular weight transition and the manipulated variable profiles, respectively, and represent the best of the ten runs for DE (solid line) and SQP (dashed line). The optimum transition horizons, t_f are 0.37 hr and 0.28 hr for DE and SQP, respectively. Although the grade changeover time is marginally higher when using DE, the rise time is same (0.23 hr). The settling times using the two methods are almost same (0.71 hr). The DE and SQP methods converge to the respective objective function values as J_1^*= 1.0189 x 10^7 and J_1^*=1.0419x 10^7. The lower values of objective functions obtained here relative to Case Study I (J_1^*=1.0698 x 10^7 for DE and SQP), indicate improved performance with both DE and SQP when the transition horizon is also optimized. Another unobvious advantage of optimization of transition horizon is observed upon comparing the manipulated variable profiles of Case Study I (Fig. 3) and the current case (Fig. 5). Optimizing the transition horizon results in achieving the new steady state for the manipulated variables faster (t_f =0.37 hr for DE and 0.28 hr for SQP) than for the case when the transition horizon is assumed fixed at 4 hr. This behavior would translate into benefits for grade transition and reduced transient operation and therefore tighter control on the product quality. Convergence of the SQP solution to a higher objective function value relative to DE demonstrates susceptibility of SQP to local minima, as also observed in Case Study I. In the current case study, both the standard deviation and average value of objective function for DE (J_{std}=2.44 x 10^5, J_{avg} =1.0446 x 10^7) were found to be lower than those for SQP (J_{std} = 8.8073 x 10^5, J_{avg} =1.13399 x 10^7). It is interesting to note that relaxation of the constraint of Equation 2f resulted in fewer number of function evaluations needed by SQP in the current Case Study relative to Case Study I. The average number of function evaluations was 569 as opposed to 1382 in Case Study I.

.

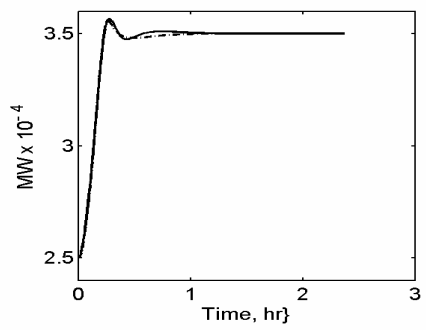

Fig. 4. CASE II. Optimal transition trajectory of molecular weight of PMMA, with DE (solid line) and SQP (dash dotted line) algorithms which optimized the transition times in addition

to the manipulated inputs and switching intervals. P = 1.

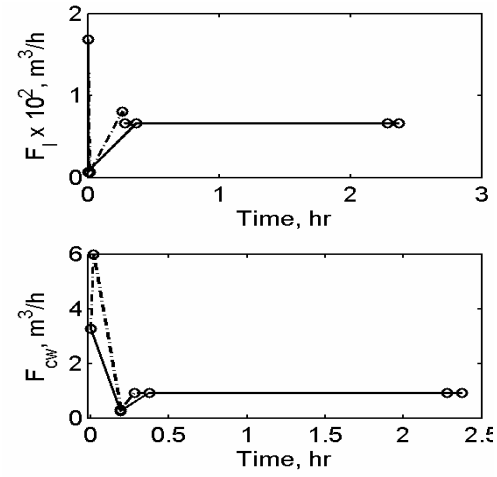

Fig. 5. CASE II. Optimal transition trajectory of manipulated inputs namely initiator and coolant flow rates, with DE (solid line) and SQP (dash dotted line) algorithms.

4. CONCLUSIONS

Optimal grade transition has emerged as a key strategy in the polymer industry, notably in gas-phase polymerization of olefins, to enable flexible operation. Our studies reveal that grade transition encountered during PMMA production results in a non-convex optimization problem having multiple optima. Thus, widely used solvers such as SQP may provide an inferior transition policy. Although DE has attributes consistent with global optimization, the DE solution is highly dependent on the mutation strategy as well as its parameters. Also, we demonstrate that explicit optimization of transition time yields solutions where the transition policies quickly converge to the new grade steady state values in addition to reducing the transition horizon.

Table 4: Tolerances used by the ODE solver (*ode15s* of MATLAB® 6.5) and the optimizer (*fmincon* in MATLAB® 6.5) in obtaining the SQP based solutions reported in Table 3.

Integration Tolerances	Absolute error	1×10^{-6}
	Relative error	1×10^{-10}
Termination Tolerances Used by Optimizer	Decision Variable	1×10^{-10}
	Constraint violation	1×10^{-9}
	Objective Function	1×10^{-8}

ACKNOWLEDGEMENTS

The authors acknowledge funding from the Department of Science and Technology, India, for financial assistance under SERC (Engg. Sciences) (#03DS015).

REFERENCES

Chatzidoukas C, Perkins JD, Pistikopoulos EN, Kiparissides C. Optimal grade transition and selection of closed-loop controllers in a gas-phase olefin polymerization fluidized bed reactor. Chemical Engineering Science. 2003;58:3643-3658.

McAuley KB, MacGregor JF. Optimal grade transition in a gas phase polyethylene reactor. AIChE J. 1992;38:1564-1576.

Takeda M, Ray WH. Optimal grade transition strategies for multistage polyolefin reactors. AIChE Journal. 1999;45:1776-1793.

Seki H, Ogawa M, Ooyama S, Akamatsu K, Ohshima M, Wang Y. Industrial application of a nonlinear model predictive control to polymerization reactors. Control Engineering Practice. 2001;8:819-828.

Wang Y, Seki H, Ohyama S, Akamatsu K, Ogawa M, Ohshima M. Optimal grade transition control for polymerization reactors. Computers and Chemical Engineering. 2000;24:1555-1561.

Daoutidis P, Soroush M, Kravaris C. Feedforward/Feedback Control of Multivariable Nonlinear Processes. AIChE Journal, 1990; 36:1471–1484.

Leis JR, Kramer MA. The simultaneous solution and sensitivity analysis of systems described by ordinary differntial equations. ACM Transactions on mathematical Software. 1988; **14**: 45-60.