1987

Developing targets for the performance index of a chemical reactor network

Luke Achenie
Carnegie Mellon University

Lorenz T. Biegler

Carnegie Mellon University.Engineering Design Research Center.

Follow this and additional works at: http://repository.cmu.edu/cheme

Developing Targets for the Performance index of a Chemical Reactor Network

by

L. Achenie and L. Biegler

DEVELOPING TARGETS FOR THE PERFORMANCE

INDEX OF A CHEMICAL REACTOR NETWORK

by

Luke Achenle and Lorenz Blegler

Presented at the AICHE Annual Meeting, New York, Nov. 1987

Department of Chemical Engineering, Carnegie Mellon University,
Pittsburgh, PA 15213.

Abstract

The target approach to reactor synthesis tries to determine an upper bound for the performance of a chemical reactor network as measured by its objective function (such as yield or reactor volume). This leads to an optimal control formulation in which the residence time distribution and the micromixing function are the control variables. The optimal control problem is solved using a gradient based algorithm that employs successive quadratic programing and adjoint variables. Examples are given to illustrate the approach taken.

1. Introduction

A chemical reactor is frequently the key unit in a chemical process. While the overall profitability of the chemical plant depends on the cost of the various units (heat exchangers, separators for example) and on everyday operating costs, in this work we focus on the reactor design in isolation because it often affects downstream concerns. An extensive amount of work has been done in this area (Levenspiel, 1972, Aris, 1961); however, most of the approaches have been based on heuristics and on graphical techniques. For example, Nestoridis et al (1986) recently used convex analysis to show that the yield of a reaction (occuring in segregated flow) is maximized if such a reaction is carried out in a combination of two plug flow reactors in parallel. While this analytic approach is encouraging the assumption of segregated flow and the convexity of the objective function (in this case yield) is limiting. The algorithmic approach that we use in our research, is a viable alternative especially when the kinetics and the performance index (yield, selectivity, reactor size etc.) are complicated functions of concentration and other process variables and parameters.

We pose the reactor synthesis problem as a two part problem as follows:

1. Given the kinetics of a homogeneous reaction what combination of ideal reactors, namely .CSTR's, recycle reactors and PFR's will maximize a given performance index?

2. What is an upper bound (target) on the performance index irrespective of the reactor type and configuration?

Optimal reactor networks for homogeneous reactions depend to a large extent on mixing, temperature effects, reaction times, and sometimes on the amount of a catalyst. To tackle problem (1) we postulate a superstructure of reactors, together with a set of decision variables that exhibit the effects of mixing, temperature and reaction times. By manipulating these variables an optimal subnetwork of reactors is derived (Achenie and Biegler, 1986). However, the optimal structure is heavily dependent on the complexity of the superstructure, which may be characterized by how many substructures it contains.

The need for an independent measure of the performance index, free of restrictions imposed by the reactor type and configuration is thus clearly present. We propose to solve the target problem by determining the optimal micro and macro mixing profiles for a reactor model that takes into account both kinds of mixing.

For isothermal homogeneous reactions occuring in a reactor with premixed feed, Adler and others (1972) have shown that two extremes of mixing put a bound on the yield of a single reaction that has a concave or convex rate function. These mixing extremes are realized in a completely segregated reactor and a "maximum mixed" reactor (Zwietering, 1959) respectively. It is immediately apparent that for more complex kinetics, and for performance indices other than yield, the two extremes of mixing will not give rise to extreme values of the performance index. In fact a reactor with an intermediate degree of mixing may result in the maximum yield or selectivity.

To predict or simulate the effect of mixing on the throughput of a reaction (which may involve many elementary steps) a number of mixing models for reactions in reactors with both premixed and unmixed feeds have been proposed (see Rao and Edwards, 1973, for a comparison of some of these models). Features that are common among these models are the macro and micro mixing concepts. Macro mixing is characterized by the residence time distribution. On the other hand the micro mixing level is a measure of how intimately molecules mix with each other, and how early or late this mixing occurs (Levenspiel, 1972). By combining these concepts all the mixing models are capable of simulating various degrees of mixing, between completely segregated and maximum mixedness, with varying degrees of success.

Three of the models that, in our opinion, are suitable for optimization are due to Rippin (1965), Villermaux (1969) and Jackson et al (1986). In Rippin's two-environment model, molecules of the fresh feed initially enter a segregated environment. A fraction of molecules transfer to the maximum mixed environment at a rate proportional to the total number of molecules in the segregated region. By varying the transfer coefficient, h, from zero to infinity complete segregation, intermediate mixing and maximum mixedness are simulated. Since this **model** is **a** one parameter model its range of applicability is somewhat limited. Jackson's model can be thought of as a number of plug flow reactors, each with a different residence time, in parallel. By allowing these reactors to exchange material at various "remaining lifes" all degrees of mixing are simulated. While this approach is elegant and one of the more realistic ones, the resulting formulation is currently too complicated for optimization.

The third model, Villermaux's model, is closely related to Rippin's. Here also, there is a segregated environment and a maximum mixed environment. He postulated that a

fraction h of the molecules that will remain in a reactor for t + At units of time will spend all this time in the maximum mixed environment. The remaining fraction, (1-h), will spend all their time in the segregated environment. By choosing a different fraction , h, for different aggregates of molecules one obtains a distribution h(t), which describes the micro mixing pattern. The residence time distribution (RTD) remains intact and it characterizes the macro mixing in the system. That there is no exchange of material between the two mixing environments, in the Villermaux model, appears unrealistic.

2. Optimal Control Problem

Of the three models the Ng and Rippin model appears to have the best tradeoff between a realistic model and one that is easy to implement in an optimization algorithm. The model is appropriate for homogeneous reactions involving low viscosity reactants, in continuous flow reactors. In the model the transfer coefficient, h, is a constant based on the assumption that all molecules in the segregated environment (irrespective of age) have the same probability of transferring to the maximum mixed environment. In our implementation we have increased the range of applicability of the model by allowing h to be dependent on the age of a molecule in the segregated environment.

With this modification, the model becomes a multi-parameter model since it is a continuous function of age. The modification is justified on the grounds that molecules that have been in the segregated environment for a longer period of time will tend to have a greater probability of transferring to the maximum mixed environment, than the younger molecules. For generality, however, we do not assume that h is a monotonic function of age.

Using the modified Rippin model as a basis, an optimal control problem (solved as a nonlinear program or NLP) can be formulated in which the objective function is the performance index, and the decision variables are the macro mixing function (the RTD) and the micro mixing function (the transfer function, h). The mass and energy balances and the bounds on the decisions form the set of constraints in this problem. To solve the optimal control problem we approximate the residence time density function (associated with the RTD) and h(t) by a set of basis functions on finite elements. NLP is solved using successive quadratic programing (SQP). With the optimal control formulation one can obtain solutions to the following problems:

    1. For a given RTD what micro mixing profile will maximize the performance index?

    2. Given a micro mixing function what RTD will maximize the performance index?

    3. What are the optimal RTD and micro mixing profiles that will give the optimal performance index?

Since the mixing models use the residual life and age concepts the optimal control problem may not yield the c ails of the reactor type and configuration. Therefore the performance index thus ob:. 3d is a bound on the performance index that can be obtained using the superstruc ure approach to reactor network design. In other words, how good the reactor superstructure is, can be measured by how close its optimal performance index is to that obtained in the residual life/age domain. Although the modified Rippin model is based on isothermal homogeneous systems, it can easily be extended to adiabatic systems. Our optimal control formulation can be made to include species dependent RTD and micro mixing functions. Thus a target for the performance index of a reactor superstructure which has component splits (by distillation for instance) can be calculated.

3. The Modified Rippin Model

The derivation of the modified Rippin model very closely parallels the original work by Ng and Rippin, and therefore will not be repeated here. However, with the modification made in the transfer function, h (Rippin uses R) it is appropriate to point out the changes that result.
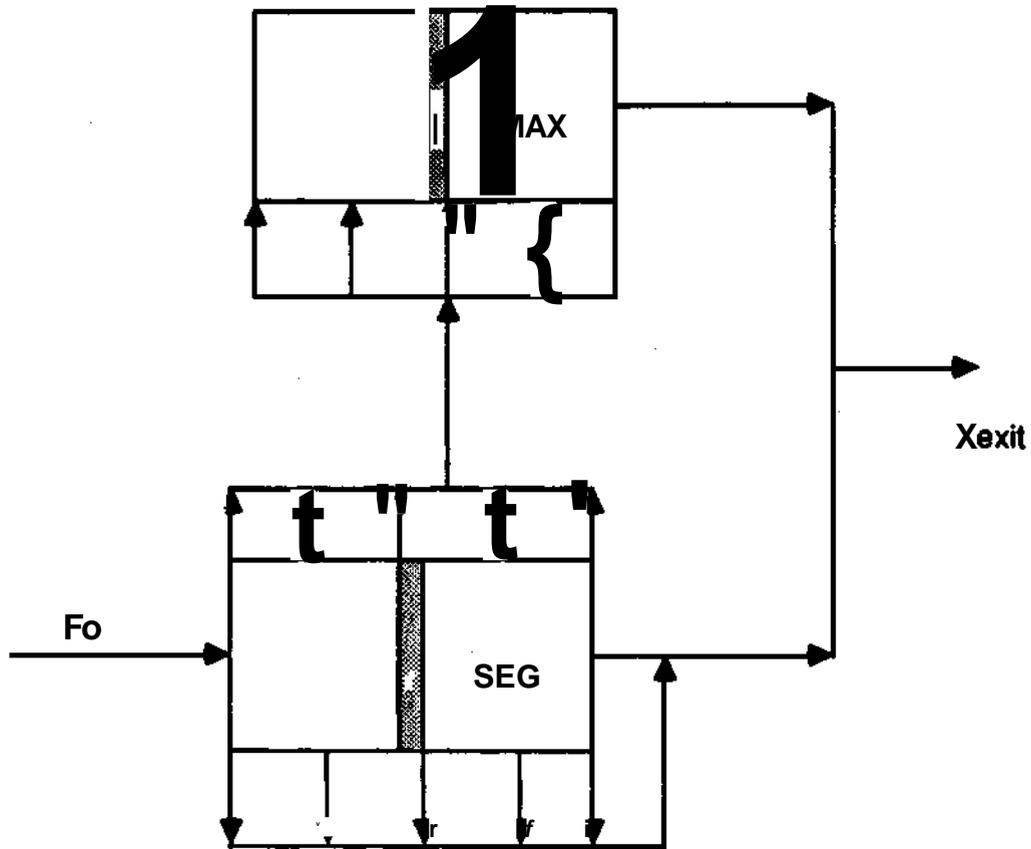
Figure 1: Schematic of Rippin Model

In figure 1 "max" and ^{N}seg^N refer to the maximum mixed and the segregated environments respectively. F_o is the fresh feed flow rate and X_{exit} is the average exit concentration vector. The key assumption is that the rate of transfer of material of age, t, from the segregated environment to the maximum mixed environment is proportional to the amount of material of age t remaining in the segregated environment. If m(t) is the amount of material of age t then

$$\frac{dm}{dt} = - h(t)m(t)$$

which implies $m = m_o\exp(-g(t))$ (1):(1)*

where $g(t) = \int_0^t h(s)ds$

and $m_o = m(O)$. Rippin's equations (4)* through (5)# then become

$$V\frac{d\lambda}{} \int^{\infty} (1-e^{-g(\alpha)})f(\alpha+\lambda)d\alpha = II.Vd\lambda$$

$$V \wedge f \ "h(a)e-8<^{a})f(a+X)da =III.VdX$$
$$x Jo$$

(3):(5)*

$$V\frac{d}{}2:f'^{o}h(a)C_{bat}(a)e-fl(^{o})f(a+X)da= IV.c_{0}VdX$$
$$X Jo$$

(4):(7)*

$$\frac{dC}{dx}\underset{}{} (X) = -mH_aC_x) + \frac{f^{\circ\circ}[C_{max}(X) - C_{bat}(a)]h(a)e-9<«)f(a_+X)da}{\int_{Jo}(1.0- e^{-g(\alpha)})f(\alpha+\lambda)d\alpha}$$

(5):(8)

then $C_{ex[t} = (1.0 - J_0^{\circ}/_0-s<^{*}>f(a)da)C_{max}(0) + J\wedge C\wedge taJe-oW)!(a)da$

(6):(9)#

where (*)* are equation numbers from Ng and Rippin's paper, and

C_{bat} = concentration profile from batch reactor kinetics (segregated environment)

C_{max} = concentration profile in the maximum mixed environment

a = age of a molecule

X = residual life of a molecule

C_o = concentration of feed to reactor

R = rate of production of a species by reaction

V = volume

f = residence time density function (note that the density function and the distribution function are used interchangeably in this paper)

h = micro mixing function

t = time

C_{exit} = average concentration at the exit of the reactor

The batch concentration profile is defined by

$$\wedge = R(C_{bat}(\alpha))$$

(7)

Equation (5)* is solved using the Zwietering boundary condition

$$\mathcal{E} 1 \wedge = 0 \text{ at } \quad X = \infty. \tag{8}$$

ok

This boundary condition makes it necessary to integrate (5) backwards using $C_{max}(\infty)$. The latter is a solution to the algebraic equation (9) resulting from equation (5)* and the boundary condition, (8).

$$0 = -R(C_{max}) + \lim_{\lambda \to} \left. {}_H -2 \frac{\int [C_{max}(\lambda) - C_{bat}(\alpha)] h(\alpha) e^{-g(\alpha)} f(\alpha+\lambda) d\alpha}{\int_{Je} P(1.0 - e^{-9(^a>)}) f(a + \lambda) d\alpha} \right. \tag{9}$$

Equation (9) is usually very difficult to solve. However, in a chemical reaction system it is generally the case that $C_{max}(oo)$ is finite. Glasser and others (1986) have shown that for some reaction systems different values for $C_{max}(oo)$ yield the same $C_{max}(0)$, which is needed in the expression for C_{exit}, the average exit concentration. Thus one can in certain cases avoid solving (9), and instead guess a value for $C_{max}(\infty)$ to be used in integrating (5)* backwards. Experience has given credence to this finding.

4. Formulation of the Nonlinear Program

For convenience we normalize the concentration vector C by letting $X = C/C_o$. Also let $\mathcal{E} = a + X$, then equation (5) becomes

$$\frac{dX_{max}(\lambda)}{d\lambda} = -R(X_{TM}) + \frac{\int_\lambda^\infty [X_{max}(\lambda) - X_{bat}(\xi-\lambda)] h(\xi-\lambda) e^{-g(\xi-\lambda)} f(\xi) d\xi}{\int (1.0 - e^{-g(\xi-\lambda)}) f(\xi) d\xi} \tag{10}$$

$^{X}maxH - {}^{finite}\ll {}^{*e} [°.\sim 1$

Also equations (6) and (7) become

$$X_{exit} = \left(1.0 - \int_0^\infty e^{-g(\alpha)} f(a) da \right) X_{max}(0) + J_{o\backslash at}(a) e^{-9<°0)} f(a) da \tag{11}$$

$$\frac{X_{bat}}{d\alpha} = R(X_{bat}(\alpha)) \tag{12}$$

$$X_{bat}(0) = X_0 \tag{13}$$

It is convenient to rewrite (11) as

$$X_{exit} - (1 - 0 - J \underset{0}{V}^{\circ} <^a> f(a)da)X_{max}(0) + I_{seg}(oo) \tag{14}$$

where we have used the fact that $f \sim f(a)da = 1$ (15)

Jo

and $^{\wedge} = e-9<\ll)f(a)X_{bat}(a)$ (16)

$I_{seg}(0) = 0$ (17)

We next approximate $t = {}^{\circ\circ}$ by ${}^{\wedge\wedge}$ and divide the interval $[0, t_{max}]$ into N finite elements,
$[t_j, t_{j+1}]$,
such that $t_j < t_{j+1} \forall J \in [1,N]$.

Here $t_1 = 0$ and $t_{N+1} = t_{max}$. The present formulation of the target problem has the following features:

(i) an objective function that depends only on exit concentrations

(ii) constraints on the average residence time.

It should be noted however, that the objective function can easily be expanded to include residence time, which is a measure of reactor volume and hence the investment cost of the reactor. Other costs associated with the reactor can be included in the objective function, as long as these costs can be expressed as functions of exit concentrations and residence time. The expanded objective function will change only the boundary conditions on the adjoint variables to be discussed in the following sections.

For the target problem we choose an objective that is a function of exit concentrations, and optimize it subject to the mass balances in the modified Rippin model. Here we define $J = J(X_{Qxit})$ as the objective function, and choose as decision variables the piecewise continuous functions h and f (being the micro and the macro mixing functions respectively).

With the above definitions the target problem can be cast from an optimal control problem to a nonlinear program defined over N finite elements as follows:

(P1) $\max J = J(X_{exit})$

subject to

$$\frac{dXmaxft)}{dt} = -\frac{R}{M}I^A_{max}J + \frac{\int^{t_{max}}_{Jt}[X_{max}(t) - X_{bat}(\xi-t)]h(\xi-t)e^{-g(\xi-t)}f(\xi)d\xi}{\int^{\wedge}_{Jt}(1.0 - e^{-g(\xi-t)})f(\xi)d\xi} \tag{18a}$$

$$\frac{x_{bat}}{dt} = R(X_{bat}(t)) \qquad\qquad te[\$t^-_{j+1}] \tag{18\&}$$

$$\frac{dl_{seg}}{dt} = e^{-g(t)}f(t)X_{bat}(t) \qquad\qquad je\ [1,N] \tag{18c}$$

with boundary conditions

$$X_{bat}(t^+_{j+1}) = X_{bat}(t^-_{j+1}) \tag{18d}$$
$$X_{bat}(t^+_0) = X_0 \tag{18e}$$

$$X_{max}(t^-_{j+1}) = X_{max}(t^+_{j+1}) \tag{18f}$$

$$X_{max}(t^-_N) = finite \tag{18g}$$

$$I_{seg}(t^+_{j+1}) = I_{seg}(t^-_{j+1}) \tag{18h}$$

$$I_{seg}(t^+_0) = 0 \tag{18i}$$

and the constraints

$$x_{exit} - f^{'TM''}\ n-0\ -e^{'g(t)}f(t)dt]X_{max}(0) + W\ W\ ) \tag{18j}$$

$$\int^{max}_{\mathbf{s}} f(t)dt = 1.0 \tag{18k}$$

$$\tau = \int^{t_{max}}_0 tf(t)dt \le t_{max}\int^{t_{max}}_0 Wdt = t_{max} \tag{18l}$$

Here x is the average residence time.

The boundary conditions (18d) through (18i) express the fact that the states (concentrations) are continuous at the knots. Equation (18k) ensures that the fraction of molecules with residence time less than infinity is one. The average residence time x may be fixed at some value or it may be allowed to find its optimum level, which can be shown to be bounded by t_{max}.

Since f(t) and h(t) are piecewise continuous they can be approximated on the j-th element by a linear combination of a set of M linearly independent basis functions (bases), (Wt). Thus,

$$f(t) \approx f_m = \sum_{j=1}^{N} \sum_{l=1}^{M} f_{ji} \phi_{ji}(t), \quad f_{ji} \geq 0 \tag{19}$$

$$h(t) \approx h_m = \sum_{j=1}^{N} \sum_{l=1}^{M} h_{ji} \phi_{ji}(t), \quad h_{ji} \wedge 0 \tag{2°}$$

The coefficients ty} and {h^ are chosen so as to maximize the objective function, J.

There is an infinite number of basis funtions that can be used. However, the choice of bases should take into account the inherent characteristics of the nonlinear program.

On an element [tj, tj^] we choose as bases

(i) $\phi_{j1} = 1.0/v$

(ii) $\langle j \rangle_{j2} = [e^{-\alpha t} - e^{-\alpha t_{j+1}}]/v$

(iii) $\phi_{j3} = [e^{\alpha t} - e^{-\alpha t_j}]/v$

(iv) $\langle \phi \rangle_{j4} = [e^{a}Vi - e^{at}]/v$

(v) $\phi_{j5} = [e^{-\alpha t_j} - e^{-\alpha t}]/v$,

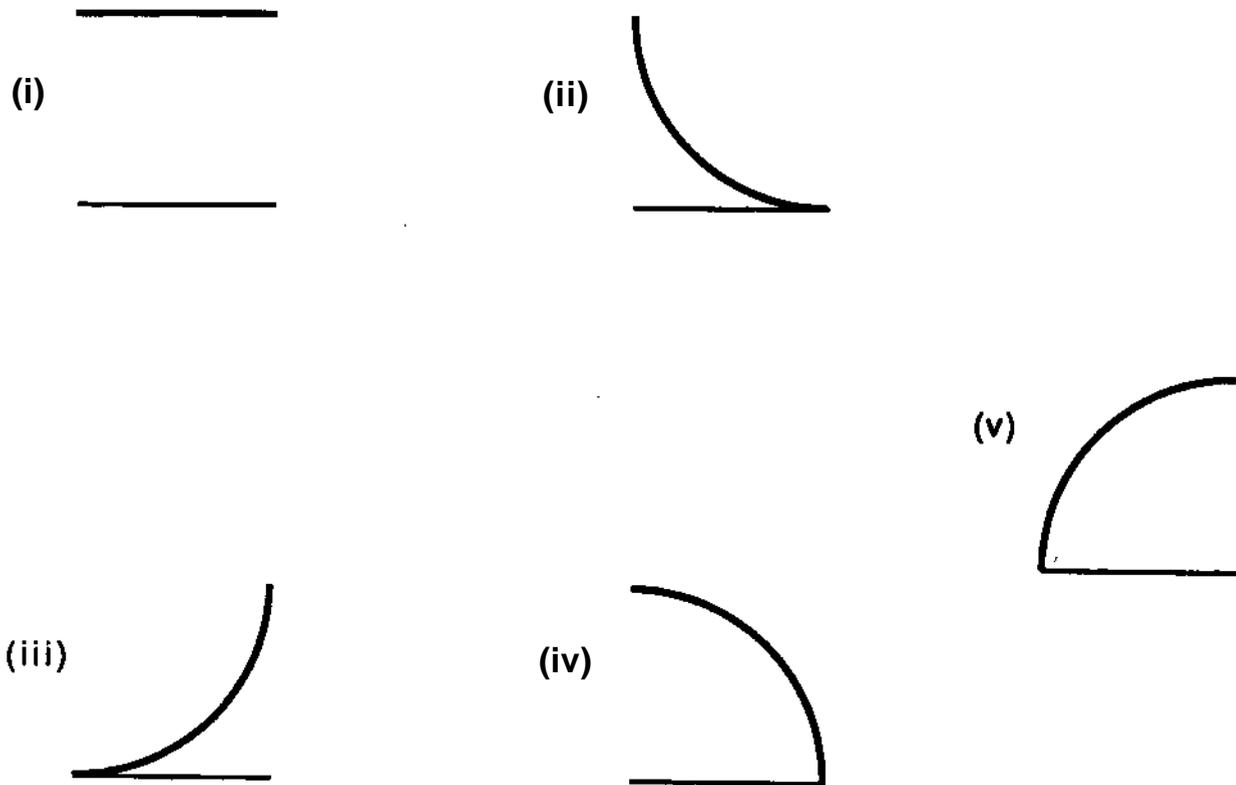where a and v are adjustable parameters. The bases are plotted in figure 2.

Figure 2: set of basis functions

The above choice of basis is motivated by the following:

(i) f and h are non-negative over [0, t_{max}]

(ii) the shapes of the bases mimic the shapes of the rtd's normally found in the chemical engineering literature

(iii) being exponentials the bases are easy to integrate or differentiate, and are generally easy to manipulate.

The bases {^(t)} can be defined globally (i.e. on the entire interval [0, t_{max}]) or locally (on an element [t_{jf} t_H] c [O.t^J).

Since residence time distributions can be discontinuous on [0, t_{max}] (eg: the PFR rtd) global interpolation is usually inappropriate. By definition a set {(^(t)} will have compact support (be non-zero) only on the j-th element.

f_m and h_m are used to denote the finite element approximation of f and h respectively. We illustrate the above concepts in Figure 3 below.
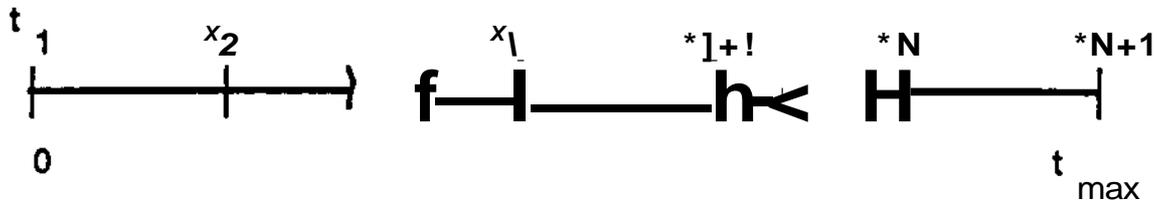
Figure 3: Partitioning of $[0, t_{max}]$ into N finite elements

In the above figure

$[tj, tj_{+1}]$sj-th element

*1-0. $t_{N+1}=t_{max}$, $t_{j+1}>tj$, $J \in [1,N]$

For global interpolation N = 1. N > 1 for finite element interpolation with N elements.

5. Method of Adjoints

There are a number of ways to solve the nonlinear optimal control problem (P1). Among these are methods based on Pontryagin's strong maximum principle and control vector parameterization coupled with direct search methods. The nonlinearity of the modified Rippin model (with respect to the control variables f and h) make the above methods impractical. The method which is described here makes use of adjoint variables, which are lagrange multipliers that are functions of time. An adjoint variable, which is a measure of the sensitivity of the objective function to a change in the decision variables, is assigned to each mass balance in the Rippin model. Let us define a hamiltonian H as

$$^H - *i (-^R(^X max> + ZmWK_m(t)) + X^R(X_{bat}) + xjx_{bat}(t)e-9m(t)f_m(t) \tag{22}$$

where

$$Z_m(t) = J \wedge P W t) -X_{bat}^\wedge t)]_m h (\text{t}e) 8 ^\wedge tf_m(4)d \tag{23}$$

$$KmW = J \wedge {}^{1}-0-e^{\wedge\wedge} m f t) * \tag{24}$$

We also define a Lagrangian L as

$$L = J(X_{exit}) + \sum_{j=1}^{AM} \mu_j^T [X_{max}(t_j^+) - X_{max}(t_j^-)] + \sum_{j=1}^{AM} \eta_j^T [I_{seg}(t_j^+) - I_{seg}(t_j^-)] +$$

$$\sum_{j=1}^{AM} \quad \quad \quad ^{AM} p_{r...} \quad ^T \cdot \quad \quad ^T \cdot \quad \quad ^T \cdot$$

where μ, TJ, v, X_A, XQ and XQ are adjoint variables. It is easy to verify that when the mass balances and boundary conditions are satisfied L = J, the objective function. A change in the states results in a change in L as follows:

dL ≈

$$(dJ/dX_{exit})^T dX_{exit} + \sum_{7=1}^{AM} \mu_j^T [dX_{max}(t_j^+) - dX_{max}(t_j^-)] + \sum_{j=1}^{AM} \eta_j^T [dI_{seg}(t_j^+) - dI_{seg}(t_j^-)] +$$

$$\sum_{j=1}^{N-1} v[[dX_{bat}(tf) - dX^O_j^-)] + g \, d(J^[H - ^ X.,,3, - XI ^ - Aj I_{seg}]dt) \quad \quad (26)$$

This equation is the basis for defining the functional form of the adjoints. More precisely the integral term determines the ODE's that λ_A, X_c and $^\lambda_D$ satisfy. The remaining terms determine the boundary conditions on the system of adjoint equations. Although it is straight forward the derivation is long and somewhat unwieldy, and will therefore not be presented in this paper. We include the knots as decision variables since they are allowed to move relative to each other. The following is a summary of the adjoint system:

$$\frac{d\lambda_A}{dt} = -\frac{\partial H}{\partial X_{max}} \quad \quad <27a>$$

$$\frac{d\lambda_c}{dt} = -\frac{\partial H}{\partial X_{bat}} \quad \quad (27b)$$

$$\frac{d\lambda_D}{dt} = -\frac{\partial H}{\partial I_{seg}} = 0 \tag{27C}$$

with boundary conditions

$$\lambda_A(t_j^+) = \lambda_A(t_j^-) \ J^e t^{2\cdot N}| \tag{27d}$$

$$\lambda_A(t_0^+) = -\left(f^{\wedge}{}'' \left[1 - 0 - e^{\wedge} J O dr_t - JL\right.\right._{/^{a*}exit} \tag{27e}$$

$$\lambda_C(t_j^-) = \lambda_C(t_j^+) \quad j \in [2, N] \tag{27f}$$

$$\lambda_C(t_{N+1}^-) = 0 \tag{27g}$$

$$\lambda_D(t_j^-) = \lambda_D(t_j^+) \tag{27h}$$

$$\lambda_D(t_{N+1}^-) = \wedge_{a\,A\ exit} - \tag{27i}$$

As we shall see later, XQ does not appear in the gradient of the lagrangian with respect to the decisions {f^, {h$, and {ty. As a result one need not solve (27b). Equation (27a) and (27f) imply

$$\lambda_D(t) = \frac{dJ}{dX_{exit}}$$

With the above definitions and analysis the gradient of the lagrangian with respect to the decisions {f;;}, {ty and {ty take the form

$$|i := f^{Vi} C(t)(dJ/dX_{exit})^T(X_{max}(0)[1.0 - e^{\ ''\ \wedge} + X_{bat}(t)e^{-g_m(t)})dt +$$

$$\int_{t_i}^{t_{i+1}} \phi \|\Gamma(t) \int_0^t \left(\lambda_A^T(\xi)[X_{max}(\xi) - X_{bat}(t-\xi)]h_m(t-\xi)e^{-g_m(t-\xi)}/K_m(\xi) - \right.$$

$$\lambda_A^T(\xi)Z_m(\xi)[1.0 - e^{-g_m(H)}]/K_m^2 \ ©y\&l \tag{28a}$$

$$\frac{\partial L}{\partial h_{ij}} = J_{j}^{t_{j+1}} \overline{\lambda_A^T}(t) \left(J_0^{V_t} K c \, JJ/dX_{exit})^T (X_{max}(0) - X_{bat}(s_+,t)) e^{-9m(^\wedge)} f_m(s_+,t) + \right.$$

$$\lambda_A^T(s)[X_{max}(s) - X_{bat}(t)] e^{-9m(t)} f_m(s_+,t)/K_m(s) \quad - \quad J_{s+t} V_{is}) Z_m(s)/K_m^2(s) +$$

$$\lambda_A^T(s)[X_{max}(s) - X_{bat}(\xi-s)] h_m(\xi-s)) e^{-9_m(\xi-s)} f_m(\xi)/K_m(s) d\xi \, ds \Bigg) dt \qquad (28b)$$

$$\S£ - H(lp - Hflf) + (dJ/dX_{ex}/^\wedge h_m(t-) - hJt + M_j^\wedge t X^\wedge O) - X_{bat}(t)] e^{-9n,(t)} f_m(t) dt +$$

$$Xmax(0)[(1-0-e^{-9m(9)} f_m(^\wedge) - (1.0 - e^{-9_m(t_j^+)} f_m(t_j^+)] \Bigg)$$

$$- JJAJCOZ^\wedge Q \Big/ [1.0 - e^{-Sm^\wedge}] f_m(tp - [1.0 - e^{-9m(tp}] f_m(tf) VK_m^2(t) dt$$

$$- \left(\int_0^{t_{max}-t_j} \int_{t+t_j}^{t_{max}} \lambda_A^T(t) Z_m(t) e^{-9_m(\xi-t)} f_m(\xi)/K_m^2(t) d\xi dt \right) [h_m(t^-) - h_m(t+)]$$

$$+ \int_0^{t_j} \lambda_A^T(t) [X_{max}(t) - X_{bat}(t_j-t)][h_m(t^--t) e^{-9_m(t_j^--t)} f_m(t_j^-) - h_m(t^+-t) e^{-9_m(t_j^+-t)} f_m(t_j^+)]/K_m(t) dt -$$

$$\frac{f}{V}{}^{\circ} \quad Jt+|_j \quad \dots \quad)[X_{max}(t) - X_{bat}(\xi-t)] h_m(\xi-t) e^{-9_m(\xi-t)} f_m(\xi)/K_m(t) d\xi dt \Bigg) [h_m(t^-) - h_m(t+)] \qquad (28c)$$

$$\frac{\partial L}{\partial t_{N+1}^-} = H(t_{N+1}^-) + (dJ/dX_{exit})^T X_{max}(0) e^{-9_m(t_{N+1})} f_m(t_{N+1}^-) -$$

$$J_0^{t_{max}} * n (W - 0 - ^\wedge \ ^\wedge \ 1 >] f_m(t N_{+1}^-)/K_m^2(t) dt$$

$$+ \int_0^{t_{max}} \lambda_A^T [X_{max}(t) - X_{bat}(t_{N+1}-t)] h_m(t_{N+1}^- - t) e^{-9_m(N+t^-)} f_m(t_{N+1}^-)/K_m(t) dt \qquad (28d)$$

for j = 1....(N+1) and i = 1....M

6. Strategy for Solving Optimal Control Problem

The target problem has been formulated as a nonlinear program (NLP). Different methods exist for its solution. For the NLP we employ a gradient based algorithm that uses successive quadratic programing (SQP) to update the decision variables. The SQP step employs the gradients that were presented in the previous section. The relevant steps in the algorithm are as follows:

Let Y be the vector of decision variables [{f^, {hjj}, {tj}] Then in terms of Y the NLP takes the simpler form

$$MAXJ = J(Y) \tag{a}$$

subject to:

$$Y_{min} \le Y \le Y_{max} \tag{b}$$

$$S_{min} \le S(Y) \le S_{max} \tag{c}$$

The nonlinear constraint set (c) contains the two integral equality constraints on f (equations 18k and 18l) and may involve some or all the decision variables. For example we may require the RTD and micro mixing functions to be continous at the knots and that $t_j < t_{j+1}$. Note also that if the knots, {tj}, are fixed then the integral constraints on f are linear in {y. Finally we propose the following strategy for solving the NLP:

(i) Provide an initial guess for the decision variables, Y

(ii) Solve the nonlinear state equations (18a) through (18c) together with the associated boundary conditions using an ODE code such as LSODI.

(iii) Evaluate the objective function J.

(iv) Using state variables from (ii) solve the linear adjoint equations (27a) through (27c).

(v) Use the set of equations (28) to calculate the derivatives of the objective function with respect to the above decision variables.

(vi) Calculate a search direction for the decision variables by setting up and solving the following quadratic program at Y_k (Y at k-th iteration):

$$\underset{d_k}{MAX}\ Q = V^T J(Y_k)d_k + 0.5djB_kd_k$$

subject to:

$$^Y min - {}^Yk * {}^dk \wedge {}^Ymax \; " \; {}^Yk$$

$$^Smin * {}^Ak{}^Yk * {}^Ak{}^dk * {}^Smax \; " \; {}^Ak{}^Yk$$

Here $A_k = V^T S(Y_k)$. B_k is a positive definite hessian matrix constructed by a quasi-Newton updating formula and serves as an approximation to $V^2 J(Y_k)$. With the search direction d_k update Y by $Y_k^\wedge{}_1 = Y_k + a \wedge$, where $c^\wedge$ is a stepsize selected so that a sufficient increase in J is found at Y_{k+1}. This Successive Quadratic Programming (SQP) algorithm is given in Han (1977) and Powell (1977). A complete description of an updated algorithm is given in Biegler and Cuthrell (1985). The above QP is solved using a code written by Gill et al.(1978)

(vii) If the Kuhn-Tucker conditions for a stationary point for the NLP are satisfied to a specified tolerance 8 STOP. Else go to (ii).


It should be noted that depending on the general nature of the objective function, J, there could be multiple local optima for the NLP. As a result to increase the likelihood of finding a global optimum it might be useful to restart the algorithm from different initial points. However, unless the objective function exhibits special features (such as convexity in the decision variables), there is no guarantee that a global optimum can be found.

7. Test Examples

Three examples have been chosen to illustrate the approach to reactor synthesis (involving isothermal homogeneous fluid phase reactions) that has been presented in this paper.

Example A:

This is an isothermal Van de Vusse reaction involving four species for which the objective is the maximization of the yield of the intermediate species B. The reaction diagram is as follows:

$$A \xrightarrow{\quad K_1 \quad} P\text{-}\ B \xrightarrow{\quad K_2 \quad} C$$

$$A \xrightarrow{K_3} D$$
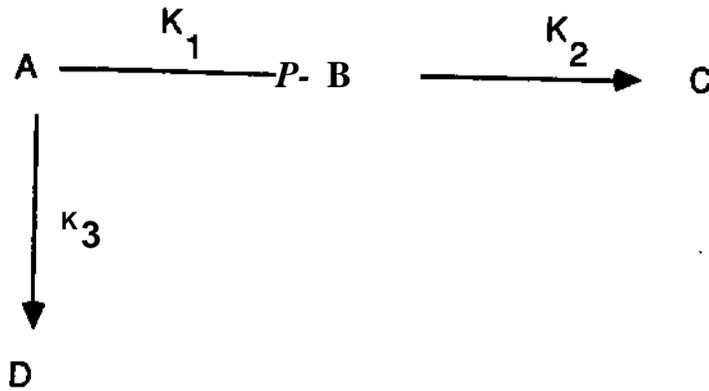
where $k_1 = 10s^{-1}$ (first order), $k_2 = 1s^{-1}$ (first order), $k_3 = 1$ L/(gmol s) (second order), molar flow rate of reactant A = 58 gmol/s, $C_{A0} = 0.58$ gmol/L.

In terms of normalized concentrations, X, the reaction rate vector is given by R(X) «[fc_A, *B *C *D]^T - I-^{fc}1 " °-5 WI $\hat{k_i}^X A - \hat{k_2}^X B$. *$_2^X$B' 0.5 $\hat{k_3}X_A^2$]^T

where $\hat{k_1} = k_{1f}$ $\hat{k_2} = kg$, $\hat{k_3}$ « $C_{A0}k_3$.

This example was solved by Chitra et al (1981) and later by Achenie and Biegler (1986), who reported an optimal reactor network of a single plug flow reactor. The target approach resulted in an optimal B-yield of 0.4426 gmol/L (versus 0.4368 Achenie et al, 0.4362 Chitra et al) with an optimal residence time of 0.2932s (versus 0.2965s, Achenie et al).

Other pertinent data are a = 2, number of elements = 15, $t_{max} = 3.0$ To reduce the number of decision variables only bases Φ_{jy} ^ and Φ_{j3} were used on each element.

As shown in figure A.2 the RTD looks very much like a delta dirac function (the PFR RTD). In this case, any values of h_m will yield the same objective and the h_m profile is nonunique and not important here. In figure A.1 where the RTD and micro mixing function have been plotted together, h_m is small wherever f_m is non-zero. Here the micromixing function suggests that the optimum mixing pattern is close to segregated flow.

VAN DE VUSSE : (0.58)
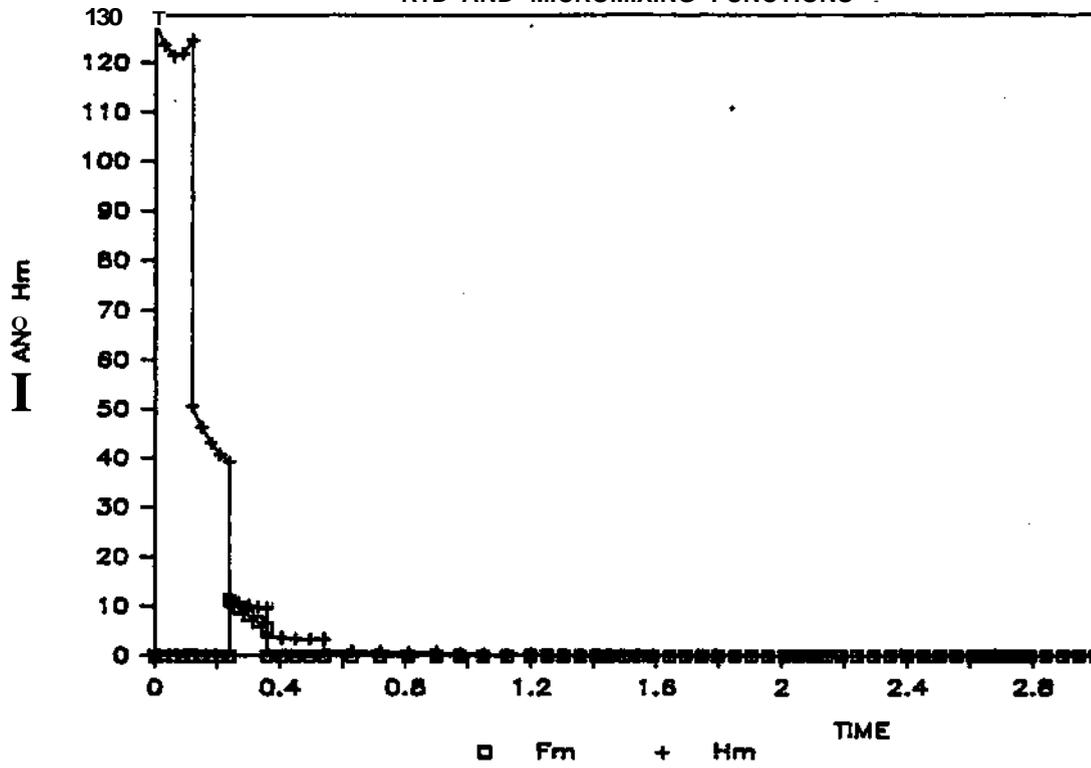
RTD AND MICROMIXING FUNCTIONS

Figure A.1: RTD and h_m for example A
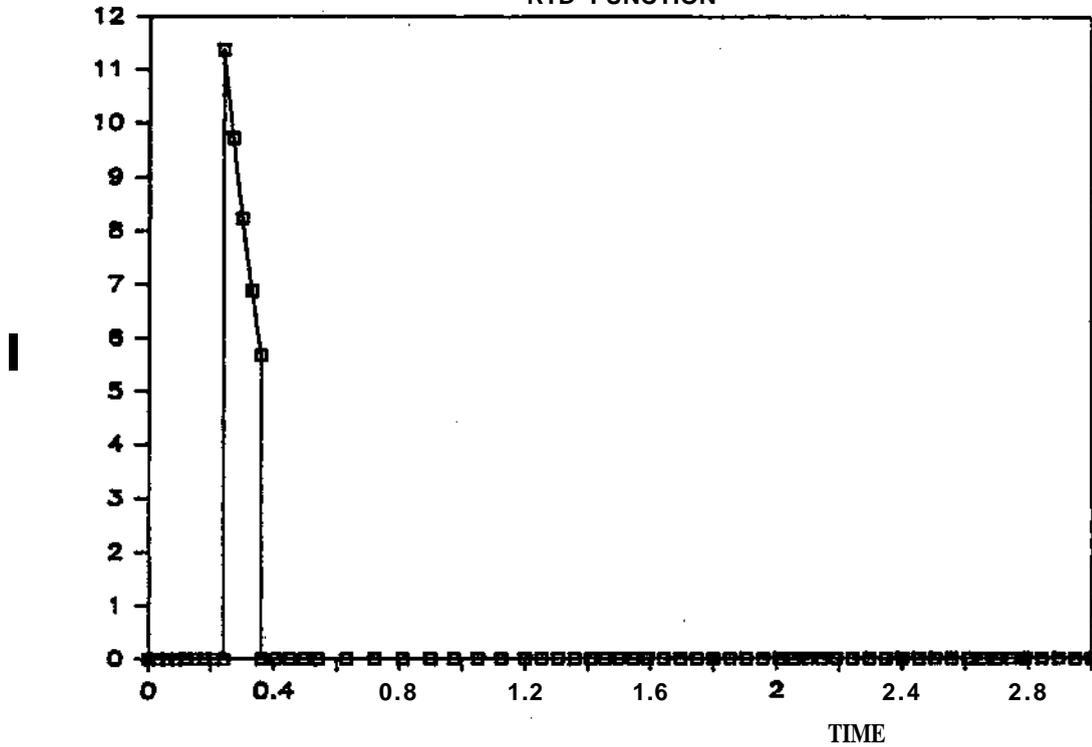
VAN DE VUSSE : (0.58)

RTD FUNCTION

Figure A.2: RTD for example A

Example B:

The reaction diagram is the same as the one in example A except that here $C_{A0} = 5.8gmol/L$ Chitra et al, Achenie at al (1986b) reported an optimal reactor network of one CSTR followed by a PFR. The optimal RTD for such a reactor train is an exponential RTD with a lag as in figure 4.
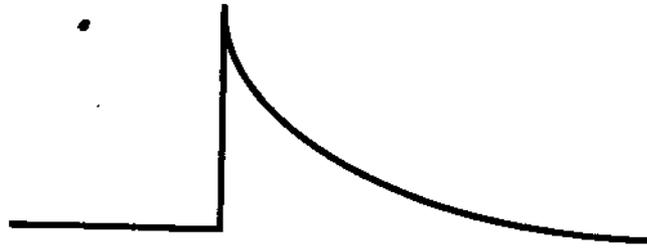
Figure 4: RTD of Optimal Network (Achenie and Biegler)

The target approach resulted in an optimal B-yield of 3.7271 gmol/L (versus 3.6806 with Achenie et al and 3.6772 with Chitra et al) and an optimal residence time of 0.2702 seconds (0.2381 with Achenie et al and 0.2802 with Chitra et al). The optimal RTD is similar to that in A, except that here the RTD is more spread out and shorter. The micro mixing profile is strongly influenced by the chosen basis functions. However, the general trend is that of a large h_m on the first half of the interval and tailing off towards the end of the interval (figures B.1 and B.2). The large h_m at the beginning points to a CSTR followed by a PFR, as opposed to a PFR followed by a CSTR (which network has the same RTD)

Figure B.1: RTD for example B

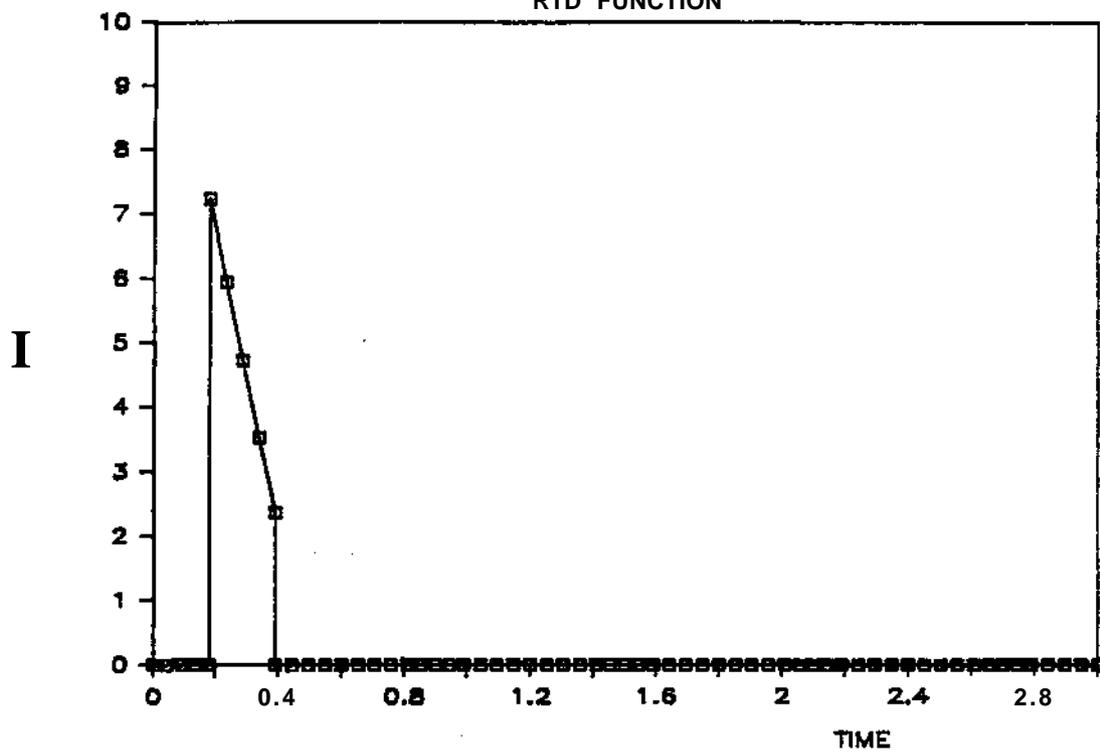## VAN DE VUSSE : (5.8)

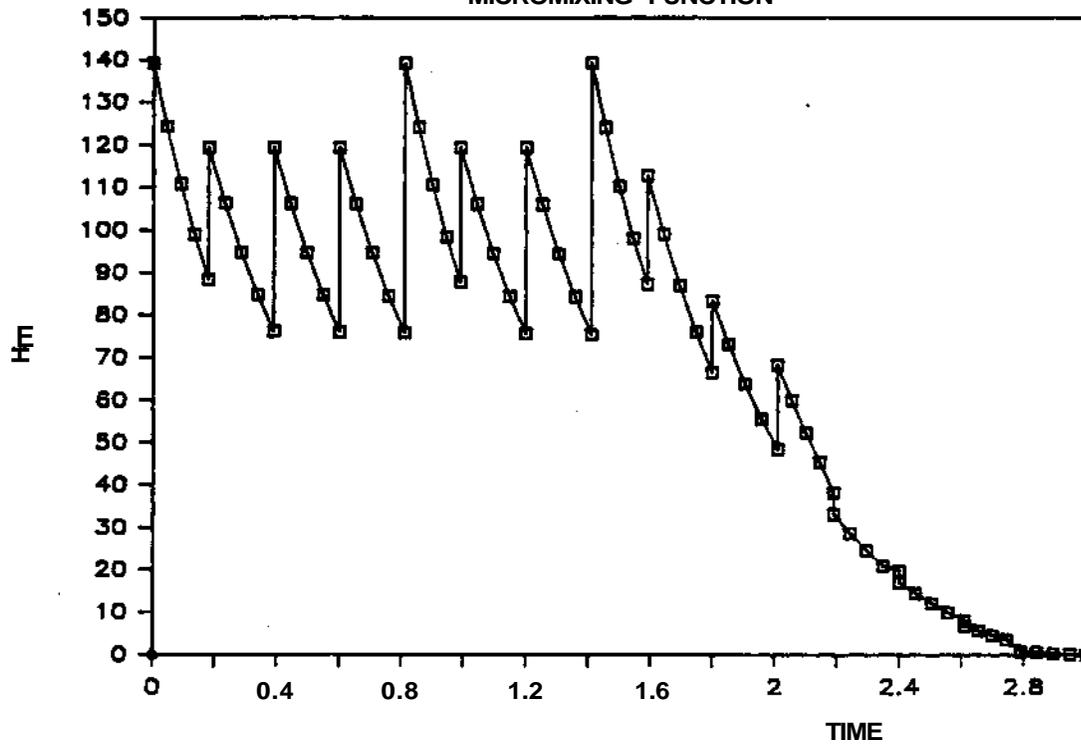**MICROMIXING FUNCTION**

Figure B.2: h_m for example B

Example C:

This example, the "plant problem", is a modified form of the Williams and Otto problem (as stated in Ray and Szekely, 1973, and in Di Bella and Stevens, 1965).
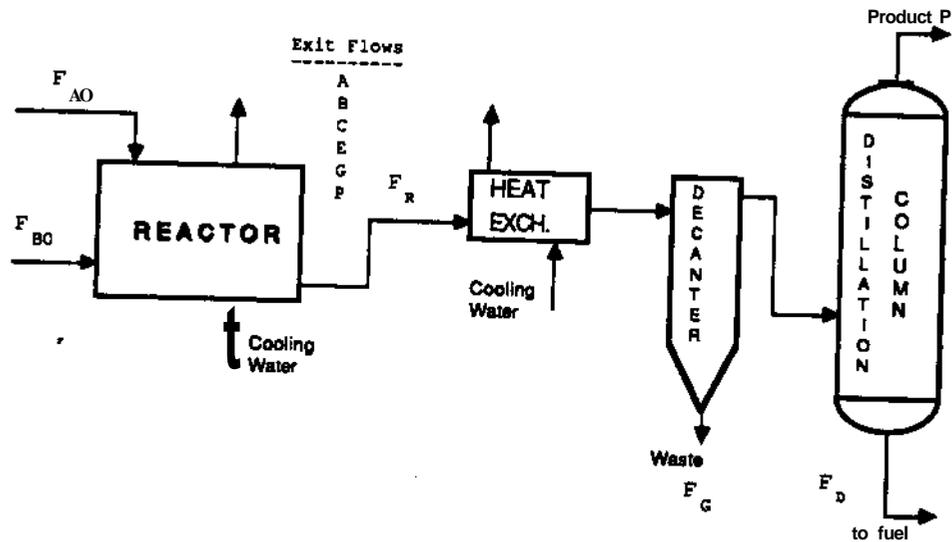
Figure 5: Flow Diagram of the Chemical Plant

This plant produces a chemical P from reactants A and B. There is a decanter that separates the heavy oily by-product G from the reaction mix before it enters a distillation column where P is separated from other components. Components C and E are intermediates that can be used as plant fuel. The reaction diagram is as follows:

A + B --*1————> C,   2nd order

C + B —fc»————> P + E,   2nd order

P + C ™ka————> G,   2nd order
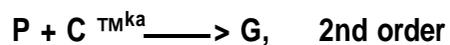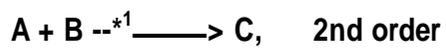
Unlike the original problem here we assume:

(i) that the reaction is isothermal,

(ii) the bottoms product is not recycled to the reactor, and

(iii) 50 million pounds of chemical P are to be produced in one year.

At the temperature of the reaction, 120°, the rate constants are lt) = 6.1074 hr wt fraction, kg = 15.0034 hr wt fraction and kg = 9.9851 hr wt fraction.

The objective function for this plant is taken to be the annual rate of retum on the investment, which can be expressed as (assuming total operating hours per year of 8400 hours).

(i) Sales Volume: $[0.3F_P + 0.0068F_D]$/hr

(ii) Raw Material Cost: $(0.02F_{A0} + 0.03F_{B0})$/hr

(iii) Waste Treatment Cost: $0.01 \hat{F}_G$/hr

(iv) Utilities Cost: $2.22F_R$/yr

(v) Sales, Administration and Research Expenses: 12.4% of sales

(vi) Plant Fixed Charge: 10% of the Plant Investment/yr

(vii) Plant Investment: 600Vp

where

$p = 50$ lb/ft^3, $V = Ox = 120.0x$, $T =$ residence time

$Fp = \hat{F}_p - 0.1\hat{F}_E$

$F_D = F_{A0} + F_{B0} - \hat{F}_G - F_P$

$F_R = \hat{F}_A + \hat{F}_B + \hat{F}_C + \hat{F}_E + \hat{F}_G + \hat{F}_P$

If there is no recycle then $F_R = F_{A0} + F_{B0}$.

$\hat{F}_j$ is the molar flow rate of species j at the reactor exit. Using the above information the objective function I(F) becomes

Max I(F) - $100[8400(0.3F_P + 0.0068F_D - 0.02F_A - 0.03F_B - 0.01 F_G) - 0.124 \times 8400 \times (0.3F_p + F_D) - 2.22F_R - 60Vp]/600Vp$

Now express all flows in terms of F_A, F_B, F_c and F_p and normalize by $F^0_{A0} =$ 1.3546x10^4 lb/hr, then we have

$J(X_{exit}) = -7.4 \times 10 - \% - 7.4 \times 10"\% - 7.4 \times 10"\% + 7.1842)\hat{} - 0.7265)\hat{\%} - 0.4541)\hat{} - 0.3932\hat{X}_{A0} - 0.6732\hat{X}_{B0}]/Q\tau - 10.0/F^0_{A0}$

as the new objective function. The reaction rate vector is given by $R(X) = [- \hat{k}_1 X_A X_B,$ $-(\hat{k}_1 X_A + \hat{k}_2 X_C)X_B,\ 2\hat{k}_1 X_A X_B - 2\hat{k}_2 X_B X_C - \hat{k}_3 X_P X_C,\ \hat{k}_2 X_B X_C - 0.5\hat{k}_3 X_P X_C,\ 2\hat{k}_2 X_B X_C,$ $1.5\hat{k}_3 X_P X_C]^T = [\hat{X}_A, \hat{X}g, \hat{X}_Q, \hat{X}_P, \hat{X}g, \hat{X}_Q]^T$

where $\hat{k}j = kj F^0_{A0}/F_R$ and F^0_{A0} is the mass flow rate of species A at the 0-th iteration. The target approach resulted in an annual rate of return on investement of 219.9% and an optimal residence time of 0.5004 hours. In addition the optimal values for F_{A0} and F_{B0} were 13546 lb/hr and 33776 lb/hr respectively. The general shape of the RTD is more spread out than in the two previous examples. h_m is zero almost everywhere, suggesting that segregated flow is favorable to the maximization of the objective function. As a result of the spread, the the maximum value the RTD takes is a lot smaller than in the previous examples.

PUNT PROBLEM

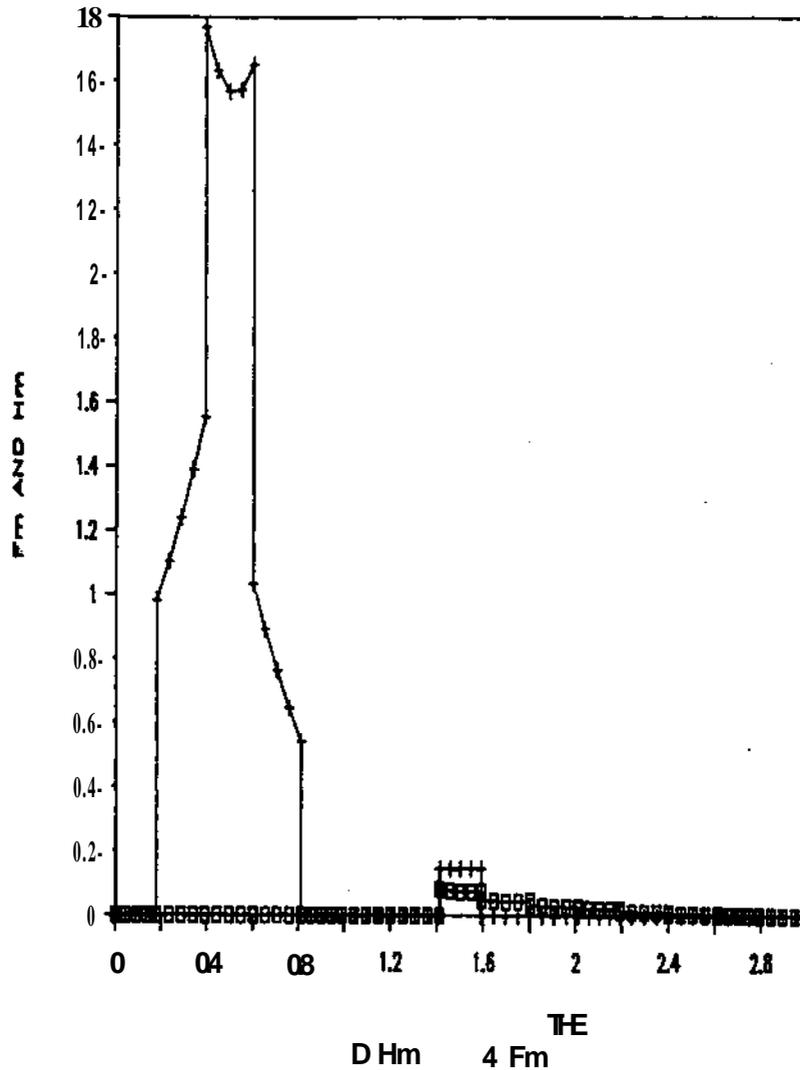### RID AND MKROMMNC FUNCTION

Figure C: RTO and h_m for example C

8. Discussion

In examples A and B the target problem yielded a slightly better objective function than the one based on a superstructure of reators. In this respect the target approach is quite successful. However, some characteristics of the target approach need to be pointed out.

(i) The target approach assumes that the RTD is species independent. It is therefore reasonable to expect that this approach may not give an upper bound for reactor networks that allow component splits. To get around this limitation the RTD will have to be made species dependent.

(ii) Besides being limited to homogeneous reaction mixtures of low viscosity, the modified Rippin model (on which the target approach is based) is restricted to isothermal and adiabatic systems. However, the model can be modified to allow non-isothermal reactions.

(iii) The local profile (i.e. on an element) of h_m and f_m are influenced by the shapes of the basis functions used. By scaling the basis functions such that no one function has undue influence alleviates that dependence. In the above examples a, the characteristic time for the bases, was set constant on each element. It is expected that by using the alternate set of bases $\{1, e'V\}$ on the k-th element and by considering o^ to be a decision variable, smoother profiles will result. However, such a modification will make the integral equality constraints on f nonlinear.

Also to avoid nonlinear constraints on f_m, the knot positions were chosen a priori (uniformly placed). The next step is to allow knot movement as part of the optimization problem. In addition the problems raised above will be addressed as part of continuing improvement on the algorithm.

References

[Achenie 86a]   Achenie, L and Biegler, L
Algorithmic Synthesis of Chemical Reactor Networks Using
     Mathematical Programming.
I & EC Fundamentals 25:621-627,1986.

[Achenie 86b]   Achenie, L and Biegler, L
An Algorithmic Method For Reactor Network Synthesis.
Engineering Design Research Center, Carnegie Mellon
     EDRC-06-20-86,1986.

[Aris61]   Aris, R.
The Optimal Design of Chemical Reactors.
Academic Press, 1961.

[Biegler 85]   Biegler, LT. and Cuthrell, J.E.
Improved Infeasible Path Optimization For Sequential Modular
     Simulators - II: The Optimization Algorithm.
Computers and Chemical Engineering 9:257-267,1985.

[Chauhan 72]   Chauhan, S.P., Bell, J.P. and Adler, R.J.
On Optimum Mixing in Continuous Homogeneous Reactors.
Chemical Engineering Science 27:585-591,1972.

[Chitra81]   Chitra, S.P. and Govind, Rakesh.
Yield Optimization for Complex Reactor Systems.
Chemical Engineering Science 36:1219-1225,1981.

[Di Bella 65]   Di Bella, C.W. and Stevens, W.F.
Process Optimization by Nonlinear Programming.
I & EC Process Design and Development 4:16-20,1965.

[Gill 78]   Gill, P.E., and Murray, W.
Numerically Stable Methods for Quadratic Programming.
Mathematical Programming 14:349-372,1978.

[Glasser 86]   Glasser, D., Crowe, C M. and Jackson, R.
Zwietering's Maximum-Mixed Reactor Model and the Existence of
     Multiple Steady States.
Chem. Eng. Commun. 40:41-48,1986.

[Han 77]   Han, S.P.
A Globally Convergent Method for Nonlinear Programming.
Journal of Optimization Theory and Applications 22(3)297-309,1977.

[Jackson 86]   Jackson, R. and Glasser, D.
A general Mixing Model for Steady Flow Chemical Reactors.
Chem. Eng. Commun. 42:17-35,1986.

[Levenspiel 72]   Levenspiel, O.
Chemical Reaction Engineering.
John Wiley and Sons, Inc., 1972.

[Nestoridis 86]   Nestoridis, V., Andreou, I., and Vayenas, C.G.
Optimal Residence Time policy for Product Yield Maximization in
Chemical Reactors.
Journal of Optimization Theory and Applications 49(2)271-287,1986.

[Ng 65]   Ng, D.Y.C. and Rippin, D.W.T.
The Effect of Incomplete Mixing on Conversion in Homogeneous
Reactions.
In *Third Symposium on Chemical Reaction Engineering.* Pergamon
Press, Oxford, 1965.

[Powell 77]   Powell, M.J.D.
A Fast Algorithm for Nonlineaiiy Constrained Optimization
Calculations.
In *Dundee Conference on Numerical Analysis.* , Dundee, Scotland,
1977.

[Rao 73]   Rao, D.P. and Edwards, LL.
Mixing Effects in Stirred Tank Reactors: a Comparison of Models.
Chemical Engineering Science 28:1179-1192,1973.

[Ray 73]   *r*   Harmon, R.W. and Szekely, J.
*Process Optimization with Applications in Metallurgy and Chemical
Engineering.*
John Wiley and Sons, Inc., 1973.

[Villermaux 69]   Villermaux, J. and Zoulalian, A.
Etat de Melange du fluide dans un Reacteur Continu. A Propos d'un
Modele de Weinstein et Adler.
Chemical Engineering Science 24:1513-1517,1969.

[Zwietering 59]   Zwietering, N.
The Degree of Mixing in Continuous Flow Systems.
Chemical Engineering Science 11:1-15,1959.