A Reduced Hessian Strategy For Sensitivity Analysis Of Optimal Flowsheets

by

N. Ganesh, L T. Biegler

EDRC-06-19-86 A

September 1986

A REDUCED HESSIAN STRATEGY FOR SENSITIVITY ANALYSIS OF OPTIMAL FLOWSHEETS

N. Ganesh and L. T. Biegler
Department of Chemical Engineering
Carnegie-Mellon University
Pittsburgh, PA 15213

To be presented at Annual AIChE Meeting

November, 1986
Miami, FL

Session on Simulation in Chemical Engineering, Paper a

ABSTRACT

An efficient and rigorous strategy is presented for evaluating the first order sensitivity of the optimal solution to changes in process parameters *or* process models. An algorithm that constructs a reduced Hessian in the null space of the equality constraints is used to solve the sensitivity equations; the resulting effort to solve these equations depends only on the space of the *decision* (independent) variables. Consequently, large computational savings can be realized because the solution procedure eliminates the need for obtaining second partial derivatives with respect to *tear* (dependent) variables explicitly. The method is applied to several flowsheeting examples in order to determine efficiently the sensitivity of the optimal solution to parametric and physical property model changes.

INTRODUCTION

Process simulation has become a widely accepted technique for carrying out design and cost estimation studies for chemical process flowsheets. Large scale process simulators are characterized by the presence of numerous complex mathematical models that constitute their functional core. Various process related properties such as vapor-liquid equilibrium constants, kinetic rates etc. are computed based on the parameters residing within these models. In reality, several model parameters such as product prices, or kinetic parameters may be uncertain or may vary over a known range. Also, competing models may exist for physical properties or unit operations with no single model being accurate over the entire range of interest. The predicted results from a simulation may therefore be subject to uncertainty, due to imprecise modeling of the process. Also, trends in changing the optimal solution are often of interest with variations in fixed parameters.

Recent developments in the area of process optimization have provided us with the ability to implement simultaneous simulation and optimization techniques using large-scale process simulators (Berna *et al.*, 1980; Jirapongphan *et al.*, 1980; Biegler & Hughes, 1982). However, because of the uncertainties and possible variations involved in process parameters or models, an optimal solution gained from a deterministic optimization problem may not, by itself, be entirely useful. Therefore, as a first step, postoptimality analysis becomes necessary to ascertain quantitatively how parametric variations and model selection affect the optimal results obtained under nominal conditions.

In the past, the parametric sensitivity problem has been addressed for simulation by several researchers. Atherton (Atherton *et al.*, 1975) proposed a statistical approach to determine sensitivity coefficients to measure the influence of uncertainties in model parameters on the solutions obtained from incorporating a particular model. In the context of flowsheet simulation, Volin and Ostrovskiy developed an approach based on setting up

and solving an adjoint flowsheet system to the nominal problem (Volin and Ostrovskiy, 1981). A more direct approach to determine the sensitivities of flowsheet variables to parametric variations utilizing the block Jacobians of unit modules was presented recently (Gallagher and Kramer, 1984). Their approach involves constructing the parametric derivative matrix by forward difference perturbations in a manner that resembles the chainruling procedure employed in other simulation and optimization studies (Stadtherr and Chen, 1983; Shivaram and Biegler, 1983).

In this paper we present an efficient and rigorous strategy for evaluating the first order sensitivity of the *optimal* solution to changes in process parameters or models. As a first step we partition the process variables into two sets, the independent (decision) variables, x and the dependent (tear) variables, *y*. When the optimum of the nominal problem satisfies the second order sufficiency conditions (local optimality), the sensitivity results for a parametric nonlinear programming problem are well known if the gradients of the active constraints are linearly independent and strict complementary slackness holds (Fiacco and McCormick, 1968; Fiacco, 1976). In Fiacco's formulation, the application of sensitivity analysis to determine the first order sensitivity of the optimal solution requires the Hessian of the Lagrange function in the combined space of the independent and dependent variables. The computational effort involved in building the Hessian can become prohibitively large even for a moderately sized flowsheet with a combined <x + </> dimensionality in the range 20-50. In this paper, we initially develop a *reduced* Hessian decomposition algorithm to solve the sensitivity equations for parametric variations. In this approach, flowsheet perturbations, most frequently employed in generating second derivative information for constructing the Hessian, need be performed only in the space of *decision* variables. Consequently, significant computational savings are realized in the evaluation of the sensitivity of the optimum to parametric variations. The procedure yields sensitivity <

infonnation on all optimal variables but the equality constraint multipliers.

We then develop the theory for analyzing the sensitivity of the optimum to *model* variations. The problem is formulated as computation of a Newton step in the space of the *new* model and the resulting set of linear equations is solved to determine the changes in the variables. In parallel with the parametric sensitivity problem, the reduced Hessian strategy can be applied for solving the linear system for model sensitivity. However, unlike the parametric case where the active set is retained under first order variations in the parameters, the model sensitivity problem must take into account changes in the active set due to model changes. To this end, we develop a novel procedure that may rigorously require solving a Mixed Integer Nonlinear Programming (MINLP) problem, although simpler approaches can be used and are demonstrated.

We illustrate the reduced Hessian procedure and the model sensitivity approach with simple analytical examples. In addition, several flowsheet optimization problems are considered in order to demonstrate the effectiveness of these approaches.

PARAMETRIC SENSITIVITY ANALYSIS

Consider the parameter based flowsheet optimization problem:

$$\begin{aligned} \text{P(p):} \quad & \underset{2}{Min} \quad \phi(z,p^{\circ}) \\ & \text{s.t} \quad giz.p^{\circ}) \ * \ 0 \\ & \qquad h(z,p^{\circ}) = 0 \end{aligned} \tag{1}$$

where

p = input parameter for base case solution

2 = process variable, $z = \begin{Bmatrix} x \\ y \end{Bmatrix}$

<fr - objective function

g - design inequality constraint

h - equality (tear) constraint

Decision variables, x, are usually adjusted by the designer and can include equipment sizes or temperatures and pressures in the flowsheet Dependent variables, y_9 can be computed from equality constraints once the decision variables are specified. These include the flowrates, pressure and enthalpy of the recycle streams in the flowsheet as well as any other variable that is specified explicitly by an equation. For convenience these are also referred to as tear variables.

At the optimal solution, when the KKT conditions are satisfied we have the following relations:

$$V\wedge(z^{\circ},p^{\circ}) \bullet a^{\circ}V\wedge(z^{o},p^{o}) \ll\bullet v^{\circ}V/Kz^{o}.p^{\circ}) = 0$$

$$u^{o}giz^{\circ}.p^{9}) = 0;\ a^{\circ} * 0;\ \leq?(r^{o}.p^{o}) \pounds 0$$

$$Mr^{\circ}.p^{\circ}) = 0 \qquad (2)$$

where

z° - base case optimal solution, $(x^{\circ},(/^{\circ}]$

a°,v° - KKT multipliers at base case optimum

This result may be interpreted to mean that for the supplied input parameters. p°. the variable vector, z°, is a local minimum of $P(p^{\circ})$ with the corresponding KKT multipliers u°,v°. In the context of flowsheet optimization the input parameter vector. p°. can include a subset of internal process parameters (e.g. kinetic rate constant terms) and externally supplied parameters (eg. feed flowrates) that are utilized for simulating the constituent process modules.

The parametric sensitivity problem addressed in this paper is to obtain first order changes in the optimal process variables and the KKT multipliers with respect to the parameters, p. The development of the mathematical formulation for sensitivity analysis is based on the classical Implicit Function Theorem (Fiacco,1976; Luenberger,1973). We start with the assumption that at the local minimum, z°, the following conditions are satisfied:

1. the functions defining $P(p^\circ)$ are continuously differentiable in (z, p) in a neighborhood of (z°,p°)

2. the constraint gradients are linearly independent at z° and, consequently, Strict Complementary Slackness holds for $P(p^\circ)$ at z° with unique KKT multipliers, u° and v° and

3. the Second Order Sufficiency conditions are met (cf. Appendix I)

From the KKT conditions at the optimum, z°, we have:

$$\begin{aligned} \nabla_z L(z^\circ,p^\circ) &= 0 \\ g_A(z^\circ,p^\circ) &= 0 \\ h(z^\circ,p^\circ) &= 0 \end{aligned} \tag{3}$$

where

L - Lagrange function

g_A - active inequality constraint

In order to satisfy these conditions for a perturbation, Δp, in the parameter p, about p°, we can find the first order corrections by noting that:

$$\begin{aligned} d\,(\nabla_z L(z^\circ,p^\circ)) &= \nabla_{zz} L^\circ\, dz + \nabla_{zp} L^\circ\, dp = 0 \\ dg_A &= \nabla_z g_A^\circ\, dz + \nabla_p g_A^\circ\, dp = 0 \\ dh &= \nabla_z h^\circ\, dz + \nabla_p h^\circ\, dp = 0 \end{aligned} \tag{4}$$

Rearranging these expressions results in **the linear system of equations:**

$$\begin{bmatrix} \nabla_{zp}L^{T\circ} \\ \nabla_{z}g_{A}^{T\circ} \\ \nabla_{z}h^{T\circ} \end{bmatrix} = - \begin{bmatrix} \nabla_{zz}L^{\circ} & \nabla_{z}g_{A}^{\circ} & \nabla_{z}h^{\circ} \\ \nabla_{z}g_{A}^{T\circ} & 0 & 0 \\ \nabla_{z}h^{T\circ} & 0 & 0 \end{bmatrix} \begin{bmatrix} \nabla_{p}x^{T\circ} \\ \nabla_{p}u^{T\circ} \\ \nabla_{p}v^{T\circ} \end{bmatrix} \tag{5}$$

In terms of the decision and tear variables Eq. 5 can be reformulated in the form:

$$\begin{bmatrix} \nabla_{xp}L^{T\circ} \\ \nabla_{J/P}L^{T\circ} \\ \nabla_{x}h^{T\circ} \\ \nabla_{x}g_{A}^{T\circ} \end{bmatrix} = - \begin{bmatrix} \nabla_{xx}L^{\circ} & \nabla_{xy}L^{\circ} & \nabla_{x}h^{\circ} & \nabla_{x}g_{A}^{\circ} \\ \nabla_{yx}L & \nabla_{yy}L^{\circ} & \nabla_{y}h^{\circ} & \nabla_{y}g_{A}^{\circ} \\ \nabla_{x}h^{T\circ} & \nabla_{y}h^{T\circ} & 0 & 0 \\ \nabla_{x}g_{A}^{T\circ} & \nabla_{y}g_{A}^{T\circ} & 0 & 0 \end{bmatrix} \begin{bmatrix} \nabla_{p}x^{T\circ} \\ \nabla_{p}y^{T\circ} \\ \nabla_{p}u^{T\circ} \\ \nabla_{p}v^{T\circ} \end{bmatrix} \tag{6}$$

If Strict Complementary Slackness holds the active set identified for the base optimal **solution is retained within an € - neighborhood of the nominal parameters** (Fiacco, 1976). **The derivative vector on the right side of Eq, 6 carries the information** regarding the **directional derivatives of the decisions and tears and the KKT multipliers at** the optimum. **This information can be used to calculate first order deviations** in **the** optímal variable **vector corresponding to a change, Ap, in the parameters, p. Therefore,** in the neighborhood **of the base case optimum we have (for CO = x,≤/,u.v):**

$$\tilde{\omega} = \omega^{\circ} + \nabla_{p}\omega^{T\circ}\,\Delta p \tag{7}$$

where

$\tilde{\omega}$ **- modified optimal solution vector**

For flowsheeting applications, the gradients of the Lagrange function, L and constraints in

Eq. 6 are usually generated by a numerical approximation so that (for \J/ = $V_X L$, V ^ . g_A, *h):*

$$\nabla_{p_k} \psi^\circ = \frac{\psi(p_k^\circ + \Delta p_k) - \psi(p_k^\circ)}{\Delta p_k} \qquad \text{<<>}$$

Alternatively, a response to a specified change in a linear combination of the parameters, Ap, can be made by computing the directional derivatives for x, ç/. a and v. From Eq. 6 and Eq. 7, we have:

$$V^2 L|ACJ] = V^2 L \hat{} 7_p 00^T \; Ap \hat{} = -V_p (V L)^T \Delta p = -\Delta(\nabla L) \qquad (9)$$

where

$$VL \equiv [V_X L \;\; \nabla_y L \;\; g_\lambda \;\; h]$$

The last term on the right hand side can be calculated by the finite difference approximation:

$$\Delta(V L^\circ) = \frac{VUp^\circ + €Ap) - VU\hat{p})}{€} \qquad (10)$$

Of course, we note that the first order sensitivities are necessarily accurate only for a small €Ap. In addition, in calculating derivatives using finite difference formulae, there are a number of factors that contribute to errors in these directional derivatives and these must be carefully controlled.

In order to solve Eq. 9 we need to construct the coefficient matrix consisting of first and second partial terms. The optimization of the parametric base case flowsheet readily supplies the gradient information V ^ , V ^ . (for ^ - ¬^. *g* h).* In addition we require the

Hessian matrix, B, given by:

$$B = \begin{bmatrix} \nabla_{xx}L^{\circ} & \nabla_{xy}L^{\circ} \\ \nabla_{yx}L^{\circ} & \nabla_{yy}L^{\circ} \end{bmatrix} \quad \text{(ID}$$

Unless the second derivatives are inexpensive to calculate and an actual constrained Newton method is used, the B matrix requires some effort to calculate. Many standard nonlinear programming algorithms approximate such matrices of second partials with quasi-Newton formulae. While these enhance the efficiency of the optimization, quasi-Newton formulae for B are inappropriate for sensitivity analysis. A justification for requiring the exact B matrix is given in Appendix II.

REDUCED HESSIAN EVALUATION

The motivation for constructing a Hessian in the reduced space of decision variables comes from recognizing that the dimensionality of the decision variables, x is often much smaller than that of the tear variables, *y*. Consequently, large savings in computations can be achieved by decomposing the linear system in Eq. 9 so that the decision and tear variables are decoupled. The smaller set of decision variable deviations can be solved independently and can then be used to solve for the larger set of tear variable deviations. By working in the reduced sj&ce, a smaller matrix is constructed by perturbing x and *y* simultaneously so that the linearizations of the equality constraints are always specified

The details of this decomposition are presented in Appendix III. Using block Gaussian elimination on the first two rows of Eq. 9. the resulting linear system is given by

$$\begin{bmatrix} a \\ b \\ f \\ e \end{bmatrix} = - \begin{bmatrix} I & 0 & E & 0 \\ 0 & I & L & M \\ 0 & 0 & H & Q \\ 0 & 0 & Q^{T} & 0 \end{bmatrix} \begin{bmatrix} A\text{«}/^{\circ} \\ AV^{\circ} \\ AX^{\circ} \\ Att^{\circ} \end{bmatrix} \quad (12)$$

where

I = Identity Matrix

$E = -A = CVyh^{1}*)-* V_x A^{To}$

$L = (\nabla_y h^\circ)^{-1} (\nabla_{yx} L^\circ - \nabla_{yy} L^\circ E>$

$M = (\nabla_y h^\circ)^{-1} \nabla_y g_A^\circ$

$Q = *xX \sim V <V^{O)''} V*$

$H = V_X x L^\circ - \nabla_{xy} L^\circ (\nabla_y h^{T\circ})^{-1} \nabla_x h^{T\circ} - \nabla_x h^\circ (\nabla_y h^\circ)^{-1} (\nabla_{yx} L^\circ + \nabla_{yy} L^\circ A)$

$a = (\nabla_y h^{T\circ})^{-1} \Delta h^\circ$

$e = \Delta g^\circ - \nabla_y g_A^{T\circ} (\nabla_y h^{T\circ})^{-1} \Delta h^\circ$

$b = (\nabla_y h^\circ)^{-1} \Delta(\nabla_y L^\circ) - (\nabla_y h^\circ)^{-1} \nabla_{yy} L^\circ (\nabla_y h^{T\circ})^{-i} Ah^\circ$

$f = A(V_X L^\circ) - \nabla_{xy} L^\circ (\nabla_y h^{T\circ})^{-1} \Delta h^\circ - \nabla_x h^\circ (\nabla_y h^\circ)^{-1} \Delta(\nabla_y L^\circ) +$
$\nabla_x h^\circ (\nabla_y h^\circ)^{-1} \nabla_{yy} L^\circ (\nabla_y h^{T\circ})^{-1} \Delta h^\circ$

The effective reduced space linear system of equations can then be written as

$$\begin{bmatrix} f \\ e \end{bmatrix} = - \begin{bmatrix} H & Q \\ Q^T & 0 \end{bmatrix} \begin{bmatrix} AX^\circ \\ Att^\circ \end{bmatrix} \qquad (13)$$

The reduced matrix, H can be constructed by simultaneously perturbing the decision and tear variables accordingly *as* (AX. $-(7yh^T)V_x h$ AX) (cf. Appendix III for details).

The number of flowsheet evaluations, $NFE_{(}$. required for the reduced form in Eq. 13 is given by the sum:

$$NFE, = (n_\chi (n_\chi + 1) / 2) + 3 n_\chi + 2 \qquad (14)$$

where

n_x - number of independent (decision) variables

In the combined space of the decision and tear variables the corresponding number of flowsheet evaluations required, NFE_2, for Eq. 9 is:

$$NFE_2 = ((n_x + n_y^*) (n_x + n_y^* + 1) / 2) + 1 \qquad (15)$$

where

Hy - number of dependent (tear) variables

The savings in the number of flowsheet evaluations compared with that required for complete Hessian evaluation is thus:

$$NFE_2 - NFE, = V^* (n_y^2 + 2 /?_x (n_y - 1) \cdot n_y) - 2 n_x - 1 \qquad (16)$$

It can be seen immediately that the savings in the number of flowsheet evaluations is directly proportional to the square of the dimension of the dependent variables, $n^$ By eliminating the need for developing explicit second partial information in the tear variable space, the reduced Hessian procedure significantly decreases the computational overhead for sensitivity analysis. The decision space second partials can be obtained directly by introducing corresponding flowsheet perturbations. Sensitivities for x and a are readily obtained; those for *y* are backed out from E Ax. The only information not obtained from Eq. 12 are sensitivities for v. Normally sensitivities for equality constrained multipliers are not as important as those for state variables.

To illustrate the reduced Hessian sensitivity approach* we first consider a small analytical example.

Example 1: Parametric Sensitivity Problem

Consider the following minimization problem:

$$\begin{aligned} \underset{x_1,y_1,y_2}{Min} \quad & \phi: x_1^2 + y_1^2 + y_2^2 \\ \text{s.t.} \quad & h_1 = 6\,x_1 + 3\,y_1 + 2\,y_2 - \alpha_1 = 0 \\ & h_2 = \alpha_2\,x_1 + y_1 - y_2 - 1 = 0 \end{aligned}$$

The problem includes one independent variable, x_1, and two dependent variables, y_1 and y_2. The terms α_1 and α_2 correspond to the parameters. We wish to analyze the sensitivity of the optimum to perturbations in these parameters about their nominal values, which in this case may be taken as, $\alpha_1^\circ = 6.0$ and $\alpha_2^\circ = 1.0$

The Lagrange function for this problem can be written as:

$$L = x_1^2+y_1^2+y_2^2 + v_1\,(6x_1+3y_1+2y_2-6) + v_2\,(x_1+y_1-y_2-1)$$

The base case optimum can be found to be at $(x_1^\circ, y_1^\circ, y_2^\circ) \equiv (0.7449,\ 0.4082,\ 0.1531)$ with the associated KKT multipliers $(v_1^\circ,\ v_2^\circ) \equiv (-0.2245,\ -0.1429)$. Also we can evaluate the following terms:

$$\nabla_x h^\circ \equiv \begin{bmatrix} 6 & \alpha_2^\circ \end{bmatrix} \equiv \begin{bmatrix} 6 & 1 \end{bmatrix}$$

$$\nabla_{y_1} h^\circ \equiv \begin{bmatrix} 3 & 1 \end{bmatrix}$$

$$\nabla_{y_2} h^\circ \equiv \begin{bmatrix} 2 & -1 \end{bmatrix}$$

To solve for the perturbed optimum due to a perturbation, $\Delta\alpha_1$ & $\Delta\alpha_2$, in the

parameters, α_1 and α_2, about their nominal values we use the reduced Hessian strategy. Using this approach (cf. Appendix III for details) the linear system to detemine the deviations in the *decision* space becomes:

$$H\ \Delta x_1 = -\ f$$

where

$$H \equiv \left[\ 7.84\ \right]$$

$$f \equiv \left[\ \Delta\alpha_2\ v_2^{\circ} - 0.88\ \Delta\alpha_1 + 0.56\ \Delta\alpha_2\ x_1^{\circ}\ \right]$$

If we assume $\Delta\alpha_1 = 0.1$ and $\Delta\alpha_2 = 0.05$, the corresponding deviations in the variable, x_1 can be found to be: $\Delta x_1 = 9.4752 \times 10^{-3}$. The parametric sensitivity analysis predicts the independent variable value at the perturbed optimum to be:

$$x_1^M = x_1^{\circ} + \Delta x_1 = 0.7544$$

Using the information regarding the sensitivity of the independent variable, we can solve for corresponding sensitivities of the dependent variables, Δy_1 and Δy_2, so that

$$y_1^M = y_1^{\circ} + \Delta y_1 = 0.4082 - 0.0101 = 0.3981$$

$$y_2^M = y_2^{\circ} + \Delta y_2 = 0.1531 + 0.0367 = 0.1898$$

Therefore perturbed optimum with $\alpha_1^M = 6.1$ and $\alpha_2^M = 1.05$ is found to be at

$$(x_1^M,\ y_1^M,\ y_2^M) \equiv (0.7544,\ 0.3981,\ 0.1898)$$

The true optimum for the same set of parameter values lies at:

$$(x_1^*,\ y_1^*,\ y_2^*) \equiv (0.7540,\ 0.3985,\ 0.1902)$$

SENSITIVITY ANALYSIS FOR MODEL VARIATIONS

Usually, the process optimum is sensitive not just to uncertain or variable process parameters but to the choice of the process models as well. Often the process models, such as those describing unit operations (Stewart, 1983; Klein 1983) or physical properties (O'Connell, 1983; Grens, 1983) can be difficult and expensive to solve. For optimization as well as for simulation, more complicated models are frequently needed because of the accuracy they provide.

A frequently asked question, that is often problem dependent, regards which models are accurate yet simple enough for process optimization. For process simulation this question is often resolved by running competing models side by side. In fact, for more efficient operation, simple models are often embedded within simulators to speed up the solution of more rigorous models. This is especially helpful for physical property calculations (Chimowitz *et a/.*, 1983; Bryan and Grens, 1983). For process optimization the use of competing models may lead to very different results even though solutions of simulation problems may be similar. On the other hand, the nature of the optimization problem may lead two competing and functionally different models to identify the same active constraint set and perhaps even the same values for the decision variables as the optimal ones.

In this paper^we develop a strategy for evaluating the sensitivity of the optimal solution to the choice of the process model (e.g. thermodynamic and/or unit operations model). We note that this problem is conceptually different from parametric sensitivity because process relationships and not parameters are being changed. Consequently, the methods discussed in the first part cannot be applied directly to this problem. Instead we consider the first order sensitivity (or direction) if one starts from the optimum of Model I and takes a Newton step for the optimality conditions in the space of Model II. Using this concept we develop a slightly different strategy that allows application of some features of Part I, in particular.

the reduced Hessian procedure.

Consider the model based optimization problem:

$$\begin{aligned} &\min_{x,y,M} && \phi'(x,y,k) \\ &\text{s.t} && g\backslash x.yM) \le 0 \\ & && tf(x.yM) = 0 \\ & && r(x.y_9k) = \text{fe} - \text{K(x,y)} = 0 \end{aligned} \tag{17}$$

where

fr - objective function

g^9 - inequality constraint

tt - equality (tear) constraint

fe - physical (model based) property

K - model for property evaluation

The property, fe, in Eq. 17 is estimated by using Model I given by, fe = $K(x.y)$. At the optimum with respect to Model I, ($x_{|t}$ $j/_t$. $u_{|t}$ v_f. $k_{\%}$). from the KKT conditions we have:

$$\begin{aligned} &\nabla_x L'(x_1.y_1.k_t) = 0 \\ &\nabla_y L'(x_1.y_1.\text{fe},) = 0 \\ &\nabla_k L'(x_1.y_1.\text{fe},) = 0 \\ &g_A'(x_1.y\wedge\text{fe},) = 0 \\ &h'(x_{|y_x}M_x) = 0 \end{aligned} \tag{18}$$

where

V - Lagrange function

g_A^9 - active inequality constraint

Once again, the first order corrections to satisfy these conditions corresponding to a small perturbation, Δk, in the property, k, about k_1, can be found from:

$$
\begin{aligned}
d\,(\nabla_x L'(x_1,y_1,k_1)) &= \nabla_{xx}L_1'\,dx + \nabla_{xy}L_1'\,dy + \nabla_{xk}L_1'\,dk = 0\\
d\,(\nabla_y L'(x_1,y_1,k_1)) &= \nabla_{yx}L_1'\,dx + \nabla_{yy}L_1'\,dy + \nabla_{yk}L_1'\,dk = 0\\
d\,(\nabla_k L'(x_1,y_1,k_1)) &= \nabla_{kx}L_1'\,dx + \nabla_{ky}L_1'\,dy + \nabla_{kk}L_1'\,dk = 0\\
dg_A' &= \nabla_x g_{A1}'\,dx + \nabla_y g_{A1}'\,dy + \nabla_k g_{A1}'\,dk = 0\\
dh' &= \nabla_x h'_1\,dx + \nabla_y h'_1\,dy + \nabla_k h'_1\,dk = 0\\
df' &= -\nabla_x K\,dx - \nabla_y K\,dy + I\,dk = 0
\end{aligned}
\tag{19}
$$

If the competing model (Model II) is given by, $k = K'(x,y)$, then we are interested in finding out how the decision variables and the active inequalities are modified with respect to the new model. The sensitivity relationships in Eq. 19 have been expressed explicitly in terms of the model equations. Since, the property evaluated by the model, k, is essentially a function of the decision and tear variables, this system can be first transformed to a form that implicitly accounts for the presence of the model. This proves advantageous when the dimension of the model variable, k, is large, as is almost always the case with stage–wise unit operations such as distillation involving multicomponent systems. The equivalent problem can be formulated as:

$$
\begin{aligned}
\underset{x,y}{min} \quad & \phi\,(x,\ y,\ K(x,y))\\
s.t. \quad & g\,(x,\ y,\ K(x,y)) \le 0\\
& h\,(x,\ y,\ K(x,y)) = 0
\end{aligned}
\tag{20}
$$

For Model I, $K(x,y)$, the optimum has been assumed to be at $\bar{z} = (x_1,\ y_1)$. However the

KKT conditions at the same point may not be satisfied with respect to Model II, fe = K*(x,y), Le.

$$\begin{bmatrix} \nabla_x L \\ \nabla_y L \ (x_x, \ y_x, \ K'ix^) \\ g_A(x_x, \ y_x, \ K'(x_t 4_t)) \\ ttx_x, \ y_x, \ K'Ot,^,)) \end{bmatrix} * 0 \qquad (21)$$

where

L g. h - Lagrange function, inequality constraint and equality constraint defined w.r.L Model II.

Let us assume that the optimum with respect to Model II is at % = (x_2, y_2). If we choose a consistent active set (we will discuss this point later) we can write a first order correction for the optimality conditions with respect to Model IL Defining a Newton step in the space of Model II leads to:

$$\begin{bmatrix} \nabla_x L_2 \\ V^* \\ g_{A2} \\ h_2 \end{bmatrix} \simeq \begin{bmatrix} V. \\ g_{A1} \\ h_1 \end{bmatrix} + \begin{bmatrix} \nabla_{xx} L_1 & \nabla_{xy} L_1 & \nabla_x g_{A1} & V.' \\ \nabla_{yx} L_1 & \nabla_{yy} L_1 & \nabla_y g_{A1} & V. \\ \nabla_x g_{A1}^T & \nabla_y g_{A1}^T & O & 0 \\ \nabla_x h_1^T & V. & 0 & 0 \end{bmatrix} \begin{bmatrix} AX \\ \Delta y \\ AU \\ Av \end{bmatrix} = 0 \qquad (22)$$

which gives a first order correction for the optimal solution for Model II. Rearranging this equation we have:

$$\begin{bmatrix} V. \\ ^A, \\ h_1 \end{bmatrix} = - \begin{bmatrix} \nabla_{xx} L_1 & \nabla_{xy} L_1 & \nabla_x g_{A1} & \nabla_x h_1 \\ \nabla_{yx} L_1 & \nabla_{yy} L_1 & \nabla_y g_{A1} & \nabla_y h_1 \\ \nabla_x g_{A1}^T & \nabla_y g_{A1}^T & 0 & 0 \\ \nabla_x h_1^T & \nabla_y h_1^T & 0 & 0 \end{bmatrix} \begin{bmatrix} AX \\ Ay \\ AU \\ AU \end{bmatrix} \qquad (23)$$

The gradients required to solve Eq. 23 are now evaluated with respect to Model II, at the optimal solution for Model I.

In developing the above relationships we first consider the case where the same active set with respect to Model I has been assumed to be retained at the optimum for Model IL The KKT muliplier $\hat{u}_j$, corresponding to an active inequality, g_j, therefore equals u^ provided g_j is active for Model II. If this is true, the Aa obtained from solving Eq. 23 is such that itj + Aa remains positive. Once again the reduced Hessian strategy, discussed previously, can be applied to the linear system in Eq. 23 so that the corrections for x and a are obtained first; the sensitivity for *y* can be computed using the Ax information. We now consider an analytical example to demonstrate this approach.

Example 2: Model Sensitivity Problem Under Active Set Retention

Consider the following model based optimization problem:

$$\underset{x,k}{Min} \quad \langle f \rangle: \; 4\,(x - 4)^2 + 9\,(k - S)^2$$

$$s.L \quad g_{\%}: \; k + 5\,x - 23.2 \leq 0$$

$$gj \;\; k - 8 \leq 0$$

$$k \leq 0 \; ; \; x \leq 0$$

The variable, k, represents the model based property in this problem. To illustrate the model sensitivity approach, we consider two simple defining equations for the property, k, defined in terms of the variable, x :

$$\text{Model I:} \quad k - 0.2\,x^2 = 0$$

$$\text{Model II:} \quad k - e^{\circ.3776x} = 0$$

We first compute the optimum for this problem with respect to Model I. The optimal solution lies at:

$$(x_1,\ k_1,\ u_1^1,\ v_1) \equiv (\ 4,\ 3.2,\ 7.8545,\ 24.5455\ )$$

The KKT multiplier, u_1^1, corresponds to the active inequality constraint, g_1, at the optimum with respect to Model I; it can be directly seen that the inequality constraint, g_2 is inactive at this point so that the corresponding KKT multiplier, $u_1^2 = 0$. Our objective is to compute the sensitivity of the optimal solution when the defining model for k is changed from Model I to Model II.

If we assume that the active set determined for the optimum with respect to Model I is retained at the optimum for Model II, we can write the Lagrange function with respect to Model II at z_1 as:

$$L = 4\ (x - 4)^2 + 9\ (k - 5)^2 + u\ (k + 5\ x - 23.2) + v\ (k - e^{0.3776x})$$

By taking a Newton step in order to determine the first order correction to the problem variables, the resulting linear system becomes:

$$\begin{bmatrix} -7.8567 & 0 & 5 & -1.7106 \\ 0 & 18 & 1 & 1 \\ 5 & 1 & 0 & 0 \\ -1.7106 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta k \\ \Delta u^1 \\ \Delta v \end{bmatrix} = - \begin{bmatrix} -2.7148 \\ 0 \\ 0 \\ -1.3295 \end{bmatrix}$$

The solution to this system predicts the following corrections for the variables and the KKT

multipliers:

$$\Delta x = -0.19813 ; \qquad \Delta k = 0.9906$$

$$\Delta u^1 = -4.3727 ; \qquad \Delta v = -13.4584$$

The predicted optimal solution with respect to Model II therefore becomes:

$$(x_2, k_2, u_2^1, v_2) \equiv (3.8019, 4.1906, 3.4818, 11.0871)$$

If we carry out the optimization directly with respect to Model II we find that the true optimum lies at

$$(x_2, k_2, u_2^1, v_2) \equiv (3.8, 4.2, 3.7109, 10.6891)$$

The assumed active set remains consistent under the model change as evidenced by the value of the KKT multiplier, u^1. Fig. 1 gives a physical picture of the change in the optimal solution from Model I to Model II; both the solutions, A and B, lie on the same active constraint, g_1.

MODEL SENSITIVITY ANALYSIS UNDER ACTIVE SET VARIATIONS

In general, the active set determined for Model I need not be retained at the modified optimum for Model II. When this occurs the correct active set is not known *a priori* to construct the sensitivity relationships. If the Newton step from the Model I optimum, $\bar{z}$, is small we can assume that the active set can be determined by first order corrections of x, y and u. Thus we have,

$$\nabla_{xx}L(\bar{z},\hat{u},v_1)\,\Delta x + \nabla_{xy}L(\bar{z},\hat{u},v_1)\,\Delta y + \nabla_x L(\bar{z},u,v) = 0$$

$$\nabla_{yx}L(\bar{z},\hat{u},v_1)\,\Delta x + \nabla_{yy}L(\bar{z},\hat{u},v_1)\,\Delta y + \nabla_y L(\bar{z},u,v) = 0$$

$$\nabla_x g^T(z)\ AX \bullet V\wedge^T(r)\ Ay \bullet g(T)\ z\ 0 \qquad (24)$$

$$V_x/?^T(r)\ AX \bullet V\wedge/?^T(r)\ Ay + /HD = 0$$

$$u = \hat{u} + Au \pounds 0\ ;\quad v = v_1 + Av$$

$$(\hat{u} + Au)\ (\ V_x 5r^T(D\ Ax \bullet V \wedge F\ )\ Ay * \$r(F)\ ) = 0$$

which can be found from the following quadratic program, using $z \equiv \begin{Bmatrix} x \\ y \end{Bmatrix}$:

$$\begin{aligned} \textit{Min}\quad & V<t>\backslash T)\ AZ \bullet \backslash Az^1\ V_{22}L(TSLv_x)\ AZ \\ \text{s.t.}\quad & g(z) + Vg^T<r>\ AZ <L\ 0 \\ & h(z) + V/?^T(F)\ Az = 0 \\ & u = \hat{a} \bullet Au\ ;\quad v\ ^s\ v_x \bullet Av \end{aligned} \qquad (25)$$

We assume that $<F) has linearly independent columns because this guarantees strict complementary slackness for the QP solution. Le.

$$a\ (\ g(T) \bullet Vflr^T(F)\ Az\ ) = 0$$

implies

$$\begin{aligned} \text{for}\quad & giT) + Vg^J(T)\ Az = 0.\quad a > 0 \\ \text{for}\quad & g(T) + V\wedge^T(F)\ Az < 0.\quad a = 0 \end{aligned} \qquad (26)$$

However the term. $\hat{a}$. in V^2/L is unknown. This gives rise to two different cases.

Case A Assume $\hat{a} = a_j$ and solve the QP to determine u. If $a_j f\ 0$. for all $a_j * 0$. then we have chosen a consistent active set This is the same situation referred to in the previous section. If this fails, we need to set some of the $\hat{a}_j$ to zero and try again.

Case B. In this case, we assume all the $\hat{a} = 0$. In doing so, we merely change the starting

point for the Newton step and let the QP select a that determine the active set This **QP** is **solvable if and only** if

$$g(T) \cdot Vg^{T}(Y)\ \mathbf{Ar} \pounds 0 \quad \text{is feasible}$$

and

$$\mathbf{Az}^{T}\ \mathbf{VZ.O\text{-}.O.v,})\ Az = -\ \mathbf{Ar}^{T}\ V\langle t\rangle(T)$$

where

$$\mathbf{Ar}^{T}\ \mathbf{Vp}_{A}\mathbf{(D = 0\ ;}\quad \mathbf{Az}^{T}\ \mathbf{VMD = 0} \tag{27}$$

Very often. Case A provides a simple enough way to estimate the correct active set. **A more rigorous way of choosing the active set** is **given** by **a Mixed** Integer Nonlinear **Programming (MINLP) formulation in Appendix IV. To see** how **the** active set changes **as a result of model variations consider a small modification of the last example.**

Example 3: Model Sensitivity under Active Set Variation

Let us once again consider the model based minimization problem:

$$\underset{x*k}{Min} \quad \$: \mathbf{4\ (x - 4)^2 \cdot 9\ (k - 5)^2}$$

$$\text{S.L} \quad g_{\%}: \mathbf{k \cdot 5\ x} - 23.2\ \pounds\ \mathbf{0}$$

$$g_2: \mathbf{k - 3.5\ \pounds\ 0}$$

$$\mathbf{k * 0\ ;\ x * 0}$$

This problem is similar to the one previously considered for model sensitivity; however, the **upper bound for the model based property, k,** has **been reduced resulting** in a change **for the inequality constraint,** g_r **Once again we wish to study the first order** correction to **the optimal solution in going from the optimum with respect to Model I, k - 0.2** x^2 **= 0. to the optimum with respect to Model II, k - e°*377** — **a**

We first assume the case when the active set is retained The solution to this case, as obtained in Example 1, clearly makes g_2 infeasible (Fig. 2). By setting $\hat{u}^1 = 0$ and introducing g_2 into the sensitivity equation, with $\hat{u}^2 = 0_9$ (Case B) we develop the modified system for the sensitivity relationship:

$$\begin{bmatrix} -7.8567 & 0 & 0 & -1.7106 \\ 0 & 18 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ -1.7106 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} Ax \\ Ak \\ AU^2 \\ Av \end{bmatrix} = - \begin{bmatrix} -41.9873 \\ -7.8545 \\ -0.3 \\ -L3295 \end{bmatrix}$$

The perturbed optimum based on the first order correction is found to be:

$$(x_2.\ k_2.\ u_2^*.\ v_2) \equiv (3.3982,\ 3.5,\ 24.2357,\ 17643)$$

The true optimum with respect to Model II is at

$$(x_2.\ k_2.\ u_2^*.\ v_2) \equiv (3.3172,\ 3.5.\ 31.1325.\ -4.1324)$$

It may be noted that the constraint g_x remains feasible; however the model change introduces a chapge in the active set

For illustration, the above examples use a full Hessian approach for the Newton steps. We note that the reduced Hessian strategy can be applied in a straightforward manner for model sensitivity analysis. In the following process examples, a reduced, Hessian approach is employed for both parametric and modd sensitivity analysis.

A Simple Flash Recycle Flowsheet

A simple flash recycle problem flowsheet serves to demonstrate the application of the reduced Hessian strategy for parametric sensitivity and the model sensitivity approach. The flowsheet is presented in Fig. 3. A light hydrocarbon feed is mixed with recycled bottoms and flashed adiabatically. The vapor is removed as a product and the liquid is split into a bottoms product and the recycle, which is pumped back to the feed. The problem specifications are presented in Table. L The process was optimized using the simulator SPAD from the University of Wisconsin-Madison on a DEC-20 computer.

The flowsheet includes two decision variables, the splitter ratio and the pressure in the flash. The six component flowrates and specific enthalpy of the recycle stream together constitute the seven tear variables for the problem. Since the outlet pressure of the pump is prefixed in this case the recycle stream pressure does not figure in the set of tear variables. The objective function for the *monotonic* optimization problem corresponds to the flowrate of the lightest component in the overheads from the flash; for the *nonlinear* case a predetermined combination of the component flows in the flash overheads was maximized.

For the parametric senstivity study, the flowrates of the components in the hydrocarbon feed stream were perturbed about their nominal values. The results of the parametric sensitivity analysis for both monotonic and nonlinear optimization problems are shown in Table. II.

In the model sensitivity case, two competing models were employed to compute the physical properties for vapor-liquid equilibrium in the flash: the Ideal Raoulfs law model and the Ghao-Seader model. The sensitivity of the Ideal Optimum on applying a first order correction procedure with respect to the Cbao-Seader model was first studied for both monotonic as well as nonlinear objectives. Going from the Chao-Seader optimum to the ideal Raoulfs Law model was considered next

The model sensitivity results are given in Table. III. For the monotonic objective function, the splitter ratio and the flash pressure stay at their respective lower bounds regardless of the modeL It may be noted that the active set is retained and the KKT multipliers corresponding to the active bounds, u = Uj+Au, remain positive. However, in going from the Raoult's Law Model to the Chao-Seader model, the nonlinear objective function gives rise to the case in which the active set is no longer retained; in this case the problem is solved by setting all $\hat{a} = 0$ and the solution to the sensitivity equation gives a feasible move for the decision variables. The requirement for satisfying strict complementarity makes the modified KKT multipliers for the inequality constraints, $a^1 = a^2 = 0$ since both decisions are no longer at their bounds. Note, however, that since the constraint corresponding to the decision variable bound happens to be linear, the reduced Hessian is not affected in the above case. Furthermore the true optimum for the Chao-Seader model makes the constraint, x, £ 0.8, active. We will comment on this after considering the following flowsheet example.

Monochlorobenzene Separation Flowsheet

This problem is adapted from an example in the FLOWTRAN manual (Fig. 1.1. Seader *et al.*% 1977). The flowsheet for this problem is shown in Fig. 4. A mixture of HC1, Benzene and Monochlorobenzene is fed to the separation process; benzene and monochlorobenzene are separated in a distillation column and a part of the bottoms from this column is split as recycle and fed to the absorber. The process was optimized using the FLOWTRAN simulator on a VAX-11/780. The base case feed flows (parameters) and the process models used for computing vapor-liqiud equilibrium properties are given in Table. IV.

Once again for the parametric case a variation in the feed was introduced at the base case optimum. The problem involves 6 decision variables and 5 tear variables. The problem was solved using the reduced Hessian procedure; the results of the parametric sensitivity analysis are presented in Table. V. The number of flowsheet evaluations in the reduced Hessian case

with 6 decision variables is 41 and the corresponding CPU time required is 250.1 seconds. A full Hessian computation for the same problem requires 67 flowsheet evaluations corresponding to an equivalent time consumption of 408.7, both times computed on a VAX-11/780. The reduction in time consumption for sensitivity analysis from the reduced Hessian strategy is of the order of 39% compared with full Hessian evaluation.

Sensitivity of the optimal solution to a variation in process model was studied by changing the thermodynamic model that evaluates the liquid phase activity coefficients for the components. The Ideal Solution option in FLOWTRAN was used as the base case model (Model I) to compute this property; the Regular Solution Model was chosen as the alternative model (Model II) to evaluate the activity coefficients. The results of the model sensitivity analysis are shown in Table. VI. In solving the sensitivity equations using the reduced Hessian strategy, the active set determined for the base case optimum has been assumed to be retained. The results indicate that this active set *is* retained at the optimal solution determined from a first order correction. However, at the true flowsheet optimum corresponding to Model II (Regular Solution) we have a different active set with the Absorber Input Temperature, x_6 reaching its lower bound. The total number of flowsheet evaluations required in this case is 52 corresponding to a CPU time requirement of 317.2 seconds whereas a full Hessian computation for model sensitivity analysis would require 78 flowsheet evaluations with an equivalent time requirement of 475.8 seconds; the savings in this case is of the order of 33%.

It may be noted that the number of tear variables in this flowsheet is in fact less than the number of recycle variables and hence the savings in computation time is limited. In general the number of tear variables exceeds the number of decision variables by a factor of 3 or 4 which makes the reduced Hessian procedure computationally very attractive for sensitivity analysis.

The treatment of model sensitivity as a Newton step in the space of the *new* model uses first order linearization in the decisions and tears, about the base case optimum to satisfy the KKT conditions at the modified optimal solution. The accuracy of this linearization dictates the prediction of the correct active set for the modified optimum. The above examples indicate that, in general, the model sensitivity algorithm performs well in determining the changes in the optimal solution to model variations. If the objective function or the constraint happens to be highly nonlinear then the linearization may not be sufficient to predict the change in the active set accurately. This can be seen from the nonlinear objective function problem (Raoult's Law to Chao-Seader), for the simple flash flowsheet. The linearization also controls the trend in the predicted change for a process variable; note that in the monochlorobenzene separation problem the actual change in the absorber input temperature, x_6, is in a direction opposite to the one obtained from model sensitivity analysis. Even in this case, the approach gives the right predictions for sensitivity directions for all other decision variables.

CONCLUSIONS

Significant computational savings are realized by applying the *reduced* Hessian algorithm to determine the sensitivity of an optimal solution of a process flowsheet to parametric variations. The reduced Hessian strategy, which yields sensitivity information on all but the equality constraint multipliers, performs the finite difference perturbations required for constructing the Hessian only in the space of the independent (decision) variables. The reduction in the number of flowsheet evaluations is seen to be proportional to the square of the number of the dependent (tear) variables. The procedure is very efficient especially for sensitivity analysis of medium to large sized flowsheets that involve multiple components and recycle loops.

The sensitivity of the optimal solution to changes in process models has been treated as a

Newton step in the space of the *new* model This results in a linear system of equations similar to the one obtained for parametric sensitivity and allows the reduced Hessian procedure to be directly extended for model sensitivity analysis. The problem of determining the correct active set can be dealt with rigorously by formulating and solving a Mixed Integer Nonlinear Program (MINLP). From limited calculations with a simple flash recycle flowsheet and a monochlorobenzene separation flowsheet, the model sensitivity approach seems to perform well in predicting the changes in the flowsheet optimum to model variations.

REFERENCES

1. **Athertoru R. W., Schainker, R. R** and **E R. Ducot, "On the** Statistical Sensitivity **Analysis of Models for Chemical Kinetics-. *AlChE J,* 21. 411 (1975).**

2. **Berna, T. J., Locke, M. H. and A. W. Westerberg. "A New Approach** to Optimization **of Chemical Processes", *AlChE J,* 26 37 (1980).**

3. **Biegler. L. T. and R. R. Hughes. "Infeasible Path Optimization of** Sequential Modular **Simulators". *AlChE* 7. 28 994 (1982).**

4. **Boggs. P. T.. Tolle, J. W. and P. Wang. "On the Local Convergence of** Quasi-Newton **Methods for Constrained Optimization". *SIAM J Control and* Optimization, 20 161 (1982).**

5. **Bryan, P. F. and E. A. Grens II. "Efficient Use of Thermodynamic** Properties for **Phase-Equilibrium Calculations in Solvent Extractions[9]*. Int.** Solvent Extraction **Conference, Denver, CO (1983).**

6. **Chimowitz, E. H., Anderson. T. F.. Macchietto. S. and L. F- Stutzman,** "Local Models **for Representing Phase Equilibria in Multicomponent Nonideal** Vapor-Liquid and **Liquid-Liquid Systems. 1. Thermodynamic Approximation Functions",** *i & EC Proc. Des. & DevU* **22 217 (1983).**

7. **Duran. M. A. and I. E Grossmann. "An Outer-Approximation Algorithm** for a Special **Class of Mixed-Integer Nonlinear Programs". *DRC Report 06-58-83,*** Carnegie-Mellon **University (1983).**

8. **Edahl, R-, "Sensitivity Analysis in Nonlinear Programming** - Applications **to Comparative Statics and Algorithmic Construction". Ph.D Thesis,** Carnegie-Mellon **University (1982).**

9. **Fiacco, A. V. and G. P. McCormick. *Nonlinear Programming:* *Sequential Unconstrained Minimization Techniques.** Wiley, New York (1968).**

10. *J* **Fiacco. A. V., "Sensitivity Analysis for Nonlinear Programming** Using Penalty **Methods", *Mathematical Programming,* 10 287 (1976).**

11. **Gallagher, S. A. and M. A. Kramer. "Parametric Sensitivity** Analysis of Chemical **Process Flowsheets", presented at AlChE Meeting. San Francisco (1984).**

11 **Grens II. E. A.. "Efficient Use of Thermodynamic Models in** Process Calculations", **2nd InL Conf. on Foundations of Computer-Aided Process** Design. Snowmass, **CO (1983).**

13. **Jirapongphan. S., Boston, J. F., Britt, H. L and L. B.** Evans. "A Nonlinear

Simulatancous Modular Algorithm for Process Flowsheet Optimization", presented at 80th AIChE Meeting, Chicago (1980).

14. Klein, H. H., -Computer Modeling of Chemical Process Reactors", 2nd InL Conf. on Foundations of Computer-Aided Process Design, Snowmass, CO (1983).

15. Luenberger, D. G., *Introduction to Linear and Nonlinear Programming,* Addison-Wesley Publishing Company, Massachusetts, (1973).

16. O'Connell, J. P., The Structure of Thermodynamics in Process Calculations", 2nd InL Conf. on Foundations of Computer-Aided Process Design, Snowmass, CO (1983).

17. Powell, M. J. D., The Convergence of Variable Metric Methods for Nonlinearly Constrained Optimization Calculations", *Nonlinear Programming* 3, O. Mangasarian, R. Meyer and S. Robinson, Eds., Academic Press. New York 27 (1978).

18. Seader, J. D., Seider, W. D. and A. C Pauls. *FLOWTRAN Simulation - An Introduction,* 2nd Edition, CACHE Corp. (1977).

19. Shivaram, S. and L. T. Biegler, "Improved Infeasible Path Methods for Sequential Modular Optimization", 3rd International Congress on "Computers and Chemical Engineering", Paris (1983).

20. Stadtherr, M. A. and H. S. Chen, "Numerical Techniques for Process Optimization by Successive Quadratic Programming", 3rd International Congress on "Computers and Chemical Engineering". Paris (1983).

21. Stewart, W. E., Levien, K. L. and M. Morari, "Collocation Methods in Distillation". 2nd InL Conf. on Foundations of Computer-Aided Process Design. Snowmass. CO (1983).

22. Volin, Y> M. and G. M. Ostrovskiy. "Sensitivity Calculation Methods for Complex Chemical Systems - 1: Algorithmization," *Comp. and Chem. Eng.*$_9$ 5 21 (1981).

APPENDIX I

The following constraint conditions are assumed to be satisfied at a *local* minimum in developing the basis for sensitivity analysis of the optimum to parametric and model variations (Edahl, 1982).

Linear Independence of Constraint Gradients

At z°, a feasible solution to P(p°), the gradients of the binding constraints form a linearly independent set, Le.

$$\sum u_j \nabla g_j(z^\circ, p^\circ) + \sum v_i \nabla h_i(z^\circ, p^\circ) = 0$$

and $g_j = 0 \quad \Rightarrow \quad u_j = v_i = 0$ for all ij

Strict Complementary Slackness [SCS]

For z°. an isolated minimizes of P(p°). SCS holds if the KKT multipliers u°. v/° are such that

$$V f^\circ + u^\circ \nabla g^\circ + v^\circ \nabla h^\circ = 0$$

$$u^{oT} g(z^\circ, p^\circ) = 0$$

$$u^\circ \geq 0$$

and $g_j(z^o, p^\circ) = 0 \Rightarrow u_j^\circ > 0$

Note: $f^\circ \Rightarrow \nabla(z^\circ, p^\circ)$; similarly g°. h°

Second Order Sufficient Condition [SOC]

SOC is said to hold for *P(p°)* at *z°* if:

1) z° is a fciasible point of *P(p°)*

There exists a°, *v°* such that

2) $\nabla_z \nabla^\circ + a^o \nabla h^\circ + v^\circ \nabla h^\circ = 0$

$u^{\circ}g(z\backslash p^{*}) = 0$

$u^{\circ} * 0$

and

3) for all non-zero q € 1R" satisfying $V\wedge_{A}{}^{T}<7 = 0$ and $V/?^{T}qr = 0$ we have

$<7^{T} V_{Z2} Ur^{\circ}, a^{o}.v^{o}.p^{\circ}) q > 0$

where $L = <j> + ug + vh$ (Lagrange function).

APPENDIX II

Let B be the Hessian of the Lagrange function (Eq. 11) and let Q^k be its approximation from a quasi-Newton update formula applied at each SQP iteration. For example, the BFGS update formula is given by :

$$Q^{k+1} = Q^k - \frac{Q^k\, d\, d^T\, Q^k}{d^T\, Q^k\, d} + \frac{\eta\, \eta^T}{\eta^T\, d} \qquad \text{II-1}$$

$$d = z^{k+1} - z^k$$

$$\eta = \nabla_z L^{k+1} - \nabla_z L^k$$

If B is positive definite, the following relation applies (Boggs *et al.*, 1982) :

$$\lim_{k \to \infty} \frac{\| z^T (B - Q^k)\, d \|}{\| d \|} \to 0 \qquad \text{II-2}$$

where z is the null space of $\left[\nabla_z g_A \quad \nabla_z h \right]^T$.

Otherwise, it has been conjectured (Powell, 1978) that the following property applies :

$$\lim_{k \to \infty} \frac{\| z^T (B - Q^k)\, z\, z^T\, d \|}{\| d \|} \to 0 \qquad \text{II-3}$$

In either case, one can show that at the limit point Q^k and B differ by A D A^T, where $A \equiv \left[\nabla_z g_A \quad \nabla_z h \right]$ and D is an unknown symmetric matrix. Edahl (1982) has shown that this difference does not affect the sensitivity of the optimal variables, but substitution of the limit point Q^k causes the sensitivity of the *multipliers* to differ by $D \left[\nabla_z g_A^T \nabla_z h^T \right]^T$.

Aside from this, Q is usually initialized arbitrarily (e.g. $Q^\circ = I$) and from Eq. II-1 one

can see that Q^k can be slow to converge to its limit point even as both d and η vanish; and the above properties hold only for Q^k at its limit point. For example, an optimization that satisfies the Kuhn–Tucker tolerance after a few iterations may have a Q^k far away from B and close to Q°. The effect of using Q^k for sensitivity is therefore inaccurate.

The only reliable way to substitute Q^k for B in the sensitivity analysis is to ensure that Q^k has converged to its limit. This could require more iterations than solving the optimization problem and therefore prove inefficient in terms of algorithmic performance. Moreover, numerical errors may prevent solution of the optimization problem to a tolerance tight enough for convergence of Q^k. Consequently, we have focused instead on an efficient scheme for calculating a reduced form of B directly.

APPENDIX HI

Reduced Hessian Decomposition Strategy

Consider the linear system of equations:

$$\begin{bmatrix} A(V_X L) \\ A(V^{\wedge} U \\ {}^{A} * {}^{A} \\ Ah \end{bmatrix} = - \begin{bmatrix} {}^{V}XX^{L} & {}^{V}xy^{L} & VA & V \\ V- & V- & V* & V \\ VI & VI & 0 & 0 \\ V^{T} & V^{T} & 0 & 0 \end{bmatrix} \begin{bmatrix} AX \\ A* \\ AU \\ AV \end{bmatrix} \qquad \text{III-l}$$

Rearranging the system of equations in Eq. III-l, we gee

$$\begin{bmatrix} Ah \\ AC7yL) \\ A(V_X L) \\ A(7_A \end{bmatrix} = - \begin{bmatrix} V & 0 & V_X/\!\!\rangle^{T} & 0 \\ {}^{v}yy^{L} & V & V^{L} & VA \\ {}^{V}x \wedge & V_X/\!\!> & {}^{V}xx^{L} & VA \\ .VI & 0 & VI & 0 \end{bmatrix} \begin{bmatrix} A* \\ AV \\ AX \\ AU \end{bmatrix} \qquad \text{III-2}$$

We note that the term, $VjJ>^{T}$. is square and non-singular. Hence on applying a Gaussian elimination to Eq. Ill—2 we get the resulting system (Berna *et a/.*, 1980):

$$\begin{bmatrix} a \\ b \\ f \\ e \end{bmatrix} = - \begin{bmatrix} I & 0 & E & 0 \\ 0 & I & L & M \\ 0 & 0 & H & Q \\ 0 & 0 & Q^{T} & 0 \end{bmatrix} \begin{bmatrix} Ay \\ AV \\ AX \\ AU \end{bmatrix} \qquad \text{111-3}$$

where the terms have been defined previously in the section on Reduced Hessian Evaluation (refer Eq. 12)

It is apparent that we have a decomposed (reduced) system only in the space of the decision

variable, x. in order to solve the sensitivity equation given by:

$$\begin{bmatrix} f \\ e \end{bmatrix} = - \begin{bmatrix} H & Q \\ Q^T & 0 \end{bmatrix} \begin{bmatrix} AX \\ Au \end{bmatrix} \qquad \text{III-4}$$

In Eq. III-4 the terms Q and Q^T are readily obtained from the information regarding the constraint gradients at the local minimum from the nominal model based optimization problem. However the terms H and f need to be derived.

Let us first consider the derivation for f. The term, $V_x/i\ (\nabla_y h)^{-1}\ \nabla_{yy} L\ (V^{\wedge T})^{*!}\ Ah.$ for f can be expanded at the base case optimum, (x, (/,)• as follows:

$$\nabla_x h\ (\nabla_y h)^{-1}\ \nabla_{yy} L\ (\nabla_y h^T)^{-1}\ Ah =$$

$$\nabla_x h\ (\nabla_y h)^{-1}\ \nabla_y L\ (y_1 + (\nabla_y h^T)^{-1} \Delta h,\ x_1) - \nabla_x h\ (\nabla_y h)^{-1}\ \nabla_y L\ \{y_x,\ x_1) \qquad \text{III-5}$$

Once again the second term on the right side of Eq. Ill—5 is directly obtained from the base case optimization results. However, the first term on the right hand side of this equation involves perturbations in the tear variable, *y*. We now proceed to show how the tear variable perturbations may be reformulated in terms of corresponding perturbations in the decision variable, x.

The equality constraints, $h = 0$, are retained under a first order perturbation in the decisions and tears Le.

$$V_X/>^T\ Ax \cdot Vyh^r\ Ay = 0$$

$$\Rightarrow \qquad Ay^T = - Ax^T\ V_x/>\ (\nabla_y h)^{-1} \qquad \text{III-6}$$

Next we expand the Lagrangc function, L, about $(x_f,\ y_x \cdot (7yh^T Y^l Ah \cdot A</)$ to first order

so that:

$$L\ (x_{,},\ y_{j} \cdot (V\wedge^{T})^{H}\Delta h + \Delta y) =$$

$$L\ (x_{1},\ y_{1} + (\nabla_{y}h^{T})^{-1}\Delta h) + \Delta y^{T}\ V\wedge L\ \langle x_{|t}\ \{\wedge \cdot (7yh^{T}r^{x}Lh) \qquad \text{IH-7}$$

We now substitute for Ay^{1} in Eq. Ill—7 from Eq. III-6 which transforms Eq. III-7 accordingly as:

$$AX^{T}\ \{\ V_{K}h\ (Vyhr^{1}\ VyL\ \langle y_{i}+\langle V_{y}/\rangle^{T})\text{-}^{1}A/\rangle.\ \dot{x},)\ \} =$$

$$L\ (x_{r}\ y_{\%} \cdot (V\wedge^{T})^{H}A/7) - L\ (x_{|t}\ c/_{1} + (7yh^{T}y^{l}Ah - (Vyh^{T})^{'lJ}7_{x}h\ AX) \qquad \text{III-8}$$

Eq. Ill—8 implies that the first term on the right hand side of Eq. Ill—5 can be constructed by introducing perturbations in the decisions alone; it is worth noting that the term. $V_{y}JL.$ need no longer be determined explicitly. Similarly, it can be shown that the term, $\nabla_{xy}X\ (V\wedge^{1})^{\cdot\cdot 1}\ A/7$, can be constructed from the relations:

$$\nabla_{xy}L\ (\nabla_{y}h^{T}r^{l}\ Ah = V_{X}L\ (x_{r}\ y_{x} + (\nabla_{y}h^{T})^{-1}\Delta h) - \nabla_{\chi}L\ (x_{1}.y_{1})$$

and

$$\bar{A}x^{T}\ V_{X}L\ (x_{r}\ y_{x} + (\nabla_{y}h^{T})^{-1}\Delta h) =$$

$$L\ (X\wedge AX.\ y_{x} + (V\wedge\bar{A}VA/J) - L\ (x_{r}\ y_{x} \cdot (V\wedge AVAA) \qquad \text{111-9}$$

Once again the term, $V_{X(}X$, need not be determined explicitly; we require only an additional n_{x} perturbations in the decision variables.

Next we consider the procedure for constructing the matrix term, H. It can be seen that H can derived from:

$$\begin{bmatrix} AX\ V\ f\ V_{XX}L\ V\ ,\ H \\ A\!/1IV^{\ L}V\ \wedge \end{bmatrix}\ I\begin{bmatrix} AX \\ \Delta y \end{bmatrix}$$

by introducing the Δy perturbations in accordance with the relation, $_{LU}$ = - (V *V V^ AX. from Eq. III-6. Substituting this result and expanding the above product term gives

$$\begin{bmatrix} AX_{,} rr\ V_{IX}L\ v\ ,\ f\ i_{1} \\ |MAx]|V_{yx}L\ V^{\wedge}\ J \end{bmatrix}\ \begin{bmatrix} \Delta x \\ i\ M\Delta x \end{bmatrix} = \Delta x^{T}\ H\ AX$$

where

$$M = -(\nabla_{y}h^{T})^{-1}\ \nabla_{x}h$$

The term $\Delta\psi$ (for • - V,L, $\nabla\ \wedge\ g_{\wedge}$, h) is computed by perturbing the parameters and evaluating the respective differentials between 0. P— - '»- case function values (cf. Eq. 10).

APPENDIX IV

Mixed Integer Nonlinear Program Formulation for Active Set Selection

A change in the optimal solution resulting from a model change may be accompanied by a change in the active set The analytical example (Example 3) serves to illustrate this point While considering the problem of choosing a consistent active set in order to define a Newton step from the base optimal point to the modified optimum previously, we identified two cases: one in which all $\hat{a} = u_j$ and the other in which all $\hat{a} = 0$.

The choice of all $\hat{a} = 0$ implies that we have a starting point for the Newton step that is farther away from choosing some $\hat{a}_j = a_v$ to determine a consistent active set However, we can rigorously formulate the problem of finding the closest starting point in the following manner

$$Max \quad \sum \hat{u}.$$

$$\begin{aligned} & V_2Z.(r,o.v) + VJ/CTA\hat{V},) \; Az = 0 \\ & g\{T) * Vg^r(T) \; A2 * 0 \\ & hiT) + V/7^T(D \; Az = 0 \qquad (1) \\ & ti = \hat{a} + Aa \wedge 0 \; ; \; v = Vj + Av \\ & 0 \pounds \hat{u} \pounds u, \; y \end{aligned}$$

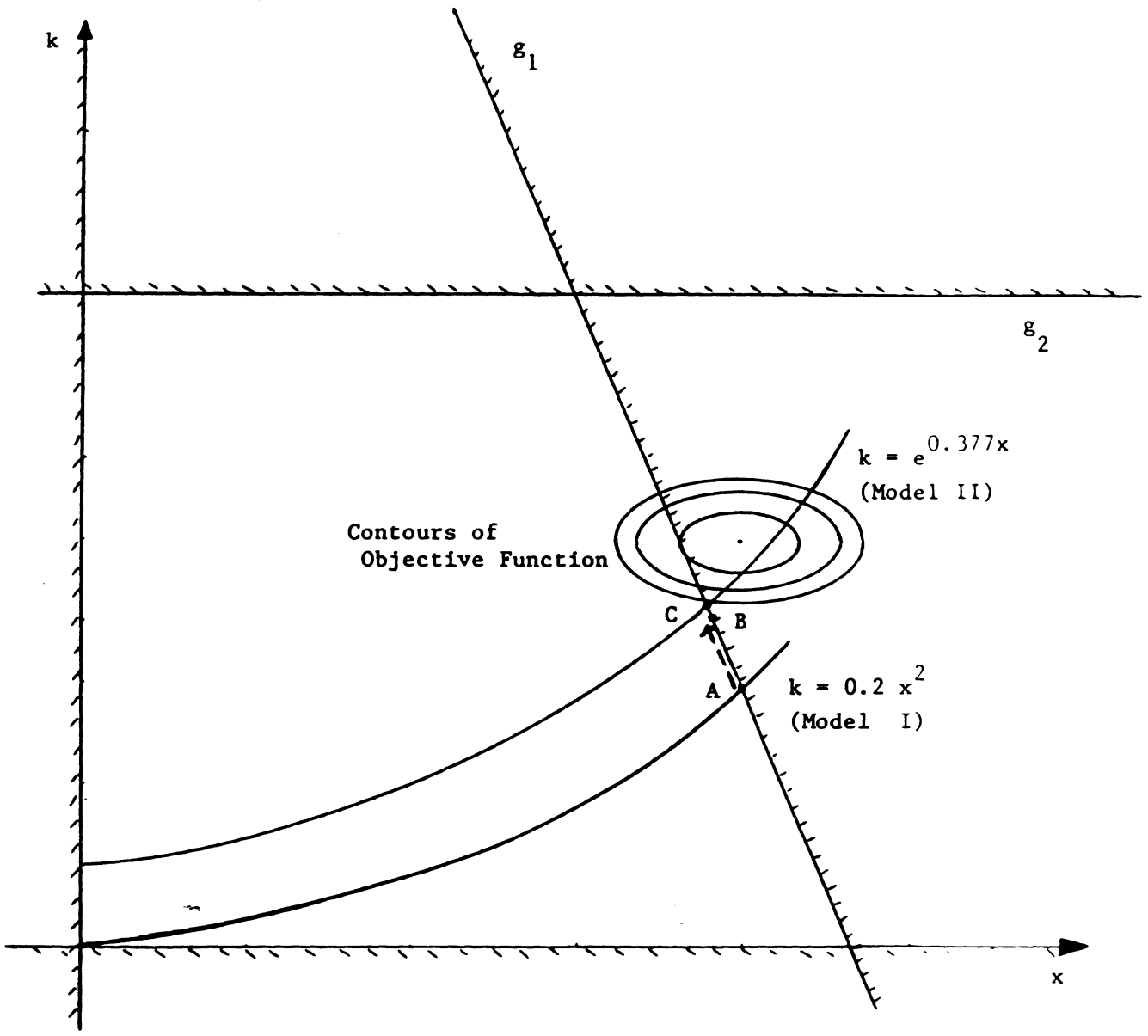
and for a large U

$$\begin{aligned} & u_j * U \; y \\ & gfj) \cdot V\wedge.^T_j(r) \; AZ \; z \; U \; (y - 1) \end{aligned}$$

The formulation leads to a Mixed Integer Nonlinear Programming (MINLP) problem. It can be seen that the constraints are linear in y and Az. A sufficient condition for a solution is that the QP is solvable with $\hat{a} = 0$. This is also a feasible lower bound. The upper bound occurs if $\hat{u} = u_t$ but this may not be feasible. This problem can be solved by

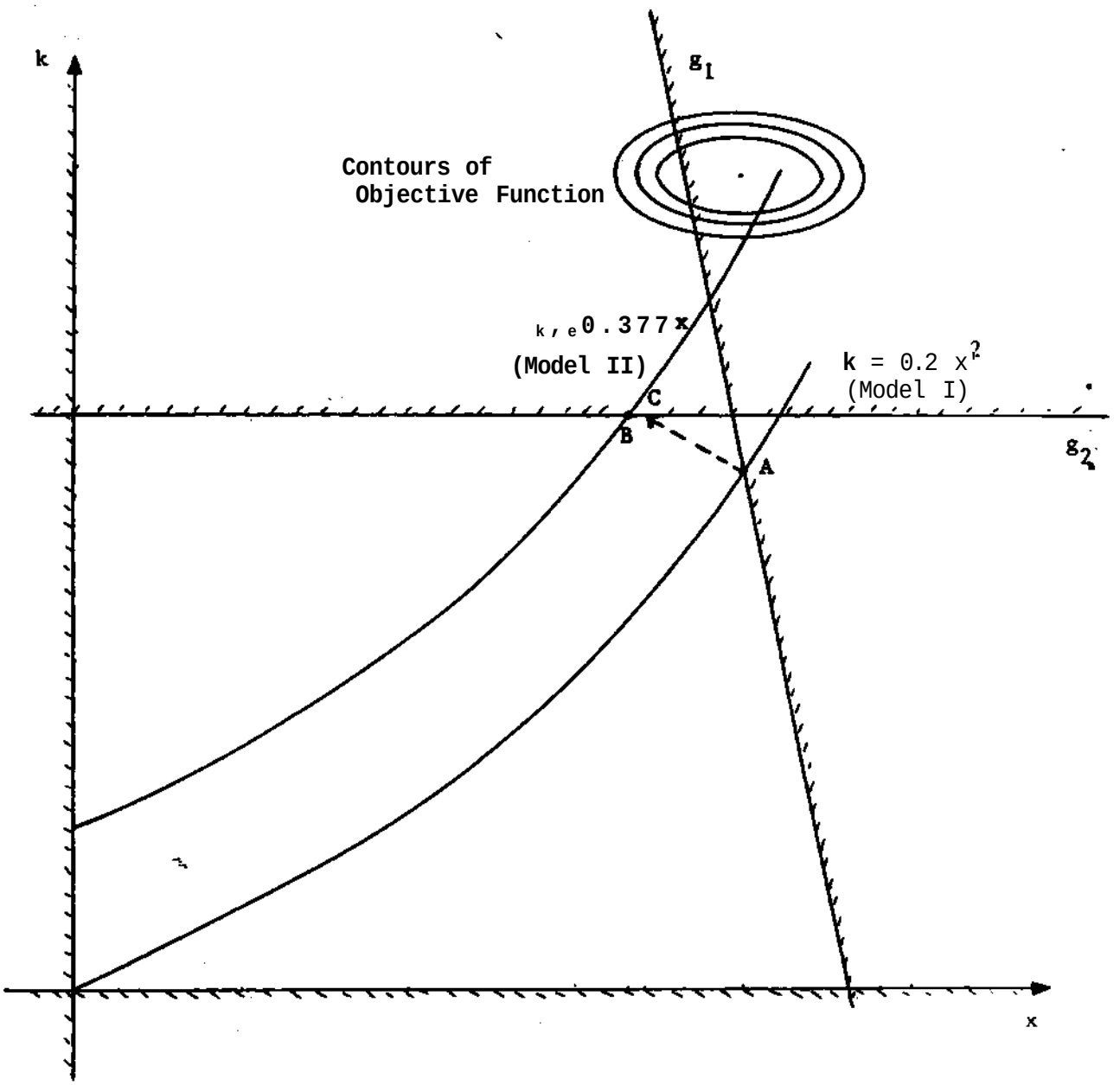
applying the Outer Approximation Approach (Duran & Grossmann, 1983). It is obvious that such as formulation gives rise to a combinatorial problem, although if the constraints are linear as with active bounds then this problem becomes nonexistent. However when nonlinear inequality constraints exist then the size of the inequalities plays an important role in the solution of the MINLP formulation. If the number of *nonlinear* inequality constraints happens to be equal to n_n, then the problem of determining the closest starting point to obtain a consistent active set may require up to 2^n discrete decisions.

Figure 1.

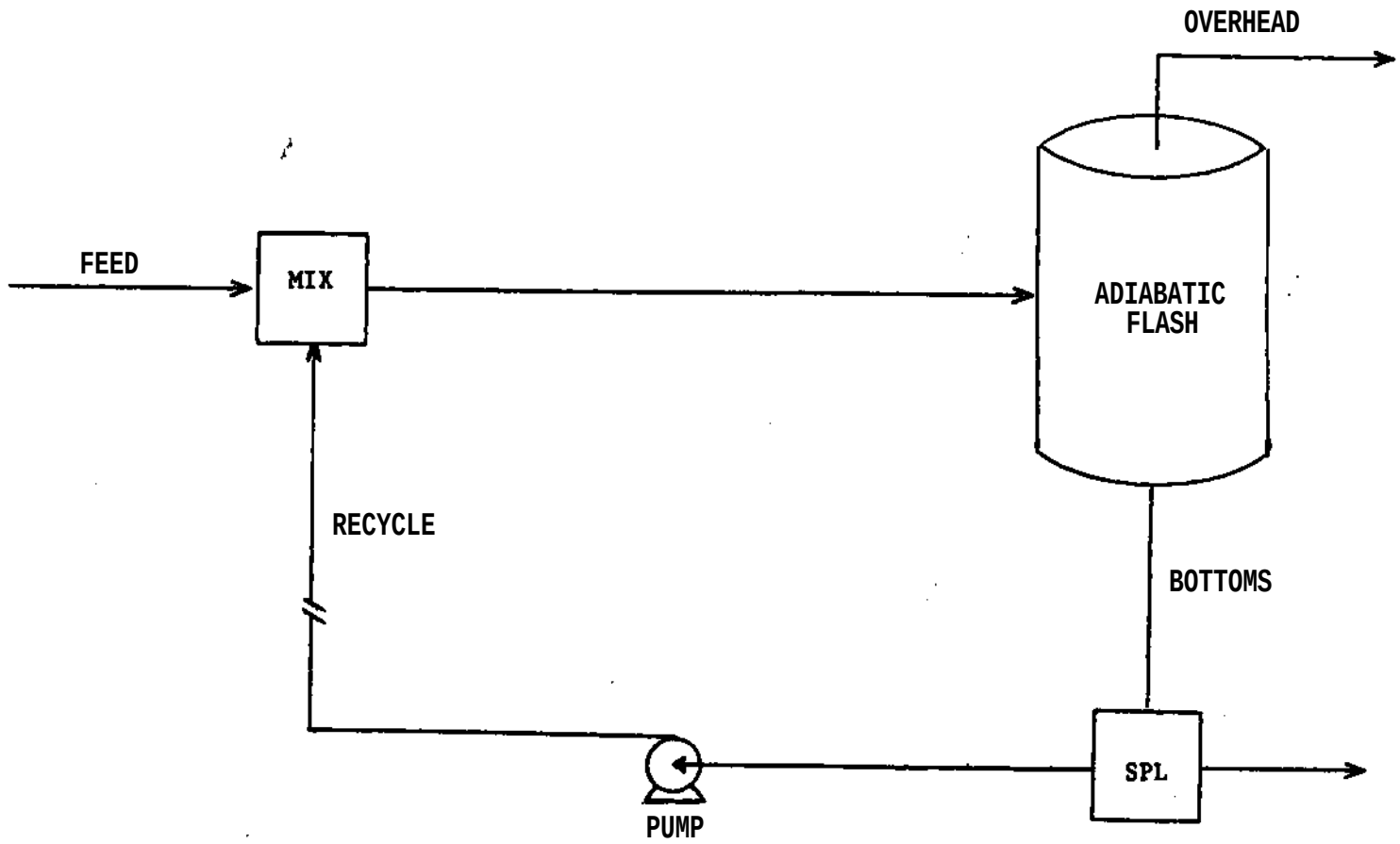


A - Optimum for Model I
B - Optimum for Model II
C - Predicted optimum for Model II from sensitivity analysis

Figure 2,



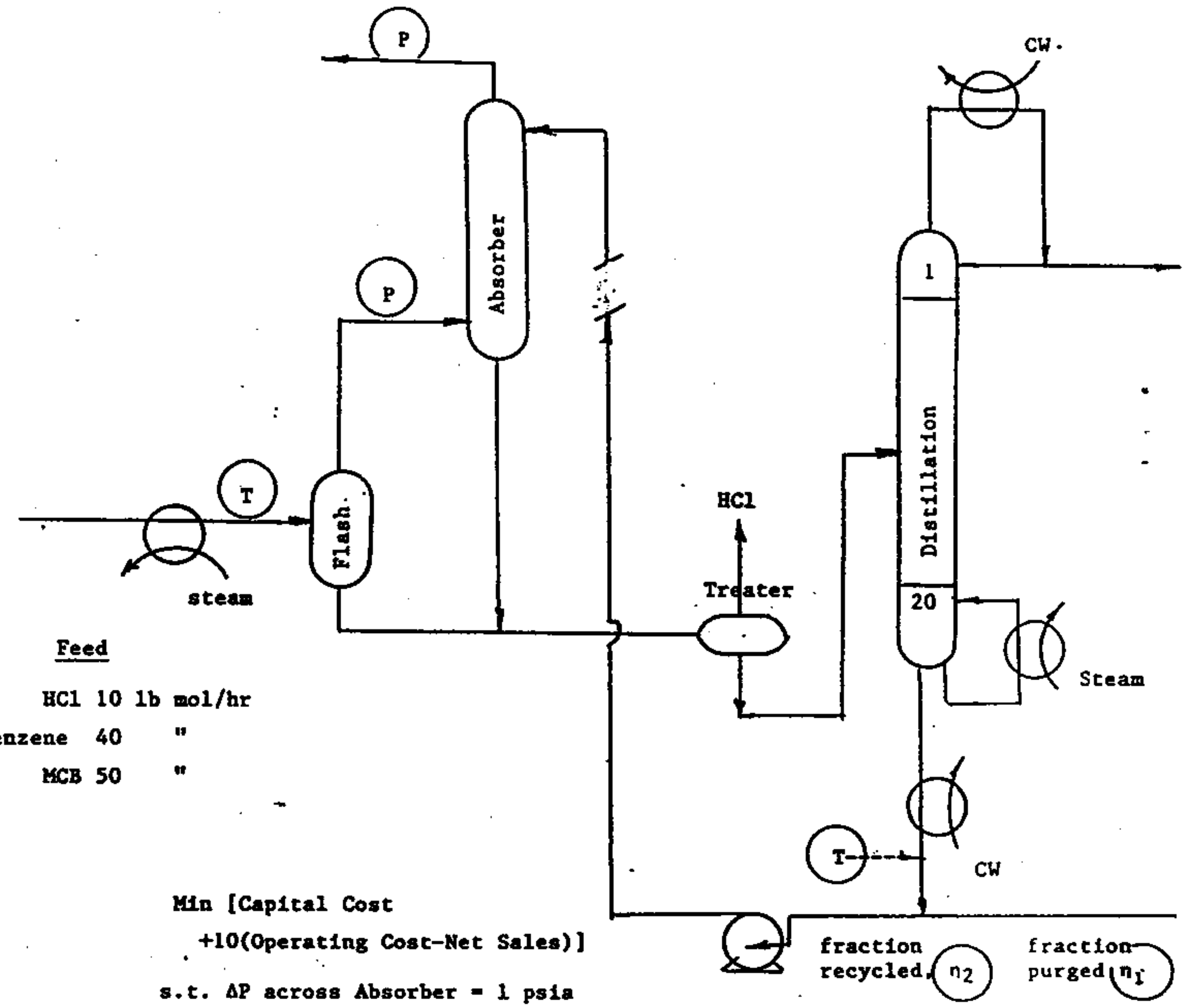
A - Optimum for Model I
B - Optimum for Model II
C - Predicted optimum for Model II from sensitivity analysis



SIMPLE FLASH RECYCLE FLOWSHEET

Figure 3.

Mono[illegible] lo[illegible] Opti[illegible]



Feed

HCl 10 lb mol/hr

Benzene 40 "

MCB 50 "

Min [Capital Cost
+10(Operating Cost-Net Sales)]

s.t. ΔP across Absorber = 1 psia

$\eta_1 + \eta_2 = 1$

Table I

Flash Recycle Flowsheet Optimization : Problem Definition

FEED **DATA**

COMPONENT	FEED FLOWRATE (mols/hr)
Propane	**10.0**
1-Butene	**15.0**
N-Butane	**20.0**
Trans-2-Butene	**20.0**
Cis-2-Butene	**15.0**
N-Pentane	**10.0**

Feed Pressure : 150 psia

Feed Temperature : 100 F

Decision Variables : Splitter Ratio, x_1 **(** $0.2 \leq x_1 \leq 0.8$ **)**

Flash Pressure, x_2 **(** $10 \leq x_2 \leq 50$ **) psia**

Tear Variables : **Component flows in Recycle,** y_i **, i = 1....6**

($0 \leq y_i \leq 100$ **) mols/hr**

Specific Enthalpy of Recycle, H

(-10000 ≤ H ≤ 10000) Btu/mol

OBJECTIVE FUNCTION

Monotonic : max e_1

Nonlinear : **max** $e_1 e_2 - e_1^2 - e_3^2 \cdot e_4 - e_5^{0.5}$

(e_i **- component flow in Flash Overhead) (Fig. 3)**

Table II

Flash Recycle Flowsheet : Sensitivity Analysis

PARAMETRIC VARIATION

COMPONENT	BASE CASE FEED (mols/hr)	PERTURBED FEED (mols/hr)
Propane	10.0	11.0
1-Butene	15.0	16.5
N-Butane	20.0	22.0
T-2-Butene	20.0	22.0
C-2-Butene	15.0	16.5
N-Pentane	10.0	11.0

MONOTONIC OBJECTIVE FUNCTION

DECISION VARIABLE	KKT MULTPR	BASE OPT	CORRECTION	PRED. OPT.	TRUE OPT.
x_1	-	0.2	0.0	0.2	0.2
x_2	-	10.0	0.0	10.0	10.0
-	u_1	1.005	- 0.45	0.55	1.652
-	u_2	0.146	0.51	0.656	0.21

NONLINEAR OBJECTIVE FUNCTION

DECISION VARIABLE	KKT MULTPR	BASE OPT	CORRECTION	PRED. OPT.	TRUE OPT.
x_1	-	0.8	0.0	0.8	0.8
x_2	-	23.785	- 0.475	23.31	23.09
-	u_1	1.654	1.877	3.53	2.56

Table III

Flash Recycle Flowsheet : Sensitivity Analysis

MODEL VARIATION

Monotonic Objective Function

Base Model : Raoult's Law (Model I)
Alt. Model : Chao-Seader (Model II)

DECISION VARIABLE	KKT MULTPR	BASE OPT	CORRECTION	PRED. OPT.	TRUE OPT.
x_1	-	0.2	0.0	0.2	0.2
x_2	-	10.0	0.0	10.0	10.0
-	u_1	1.005	1.065	2.07	1.35
-	u_2	0.146	0.014	0.16	0.17

Base Model : Chao-Seader (Model I)
Alt. Model : Raoult's Law (Model II)

DECISION VARIABLE	KKT MULTPR	BASE OPT	CORRECTION	PRED. OPT.	TRUE OPT.
x_1	-	0.2	0.0	0.2	0.2
x_2	-	10.0	0.0	10.0	10.0
-	u_1	1.35	- 0.23	1.12	1.005
-	u_2	0.17	- 0.024	0.144	0.146

contd..

Table III

Flash Recycle Flowsheet : Sensitivity Analysis

MODEL VARIATION

Nonlinear Objective Function

Base Model : Raoult'a Law (Model I)
Alt. Model : Chao-Seader (Model II)

DECISION VARIABLE	KKT MULTPR	BASE OPT	CORRECTION	PRED. OPT.	TRUE OPT.
x_1	-	0.8	- 7.55E-03	0.793	0.8
x_2	-	23.78	- 0.3107	23.47	23.03
-	u_1	1.65	-	0.0	1.43

Base Model : Chao-Seader (Model I)
Alt. Model : Raoult's Law (Model II)

DECISION VARIABLE	KKT MULTPR	BASE OPT	CORRECTION	PRED. OPT.	TRUE OPT.
x_1	-	0.8	0.0	0.8	0.8
x_2	-	23.03	1.28	24.31	23.79
-	u_1	1.43	- 0.95	0.48	1.65

Table IV.

MONOCHLOROBENZENE SEPARATION FLOWSHEET

FEED DATA

COMPONENT	FLOW RATES (mols/hr)
HCl	10
Benzene	40
MCB	50

Feed Pressure : 37 psia

Feed Temperature : 80 °F

PHYSICAL PROPERTY MODELS

Vapor Pressure Model : Cavett Equation

Vapor Fugacity Model : Redlich-Kwong Equation

Liquid Fugacity Model : Redlich-Kwong and Poynting Equations

Liquid Activity Coefficient Model: Ideal Solution

DECISION VARIABLES

Abscrber Pressure (Bottom)	:	$25 \leq x_1 \leq 35$	psia
Absorber Pressure (Top)	:	$25 \leq x_2 \leq 35$	psia
Split Fraction to Recycle	:	$0 \leq x_3 \leq 1$	
Split Fraction to Outside	:	$0 \leq x_4 \leq 1$	
Flash Input Stream Temp.	:	$200 \leq x_5 \leq 290$	°F
Absorber Input Stream Temp. (Recycle)	:	$100 \leq x_6 \leq 300$	°F

TEAR VARIABLES

Recycle HCl Flow

Recycle Benzene Flow

Recycle MCB Flow

Recycle Pressure

Recycle Temperature