A Reduced Hessian Strategy for Sensitivity Analysis of Optimal Flowsheets

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.

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

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, and the like, 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 and 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 under normal conditions.

In the past, the parametric sensitivity problem has been addressed for simulation by several researchers. Bard (1974) and Atherton et al. (1975) proposed methods to determine sensitivity coefficients to measure the influence of uncertainties in model parameters on the solutions obtained from incorporating a particular ODE model. In the context of flowsheet simulation, Volin and Ostrovskiy (1981) developed an approach based on setting up and solving an adjoint flowsheet system to the nominal problem. A more direct approach to determine the sensitivities of flowsheet variables to parametric variations utilizing the block Jacobians of unit modules was presented recently by 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; Biegler, 1985).

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

Correspondence concerning this paper should be addressed to L. T. Biegler.

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 moderate-size flowsheet with a combined $(x + y)$ dimensionality in the range of 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, which are 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 information on all optimal variables but the equality constraint multipliers.

We then develop the theory for analyzing the sensitivity of the optimum to model variations. We formulate this problem 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 process 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 procedure that identifies the modified active set by solving a set of quadratic programs iteratively. We present conditions under which this procedure is rigorous and formulate a mixed-integer nonlinear program (MINLP) to deal with the more general cases.

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:

$$\begin{aligned} \boldsymbol{P}(p): \quad & \min_z \quad \phi(z,p) \\ & \text{s.t.} \quad g(z,p) \le 0 \\ & \qquad\;\; h(z,p) = 0 \end{aligned} \tag{1}$$

where

p = input parameter

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

ϕ = objective function

g = design inequality constraint

h = equality (tear) constraint

The input parameter p may be uncertain or may vary over a known range. Decision variables x are usually adjusted by the designer and can include equipment sizes or temperatures and pressures in the flowsheet. Dependent variables y can be computed from equality constraints once the decision variables are specified. These include the flow rates, 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 Karush-Kuhn-Tucker (KKT) conditions are satisfied we have the following relations:

$$\begin{aligned} \nabla\phi(z^o,p^o) + \nabla g(z^o,p^o)u^o + \nabla h(z^o,p^o)v^o = 0 \\ u^{o^T} g(z^o,p^o) = 0; \quad u^o \ge 0; \quad g(z^o,p^o) \le 0 \\ h(z^o,p^o) = 0 \end{aligned} \tag{2}$$

where z^o = base case optimal solution, $\{x^o, y^o\}$, and u^o, v^o = KKT multipliers at base case optimum.

This result may be interpreted to mean that for the supplied input parameters p^o, the variable vector z^o is a local minimum of $\boldsymbol{P}(p^o)$ with the corresponding KKT multipliers u^o, v^o. In the context of flowsheet optimization the input parameter vector p^o can include a subset of internal process parameters (e.g., kinetic rate constant terms) and externally supplied parameters (e.g., feed flow rates) 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^o the following conditions are satisfied:

1. The functions defining $\boldsymbol{P}(p)$ are at least twice continuously differentiable in z and at least once in p for a neighborhood of (z^o, p^o).
2. The constraint gradients are linearly independent at z^o and strict complementary slackness holds for $\boldsymbol{P}(p^o)$ at z^o with unique KKT multipliers u^o and v^o.
3. The second-order sufficiency conditions are met (see Appendix A).

These assumptions are sufficiently general for most process optimization problems. Relaxation of these assumptions for sensitivity analysis is discussed by Fiacco (1983), but only limited results are available without these assumptions. From the KKT conditions at the optimum, z^o, we have:

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

where L is the Lagrange function and g_A is the active inequality constraint.

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

$$\begin{aligned} \mathrm{d}[\nabla_z \mathrm{L}(z^o,p^o)] &= \nabla_{zz}\mathrm{L}^o \mathrm{d}z + \nabla_z g_A^o \mathrm{d}u \\ &\quad + \nabla_z h^o \mathrm{d}v + \nabla_{pz}\mathrm{L}^{To}\mathrm{d}p = 0 \\ \mathrm{d}g_A(z^o,p^o) &= \nabla_z g_A^{To}\mathrm{d}z + \nabla_p g_A^{To}\mathrm{d}p = 0 \\ \mathrm{d}h(z^o,p^o) &= \nabla_z h^{To}\mathrm{d}z + \nabla_p h^{To}\mathrm{d}p = 0 \end{aligned} \tag{4}$$

Rearranging these expressions results in the linear system of equations:

$$\begin{bmatrix} \nabla_{pz}L^{To} \\ \nabla_p g^{To} \\ \nabla_p h^{To} \end{bmatrix} = -\begin{bmatrix} \nabla_{zz}L^o & \nabla_z g_A^o & \nabla_z h^o \\ \nabla_z g_A^{To} & 0 & 0 \\ \nabla_z h^{To} & 0 & 0 \end{bmatrix} \begin{bmatrix} \nabla_p z^{To} \\ \nabla_p u^{To} \\ \nabla_p v^{To} \end{bmatrix} \quad (5)$$

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

$$\begin{bmatrix} \nabla_{px}L^{To} \\ \nabla_{py}L^{To} \\ \nabla_p g_A^{To} \\ \nabla_p h^{To} \end{bmatrix} = -\begin{bmatrix} \nabla_{xx}L^o & \nabla_{xy}L^o & \nabla_x g_A^o & \nabla_x h^o \\ \nabla_{yx}L^o & \nabla_{yy}L^o & \nabla_y g_A^o & \nabla_y h^o \\ \nabla_x g_A^{To} & \nabla_y g_A^{To} & 0 & 0 \\ \nabla_x h^{To} & \nabla_y h^{To} & 0 & 0 \end{bmatrix} \begin{bmatrix} \nabla_p x^{To} \\ \nabla_p y^{To} \\ \nabla_p u^{To} \\ \nabla_p v^{To} \end{bmatrix} \quad (6)$$

If strict complementary slackness holds, the active set identified for the base optimal solution is retained within an ϵ neighborhood of the nominal parameters (Fiacco, 1976). The derivative matrix on the righthand 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 optimal variable vector corresponding to a change Δp in the parameters p. Therefore, in the neighborhood of the base case optimum we have (for $\omega = x, y, u, v$):

$$\tilde{\omega} = \omega^o + \nabla_p \omega^{To} \Delta p = \omega^o + \Delta\omega \quad (7)$$

where $\tilde{\omega}$ is the 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 $\psi = \nabla_x L, \nabla_y L, g_A, h$):

$$\frac{\partial \psi^o}{\partial p_k} = \frac{\psi(p_k^o + \Delta p_k) - \psi(p^o)}{\Delta p_k} \quad (8)$$

The scalar, Δp_k, in Eq. 8 corresponds to perturbation of the kth element of the input parameter vector p. Alternatively, a response to a specified change in a linear combination of the parameters Δp can be made by computing the directional derivatives for x, y, u, and v. From Eqs. 6 and Eq. 7, we have:

$$\nabla^2 L^o \Delta\omega = \nabla^2 L^o \nabla_p \omega^{To} \Delta p = -\nabla_p(\nabla L)^T \Delta^o p = -\Delta(\nabla L^o) \quad (9)$$

where

$$\nabla L^o \equiv [\nabla_x L \; \nabla_y L \; g_A \; h]^T$$

The last term on the righthand side of Eq. 9 can be calculated by the finite-difference approximation along the direction Δp:

$$\Delta(\nabla L^o) = \frac{\nabla L(p^o + \epsilon\Delta p) - \nabla L(p^o)}{\epsilon} \quad (10)$$

Of course, we note that the first-order sensitivities are necessarily accurate only for a small $\epsilon\Delta p$. 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 $\nabla_x\psi$, $\nabla_y\psi$, (for $\psi = \phi, g, h$). In addition we require the Hessian matrix $\boldsymbol{B}$, given by:

$$\boldsymbol{B} = \begin{bmatrix} \nabla_{xx}L^o & \nabla_{xy}L^o \\ \nabla_{yx}L^o & \nabla_{yy}L^o \end{bmatrix} \quad (11)$$

Unless the second derivatives are inexpensive to calculate and an actual constrained Newton method is used, the $\boldsymbol{B}$ 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 $\boldsymbol{B}$ are inappropriate for sensitivity analysis. A justification for requiring the exact $\boldsymbol{B}$ matrix is given in Appendix B.

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 space, a smaller matrix is constructed by perturbing x and y simultaneously so that the linearizations of the equality constraints are always satisfied.

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

$$\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \\ \boldsymbol{f} \\ \boldsymbol{e} \end{bmatrix} = -\begin{bmatrix} \boldsymbol{I} & 0 & \boldsymbol{E} & 0 \\ 0 & \boldsymbol{I} & \boldsymbol{L} & \boldsymbol{M} \\ 0 & 0 & \boldsymbol{H} & \boldsymbol{Q} \\ 0 & 0 & \boldsymbol{Q}^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y^o \\ \Delta v^o \\ \Delta x^o \\ \Delta u^o \end{bmatrix} \quad (12)$$

where

$\boldsymbol{I}$ = identity matrix

$$\boldsymbol{E} = -\boldsymbol{A} = (\nabla_y h^{To})^{-1} \nabla_x h^{To}$$

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

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

$$\boldsymbol{Q} = \nabla_x g_A^o - \nabla_x h^o (\nabla_y h^o)^{-1} \nabla_y g_A^o$$

$$\boldsymbol{H} = \nabla_{xx} L^o - \nabla_{xy} L^o (\nabla_y h^{To})^{-1} \nabla_x h^{To} - \nabla_x h^o (\nabla_y h^o)^{-1} (\nabla_{yx} L^o + \nabla_{yy} L^o \boldsymbol{A})$$

$$\boldsymbol{a} = (\nabla_y h^{To})^{-1} \Delta h^o$$

$$\boldsymbol{e} = \Delta g^o - \nabla_y g_A^{To} (\nabla_y h^{To})^{-1} \Delta h^o$$

$$b = (\nabla_y h^o)^{-1}\Delta(\nabla_y L^o) - (\nabla_y h^o)^{-1}\nabla_{yy}L^o(\nabla_y h^{To})^{-1}\Delta h^o$$

$$f = \Delta(\nabla_x L^o) - \nabla_{xy}L^o(\nabla_y h^{To})^{-1}\Delta h^o - \nabla_x h^o(\nabla_y h^o)^{-1} \cdot \Delta(\nabla_y L^o) + \nabla_x h^o(\nabla_y h^o)^{-1}\nabla_{yy}L^o(\nabla_y h^{To})^{-1}\Delta h^o$$

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

$$\begin{bmatrix} f \\ e \end{bmatrix} = -\begin{bmatrix} H & Q \\ Q^T & 0 \end{bmatrix}\begin{bmatrix} \Delta x^o \\ \Delta u^o \end{bmatrix} \tag{13}$$

The reduced matrix, H can be constructed by simultaneously perturbing the decision and tear variables accordingly as $[\Delta x, -(\nabla_y h^T)^{-1}\nabla_x h^t \Delta x]$ (see Appendix C for details).

The number of flowsheet evaluations, NFE_1, required for the reduced form in Eq. 13 is given by the sum:

$$NFE_1 = [n_x(n_x + 1)/2] + 3n_x + 2 \tag{14}$$

where n_x is the 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:

$$NFE_2 = [(n_x + n_y)(n_x + n_y + 1)/2] + 1 \tag{15}$$

where n_y is the number of dependent (tear) variables.

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

$$NFE_2 - NFE_1 = \tfrac{1}{2}[n_y^2 + 2n_x(n_y - 1) + n_y] - 2n_x - 1 \tag{16}$$

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_y. 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 u are readily obtained; those for y are backed out from $E\Delta x$ and a. The only information not obtained from Eq. 12 are sensitivities for v. Normally sensitivities for equality constraint multipliers are not required as often as those for state variables.

Example 1: Parametric sensitivity problem

To illustrate the reduced Hessian sensitivity approach, we first consider a small analytical example. Consider the following minimization problem:

$$\min_{x_1, y_1, y_2} \quad \phi: x_1^2 + y_1^2 + y_2^2$$

$$\text{s.t.} \quad h_1 = 6x_1 + 3y_1 + 2y_2 - p_1 = 0$$

$$h_2 = p_2 x_1 + y_1 - y_2 - 1 = 0$$

The problem includes one independent variable, x_1, and two dependent variables, y_1 and y_2. The terms p_1 and p_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 $p_1^o = 6.0$ and $p_2^o = 1.0$.

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 optimal solution is $(x_1^o, y_1^o, y_2^o) \equiv (0.7449, 0.4082, 0.1531)$ with the associated KKT multipliers $(v_1^o, v_2^o) \equiv (-0.2245, -0.1429)$. Also we can evaluate the following terms:

$$\nabla_x h^o \equiv [6 \quad p_2^o] \equiv [6 \quad 1]$$

$$\nabla_{y_1} h^o \equiv [3 \quad 1]$$

$$\nabla_{y_2} h^o \equiv [2 \quad -1]$$

To solve for the perturbed optimum due to a perturbation Δp_1 and Δp_2 in the parameters p_1 and p_2 about their nominal values, we use the reduced Hessian strategy. Using this approach (see Appendix C for details) the linear system to determine the deviations in the decision space becomes:

$$H\Delta x_1 = -f$$

where

$$H \equiv [7.84]$$

$$f \equiv [\Delta p_2 v_2^o - 0.88\Delta p_1 + 0.56\Delta p_2 x_1^o]$$

If we assume $\Delta p_1 = 0.1$ and $\Delta p_2 = 0.05$, the corresponding deviation 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: $\tilde{x}_1 = x_1^o + \Delta x_1 = 0.7544$.

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

$$\tilde{y}_1 = y_1^o + \Delta y_1 = 0.4082 - 0.0101 = 0.3981$$

$$\tilde{y}_2 = y_2^o + \Delta y_2 = 0.1531 + 0.0367 = 0.1898$$

Therefore the perturbed optimum with $p_1 = 6.1$ and $p_2 = 1.05$ is found to be at $(\tilde{x}_1, \tilde{y}_1, \tilde{y}_2) \equiv (0.7544, 0.3981, 0.1898)$.

The actual 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 et al., 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. In general, such local simplified models derive from a common rigorous model and often resemble each other closely in structure and function (Bryan and Grens, 1983; Chimowitz et al., 1983, 1984). This is especially helpful for physical property calculations. For process optimization the use of such competing approximate local models can lead to locally different solutions even though solutions of simulation problems may be the same (Biegler et al., 1985). 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 section 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 the previous section, in particular, the reduced Hessian procedure.

Consider the model-based optimization problem:

$$\begin{aligned} \min_{x,y,k} \quad & \phi(x, y, k) \\ \text{s.t.} \quad & g(x, y, k) \leq 0 \\ & h(x, y, k) = 0 \\ & f(x, y, k) = k - K(x, y) = 0 \end{aligned} \tag{17}$$

where

ϕ = objective function

g = inequality constraint

h = equality (tear) constraint

k = physical (model based) property

K = model for property evaluation

The property k in Eq. 17 is estimated by using model I that is given by $k = K(x, y)$. At the optimum with respect to model I, $(x_1, y_1, u_1, v_1, k_1)$, from the KKT conditions we have:

$$\begin{aligned} \nabla_x \mathrm{L}(x_1, y_1, k_1) &= 0 \\ \nabla_y \mathrm{L}(x_1, y_1, k_1) &= 0 \\ \nabla_k \mathrm{L}(x_1, y_1, k_1) &= 0 \\ g_A(x_1, y_1, k_1) &= 0 \\ h(x_1, y_1, k_1) &= 0 \end{aligned} \tag{18}$$

where L is the Lagrange function and g_A is the 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:

$$\begin{aligned} \mathrm{d}[\nabla_x \mathrm{L}(x_1, y_1, k_1)] &= \nabla_{xx}\mathrm{L}_1 \mathrm{d}x + \nabla_{xy}\mathrm{L}_1 \mathrm{d}y + \nabla_x g_{A1} \mathrm{d}u \\ &\quad + \nabla_x h_1 \mathrm{d}v + \nabla_{xk}\mathrm{L}_1 \mathrm{d}k = 0 \\ \mathrm{d}[\nabla_y \mathrm{L}(x_1, y_1, k_1)] &= \nabla_{yx}\mathrm{L}_1 \mathrm{d}x + \nabla_{yy}\mathrm{L}_1 \mathrm{d}y + \nabla_y g_{A1} \mathrm{d}u \\ &\quad + \nabla_y h_1 \mathrm{d}v + \nabla_{yk}\mathrm{L}_1 \mathrm{d}k = 0 \\ \mathrm{d}[\nabla_k \mathrm{L}(x_1, y_1, k_1)] &= \nabla_{kx}\mathrm{L}_1 \mathrm{d}x + \nabla_{ky}\mathrm{L}_1 \mathrm{d}y + \nabla_k g_{A1} \mathrm{d}u \\ &\quad + \nabla_k h_1 \mathrm{d}v + \nabla_{kk}\mathrm{L}_1 \mathrm{d}k = 0 \\ \mathrm{d}g_A(x_1, y_1, k_1) &= \nabla_x g_{A1}^T \mathrm{d}x + \nabla_y g_{A1}^T \mathrm{d}y + \nabla_k g_{A1}^T \mathrm{d}k = 0 \\ \mathrm{d}h(x_1, y_1, k_1) &= \nabla_x h_1^T \mathrm{d}x + \nabla_y h_1^T \mathrm{d}y + \nabla_k h_1^T \mathrm{d}k = 0 \\ \mathrm{d}f(x_1, y_1, k_1) &= -\nabla_x K^T \mathrm{d}x - \nabla_y K^T \mathrm{d}y + I \mathrm{d}k = 0 \end{aligned} \tag{19}$$

If the competing model (model II) is given by $k = \tilde{K}(x, y)$, 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. Here the sensitivity of x, y, and k, to the new model can be calculated by substituting $\tilde{K}$ for K in the last relation of Eq. 19. On the other hand, since the property evaluated by the model, k, is essentially a function of the decision and tear variables, this system can first be 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 stagewise unit operations. The equivalent problem can be formulated as:

$$\begin{aligned} \min_{x,y} \quad & \phi[x, y, K(x, y)] \\ \text{s.t.} \quad & g[x, y, K(x, y)] \leq 0 \\ & h[x, y, K(x, y)] = 0 \end{aligned} \tag{20}$$

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, $k = \tilde{K}(x, y)$, i.e.,

$$\begin{bmatrix} \nabla_x L[x_1, y_1, \tilde{K}(x_1, y_1)] \\ \nabla_y L[x_1, y_1, \tilde{K}(x_1, y_1)] \\ g_A[x_1, y_1, \tilde{K}(x_1, y_1)] \\ h[x_1, y_1, \tilde{K}(x_1, y_1)] \end{bmatrix} \neq 0 \tag{21}$$

where L, g, and h are respectively the Lagrange function, inequality constraint, and equality constraint defined with respect to model II.

Let us assume that the optimum with respect to model II is at $\tilde{z} = (x_2, y_2)$. If we choose a consistent active set (we discuss this point later) we can write a first-order correction for the optimality conditions with respect to model II. Defining a Newton step in the space of model II for x, y, u, and v leads to:

$$\begin{bmatrix} \nabla_x L_2 \\ \nabla_y L_2 \\ g_{A2} \\ h_2 \end{bmatrix} \simeq \begin{bmatrix} \nabla_x L_1 \\ \nabla_y L_1 \\ g_{A1} \\ h_1 \end{bmatrix} + \begin{bmatrix} \nabla_{xx} L_1 & \nabla_{xy} L_1 & \nabla_x g_{A1} & \nabla_x h_1 \\ \nabla_{yx} L_1 & \nabla_{yy} L_1 & \nabla_y g_{A1} & \nabla_y h_1 \\ \nabla_x g_{A1}^T & \nabla_x g_{A1}^T & 0 & 0 \\ \nabla_x h_1^T & \nabla_y h_1^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta u \\ \Delta v \end{bmatrix} = 0 \tag{22}$$

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

$$\begin{bmatrix} \nabla_x L_1 \\ \nabla_y L_1 \\ g_{A1} \\ h_1 \end{bmatrix} = - \begin{bmatrix} \nabla_{xx} L_1 & \nabla_{xy} L_1 & \nabla_x g_{A1} & \nabla_x h_1 \\ \nabla_{yx} L_1 & \nabla_{yy} L_1 & \nabla_y g_{A1} & \nabla_y h_1 \\ \nabla_x g_{A1}^T & \nabla_y g_{A1}^T & 0 & 0 \\ \nabla_x h_1^T & \nabla_y h_1^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta u \\ \Delta v \end{bmatrix} \tag{23}$$

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 II. Here, the KKT multiplier u_1, corresponding to an active inequality g_j, is positive provided g_j is active for model II and the Δu from Eq. 23 is also such that $u_1 + \Delta u$ remains positive. Again, the reduced Hessian strategy can be applied to the linear system in Eq. 23 so that the corrections for x and u are obtained first; the sensitivity for y can be computed using the Δx 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:

$$\begin{aligned} \min_{x,k} \quad & \phi\colon\ 4(x-4)^2 + 9(k-5)^2 \\ \text{s.t.} \quad & g_1\colon\ k + 5x - 23.2 \le 0 \\ & g_2\colon\ k - 8 \le 0 \\ & k \ge 0; \quad x \ge 0 \end{aligned}$$

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:

$$\text{Model I: } k - 0.2x^2 = 0$$
$$\text{Model II: } k - e^{0.3776x} = 0$$

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

$$\Delta x = -0.19813; \quad \Delta k = 0.9906$$
$$\Delta u^1 = -4.3727; \quad \Delta v = -13.4584$$

The estimated optimal solution with respect to model II is: $(x_2, k_2, u_2^1, v_2) \equiv (3.8019, 4.1906, 3.4818, 11.0871)$. If we carry out the optimization directly with respect to model II we find that the actual 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. Figure 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,

$$\begin{aligned} \nabla_{xx} L(\bar{z}, u_1, v_1)\Delta x + \nabla_{xy} L(\bar{z}, u_1, v_1)\Delta y + \nabla_x L(\bar{z}, u, v) &= 0 \\ \nabla_{yx} L(\bar{z}, u_1, v_1)\Delta x + \nabla_{yy} L(\bar{z}, u_1, v_1)\Delta y + \nabla_y L(\bar{z}, u, v) &= 0 \\ \nabla_x g^T(\bar{z})\Delta x + \nabla_y g^T(\bar{z})\Delta y + g(\bar{z}) &\le 0 \\ \nabla_x h^T(\bar{z})\Delta x + \nabla_y h^T(\bar{z})\Delta y + h(\bar{z}) &= 0 \\ u = u_1 + \Delta u \ge 0; \quad v = v_1 + \Delta v & \\ (u_1 + \Delta u)^T[\nabla_x g^T(\bar{z})\Delta x + \nabla_y g^T(\bar{z})\Delta y + g(\bar{z})] &= 0 \end{aligned} \tag{24}$$

which can be found from the following quadratic program (QP),

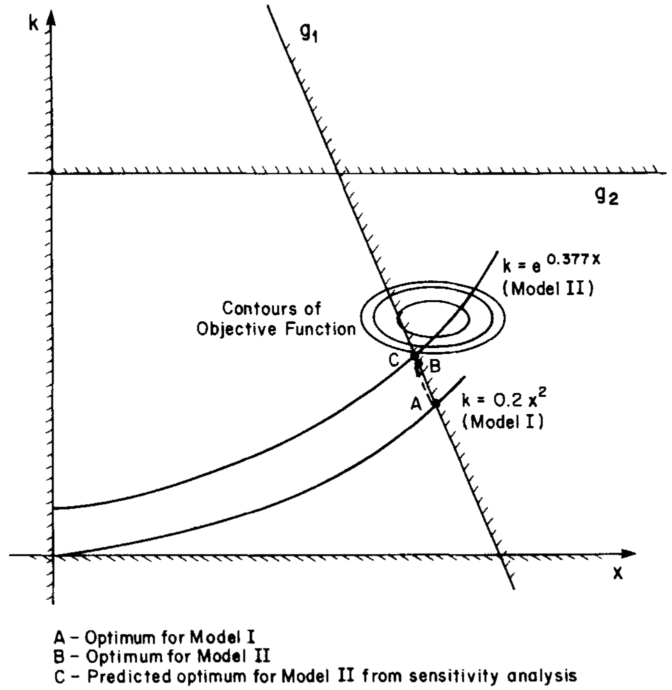


Figure 1. Analytic model sensitivity example.

using $z \equiv \{^x_y\}$:

$$\begin{aligned} \min \quad & \nabla\phi^T(\bar{z})\Delta z + \tfrac{1}{2}\Delta z^T \nabla_{zz} L(\bar{z}, u_1, v_1)\Delta z \\ \text{s.t.} \quad & g(\bar{z}) + \nabla g^T(\bar{z})\Delta z \le 0 \\ & h(\bar{z}) + \nabla h^T(\bar{z})\Delta z = 0 \\ & u = u_1 + \Delta u; \quad v = v_1 + \Delta v \end{aligned} \tag{25}$$

We also assume strict complementary slackness for the QP solution, i.e.,

$$u^T(g(\bar{z}) + \nabla g^T(\bar{z})\Delta z) = 0$$

implies

$$\begin{aligned} \text{for} \quad & g_j(\bar{z}) + \nabla g_j^T(\bar{z})\Delta z = 0, \quad u^j > 0 \\ \text{for} \quad & g_j(\bar{z}) + \nabla g_j^T(\bar{z})\Delta z < 0, \quad u^j = 0 \end{aligned} \tag{26}$$

Solution of the QP corresponding to Eq. 25 amounts to taking a Newton step for z, u, and v in the space of model II from the solution of model I, with the added restriction that u never becomes negative and the inequalities are satisfied to first order. Therefore the KKT conditions are correct to first order for this step.

We note, however, that the Lagrange function is nondifferentiable in z at the point where the active set changes. This occurs because u changes from a positive value to zero or vice versa. Clearly, this nonlinear feature may affect the accuracy of the sensitivities obtained from the Newton step. To remedy this, note that the nondifferentiability disappears if both u^j and u_1^j are positive (or zero) together. Under this condition we have a consistent active set, a smooth Lagrange function, and consequently more accurate sensitivity estimates from the system in Eq. 25. In fact, solutions from the Newton step corresponding to Eq. 23 and Eq. 25 would be identical.

An alternative way of dealing with this problem would be to substitute u_1, in general, by $\hat{u}$ in Eq. 25 such that $\hat{u} \ge 0$. Note that u_1 or $\hat{u}$ appears only in $\nabla_{zz}L$. First we rewrite the quadratic program, Eq. 25, as:

$$\begin{aligned} \min \quad & \nabla\phi^T(\bar{z})\Delta z + \tfrac{1}{2}\Delta z^T \nabla_{zz} L(\bar{z}, \hat{u}, v_1)\Delta z \\ \text{s.t.} \quad & g(\bar{z}) + \nabla g^T(\bar{z})\Delta z \le 0 \\ & h(\bar{z}) + \nabla h^T(\bar{z})\Delta z = 0 \end{aligned} \tag{27}$$

We now develop a way of determining $\hat{u}$ so that a consistent active set is obtained. To do this we consider the following cases:

- All g_j are linear. In this case, $\hat{u}$ does not appear in Eq. 27 because the inequalities do not contribute to the curvature of the Lagrange function in z. Therefore a consistent active set is obtained simply by solving Eq. 25.
- Some g_j are nonlinear. Here we propose the following algorithm to obtain a consistent active set.

1. Set $\hat{u} = u_1$.
2. Solve the quadratic program in Eq. 27 to obtain Δz, u, and v.
3. If both $\hat{u}^j$ and u^j are positive or zero, stop.
4. Otherwise, set $\hat{u} = u$ and go back to step 2.

We found this algorithm to work well on a number of examples, with convergence dependent on the magnitude of the contribution of $\hat{u}^j \nabla_{zz} g_j$ on $\nabla_{zz}L$ in Eq. 27. We admit, however, that in some cases the above algorithm may perform poorly and thus note this combinatorial problem can be posed formally as a mixed-integer nonlinear program (MINLP) as shown in Appendix D. Since this MINLP involves only bilinear functions in the continuous variables, it can be solved using a number of strategies such as the generalized benders decomposition or outer approximation (Duran and Grossmann, 1986). We repeat, however, that in many cases the solution can be obtained in a straightforward manner using the above algorithm.

To illustrate these concepts, we consider two analytical examples before presenting the process examples.

Example 3: Model sensitivity under active set variation—linear case

Consider the model-based minimization problem:

$$\begin{aligned} \min_{x,k} \quad & \phi\colon\ 4(x-4)^2 + 9(k-5)^2 \\ \text{s.t.} \quad & g_1\colon\ k + 5x - 23.2 \le 0 \\ & g_2\colon\ k - 3.5 \le 0 \\ & k \ge 0;\ x \ge 0 \end{aligned}$$

This problem is similar to that in example 2. However, the upper bound for the model-based property k has been reduced and results in a change for the linear inequality constraint g_2. Here we wish to study the first-order correction to the optimal solution in going from the optimum with respect to model I, $k - 0.2x^2 = 0$, to the optimum with respect to model II, $k - e^{0.3776x} = 0$.

If we first assume that the active set is retained, the solution, as obtained in example 2, clearly makes g_2 infeasible, Figure 2a. However by solving Eq. 25 directly, we include g_2 as an active constraint and develop the KKT conditions of Eq. 25 for the sensitivity relationship:

$$\begin{bmatrix} -7.8567 & 0 & 0 & -1.7106 \\ 0 & 18 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ -1.7106 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta k \\ \Delta u^2 \\ \Delta v \end{bmatrix} = - \begin{bmatrix} -41.9873 \\ -7.8545 \\ -0.3 \\ -1.3295 \end{bmatrix}$$

The estimated optimum based on the first order correction is thus: $(x_2, k_2, u_2^2, v_2) \equiv (3.3982, 3.5, 24.2357, 2.7643)$. The actual optimum with respect to Model II is at $(x_2, k_2, u_2^2, v_2) \equiv (3.3172, 3.5, 31.1325, -4.1324)$. It may be noted that the constraint g_1 remains feasible, and the model change introduces a change in the active set.

To consider the case of nonlinear inequalities we next present an analytical example solved using the algorithm developed for the quadratic program in Eq. 27.

Example 4: Model sensitivity under active set variation—nonlinear case

Consider the problem:

$$\min_{x,k} \quad \phi: 4(x-4)^2 + 9(k-5)^2$$
$$\text{s.t.} \quad g_1: k + 5x - 23.2 \le 0$$
$$g_2: k^2 + x^2 - 27 \le 0$$
$$k \ge 0; \quad x \ge 0$$

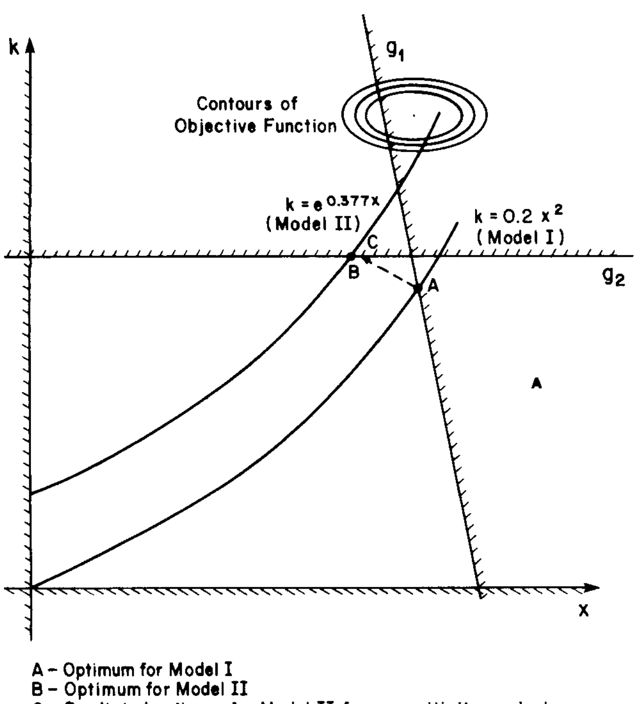


Figure 2a. Model sensitivity: linear inequality constraints.

with

$$\text{Model I: } k - 0.2x^2 = 0$$
$$\text{Model II: } k - e^{0.3776x} = 0$$

For model I, the optimal solution to this problem is: $(x_1, k_1, u_1^1, u_1^2, v_1) \equiv (4, 3.2, 7.8545, 0, 24.546)$. We now substitute model II for model I and apply the quadratic program of Eq. 27 with the above solution vector. By setting $\hat{u}^1 = u_1^1$ and $\hat{u}^2 = u_1^2$, we obtain the following solution: $(x_2, k_2, u_2^1, u_2^2, v_2) \equiv (3.65, 3.93, 0, 4.55, 10.16)$.

This is the same solution as with Eq. 25. Note that the nonlinear constraint g_2 has been selected to be active. We now set $\hat{u}^1 = 0$ and $\hat{u}^2 = 4.55$ and solve Eq. 27 again. This results in the following estimate of the optimum: $(x_2, k_2, u_2^1, u_2^2, v_2) \equiv (3.65, 3.93, 0, 1.53, 2.92)$, and the active set remains consistent, Figure 2b. The actual optimum vector for model II is: $(x_2, k_2, u_2^1, u_2^2, v_2) \equiv (3.54, 3.8, 0, 1.92, 6.89)$.

Note that the estimates using either Eq. 25 or Eq. 27 give equally reasonable predictions of x and k to changes in the model. The solution using Eq. 27 to account for a consistent active set also gives a slightly better estimate of the multipliers. For the nonlinear case the estimated optimum from sensitivity analysis satisfies the linearized form of the active inequality constraint g_2, Figure 2b.

For illustration, the above examples use a full Hessian approach for the Newton steps. We note that the reduced Hessian strategy can also be applied in a straightforward manner for model sensitivity analysis. Here we consider the reduced quadratic program in Δx, where x, as defined previously, refers to

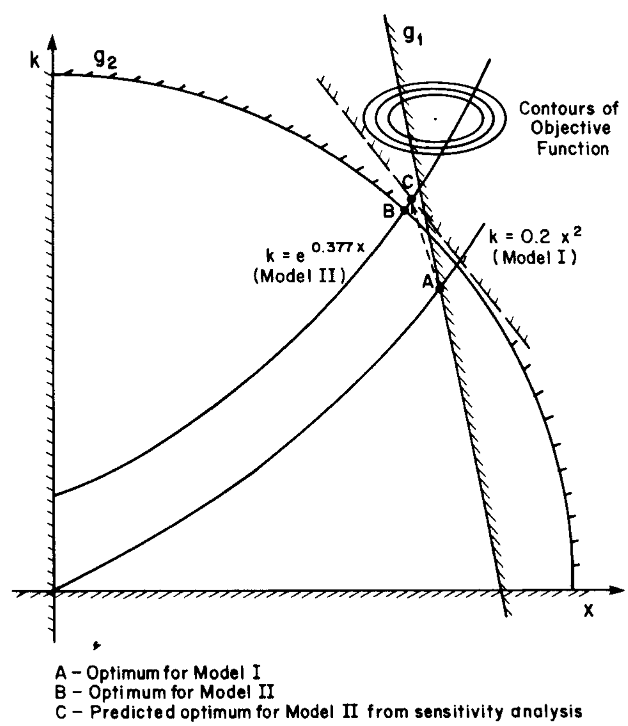


Figure 2b. Model sensitivity: nonlinear constraints.

the independent (decision) variables:

$$\min \quad \phi^T \Delta x + \tfrac{1}{2} \Delta x^T H \Delta x$$
$$\text{s.t.} \quad e + Q^T \Delta x \leq 0 \tag{28}$$

(see Eq. 13), and calculate all but the sensitivities of the equality constraint multipliers Δv with far fewer perturbations than in the full Hessian case.

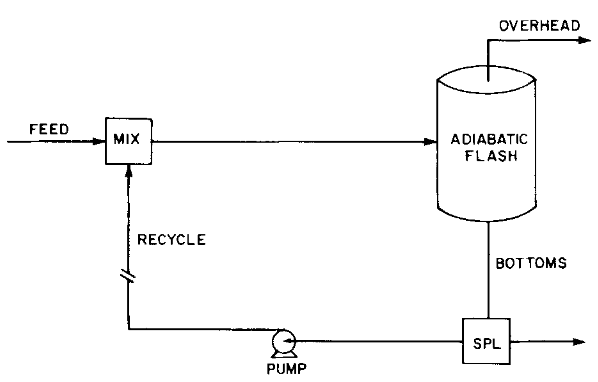
In the following process examples, a reduced Hessian approach is employed for both parametric and model sensitivity analysis.

A simple flash recycle flowsheet

A simple flash recycle problem flowsheet serves to demonstrate the application of the reduced Hessian strategy for the parametric sensitivity and model sensitivity approaches. The flowsheet is presented in Figure 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 1. The process was optimized on a DEC-20 computer using the SPAD simulator from the University of Wisconsin–Madison.

The flowsheet includes two decision variables, the splitter ratio and the pressure in the flash. The six component flow rates and specific enthalpy of the recycle stream together constitute the seven tear variables for the problem. Since the outlet pressure of the pump is fixed in this case, the recycle stream pressure need not belong to the set of tear variables. The objective function for the monotonic optimization problem corresponds to the flow rate 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 sensitivity study, the flow rates 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 given in Table 2. The reduced Hessian strategy requires only 11 evaluations of the flash recycle flowsheet; on the other hand, a complete Hessian based analysis would require 46 flowsheet evaluations. The reduced Hessian strategy results in a 76% saving in the computational time required for sensitivity analysis, Table 4.



SIMPLE FLASH RECYCLE FLOWSHEET

Figure 3. Flash loop example.

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

The results for the model sensitivity analysis are given in

Table 1. Flash Recycle Flowsheet Optimization: Problem Definition

Feed Data

Component	Flow Rate mol/h
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 press. 1,034.25 kPa
Feed temp. 37.78°C

Decision Variables
Splitter ratio, x_1 ($0.2 \leq x_1 \leq 0.8$)
Flash press. x_2 ($69 \leq x_2 \leq 345$) kPa

Tear Variables
Component flows in recycle,
$y_i, i = 1, \ldots, 6$ ($0 \leq y_i \leq 100$) mol/h
Specific enthalpy of recycle,
H ($-1.055\text{E} + 07 \leq H \leq 1.055\text{E} + 07$) J/mol

Objective Function
Monotonic, max e_1
Nonlinear, max $e_1^2 e_2 - e_1^2 - e_3^3 + e_4 - e_5^{0.5}$
(e_i = component flow in flash overhead; Figure 3)

Table 2. Flash Recycle Flowsheet, Sensitivity Analysis: Parametric Variation

Component		Base Case Feed mol/h		Perturbed Feed mol/h	
Propane		10.0		11.0	
1-Butene		15.0		16.5	
n-Butane		20.0		22.0	
trans-2-butene		20.0		22.0	
cis-2-butene		15.0		16.5	
n-Pentane		10.0		11.0	
Decision Variable	**KKT Mult.**	**Base Opt.**	**Corr.**	**Pred. Opt.**	**True Opt.**
		Monotonic Objective Function			
x_1	—	0.2	0.0	0.2	0.2
x_2	—	69.0	0.0	69.0	69.0
—	u_1	1.005	1.04	2.045	1.652
—	u_2	0.146	0.51	0.656	0.21
		Nonlinear Objective Function			
x_1	—	0.8	0.0	0.8	0.8
x_2	—	164.0	−3.3	160.7	159.2
—	u_1	1.654	1.877	3.53	2.56

Table 3. Table 4 shows the number of flowsheet evaluations and CPU times corresponding to the reduced Hessian and complete Hessian strategies, respectively, for the monotonic and nonlinear objective functions. It can be seen that the reduced Hessian procedure results in a 64% reduction in the computational time required for sensitivity analysis. 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, $u = u_1 + \Delta u$, 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 from Eq. 28 and this gives a feasible move for the decision variables. Note also 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 actual optimum for the Chao-Seader model makes the constraint $x_1 \leq 0.8$ 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 (Seader *et al.*, 1977; Figure 1.1). The flowsheet for this problem is shown in Figure 4. A mixture of HCl, 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-liquid equilibrium properties are given in Table 5.

Table 3. Flash Recycle Flowsheet, Sensitivity Analysis: Model Variation

Decision Variable	KKT Mult.	Base Opt.	Corr.	Pred. Opt.	True Opt.
Monotonic Objective Function					
Base model, Raoult's law (model I)					
Alt. model, Chao-Seader (model II)					
x_1	—	0.2	0.0	0.2	0.2
x_2	—	69.0	0.0	69.0	69.0
—	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)					
x_1	—	0.2	0.0	0.2	0.2
x_2	—	69.0	0.0	69.0	69.0
—	u_1	1.35	−0.23	1.12	1.005
—	u_2	0.17	−0.024	0.144	0.146
Nonlinear Objective Function					
Base model, Raoult's law (model I)					
Alt. model, Chao-Seader (model II)					
x_1	—	0.8	−7.55E-03	0.793	0.8
x_2	—	164.0	−2.13	161.87	158.79
—	u_1	1.65	—	0.0	1.43
Base model, Chao-Seader (model I)					
Alt. model, Raoult's law (model II)					
x_1	—	0.8	0.0	0.8	0.8
x_2	—	158.79	8.83	167.62	164.0
—	u_1	1.43	−0.95	0.48	1.65

Table 4. Flash Recycle Problem, Sensitivity Analysis: No Flowsheet Evaluations and CPU Time

Type of Hessian Eval.	NFE	CPU Time, s
Parametric Sensitivity Analysis		
Monotonic Objective Function		
Reduced	11	2.4
Complete	46	9.99
Nonlinear Objective Function		
Reduced	11	2.5
Complete	46	10.4
Model Sensitivity Analysis		
Monotonic Objective Function		
Case A		
Reduced	20	30.0
Complete	55	82.5
Case B		
Reduced	20	4.3
Complete	55	11.95
Nonlinear Objective Function		
Case A		
Reduced	20	30.4
Complete	55	83.5
Case B		
Reduced	20	4.5
Complete	55	12.4

NFE, Number of flowsheet evaluations
Case A, Change from Raoult's law (model I) to Chao-Seader (model II)
Case B, Change from Chao-Seader (model I) to Raoult's law (model II)

Once again for the parametric case a variation in the feed was introduced at the base case optimum. The problem involves six decision variables and five tear variables. The problem was solved using the reduced Hessian procedure; the results of the parametric sensitivity analysis are presented in Table 6. The number of flowsheet evaluations in the reduced Hessian case with six decision variables is 41 and the corresponding CPU time required is 250.1 s. A full Hessian computation for the same problem requires 67 flowsheet evaluations corresponding to an equivalent time consumption of 408.7 s, both times computed on a VAX-11/780. The reduction in time consumption for

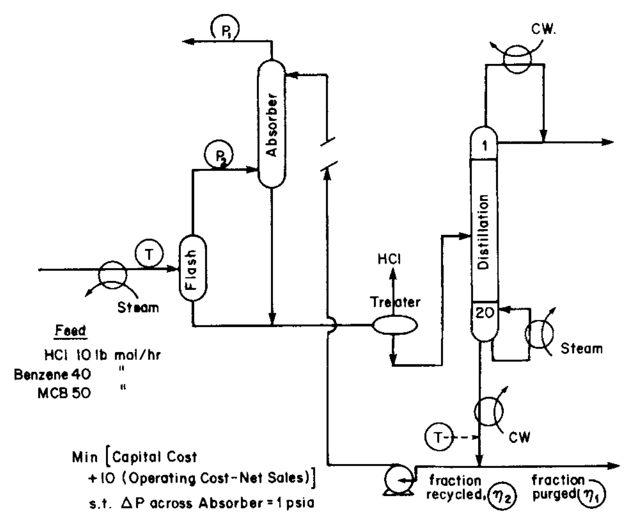


Figure 4. Monochlorobenzene flowsheet example.

Table 5. Monochlorobenzene (MCB) Separation Flowsheet

Feed Data

Component	Flow Rate mol/h
HCl	10
Benzene	40
MCB	50
Feed Press.,	255.12 kPa
Feed Temp.	26.67°C

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

Absorber press. (bottom)	$172.4 \le x_1 \le 241.33$ kPa
Absorber press. (top)	$172.4 \le x_2 \le 241.33$ kPa
Split ratio to recycle	$0 \le x_3 \le 1$
Split ratio to outside	$0 \le x_4 \le 1$
Flash input stream temp.	$93.33 \le x_5 \le 143.33$°C
Absorber input stream temp. (recycle)	$37.78 \le x_6 \le 148.89$°C

Tear Variables
- Recycle HCl flow
- Recycle benzene flow
- Recycle MCB flow
- Recycle press.
- Recycle Temp.

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 7. 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 in fact 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 s, whereas a full Hessian computation for model sensitivity analysis would require 78 flowsheet evaluations with an equivalent time requirement of 475.8 s. The savings in this case is of the order of 33%.

We note in this example that even though the number of tear variables in this flowsheet is in fact less than the number of decision variables, we still obtain noticeable savings in computational effort. In the flash recycle example, and in general, the number of tear variables may exceed the number of decision variables by a factor of three or four, and this 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 or the constraint functions happen to be highly nonlinear or the optima are far apart, 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

Table 6. Monochlorobenzene (MCB) Separation Flowsheet, Sensitivity Analysis: Parametric Variation

Component	Base Case Feed mol/h		Perturbed Feed mol/h	
HCl	10.0		10.1	
Benzene	40.0		40.4	
MCB	50.0		50.5	
Decision Variable	**Base Opt.**	**Corr.**	**Pred. Opt.**	**True Opt.**
Absbr. press. (bottom), kPa	222.43	−0.69	221.74	221.28
Absbr. press (top), kPa	215.54	−0.69	214.85	214.38
Split ratio (Recycle)	0.3155	−2.21E-02	0.3132	0.3102
Split ratio (Outside)	0.6845	2.21E-02	0.6867	0.6898
Input temp. (flash)°C	93.33*	0.0	93.33	93.33
Input temp (absbr.)°C	38.95	−0.28	38.67	38.31

*—variable at its bound (active inequality)

Table 7. Monochlorobenzene (MCB) Separation Flowsheet, Sensitivity Analysis: Activity Coefficient Model Variation

Decision Variable	Base Opt.	Corr.	Pred. Opt.	True Opt.
Base Model, Ideal solution option (FLOWTRAN) (model I)				
Alt. Model, Regular solution option (FLOWTRAN) (model II)				
Absbr. press (bottom) kPa	222.43	0.35	222.78	222.52
Absbr. Press (top) kPa	215.54	0.35	215.89	215.62
Split ratio (recycle)	0.3155	0.0716	0.3871	0.3319
Split ratio (outside)	0.6845	−0.0716	0.6129	0.6681
Input temp. (flash)°C	93.33*	0.0	93.33	93.33
Input temp (absbr.)°C	38.95	0.6	39.55	37.78*

*—variable at its bound (active inequality)

change for a process variable. For example, note that in the monochlorobenzene separation problem the actual change in the absorber input temperature x_6 is in a direction opposite to that 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-size 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 by solving quadratic programs to maintain a consistent active set. It can also be stated rigorously as a mixed-integer nonlinear program. 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.

It should be emphasized that one must be cautious while interpreting the results for changes in the optimal solution from a model sensitivity analysis. If the search direction is large, the first-order information may be overshadowed by higher order effects. This indicates that the solution is sensitive to the model used and that the problem need be optimized with respect to the new model in order to determine the actual optimum. In this case the search direction by itself may not provide a meaningful estimate of the true optimum. However, following along the search direction will, under mild conditions, lead to an improved point in the space of the new model, and for a finite neighborhood, repeated point-to-point application of this approach (Eq. 25 or 27) will lead to convergence to the optimum for model II. On the other hand, when the search direction is small this analysis gives a very useful indication of the changes in the variables in the direction of the new optimum. For example, when local simplified models are used, this strategy can be effective in generating meaningful estimates of the deviations in the optimal solution resulting locally from the use of one simplified model vs. another. Thus the best use of the model sensitivity approach is that it provides an efficient quantitative means of evaluating how sensitive the optimum is to the choice of process models.

Acknowledgment

Financial aid from the National Science Foundation under Grant No. ENG 8451058 is gratefully acknowledged. The authors also acknowledge the assistance of Yi-dong Lang with the FLOWTRAN simulator.

Notation

g = inequality constraint
h = equality constraint
k = physical property (model-dependent)
$K, \tilde{K}$ = base case model and new model, respectively, for physical properties
L = Lagrange function
L = Lagrange function defined with respect to model II in model sensitivity
p = input parameter (uncertain or variable over a range)
u = Lagrange (KKT) multiplier for inequality constraint
v = Lagrange (KKT) multiplier for equality constraint
x = independent (decision) variable
y = dependent (tear) variable
z = process variable (includes decisions and tears)

Greek letters

ϕ = objective function
ω = optimal solution vector, (x, y, u, v)
ϵ = perturbation factor for generating derivatives by forward difference approximation
Δ = delta operator, difference from a reference value for a variable
∇_i = vector or matrix of differentials with respect to variable i of row dimension equal to dim $\{i\}$
∇_{ij} = matrix of differentials of order, dim $\{i\} \times$ dim $\{j\}$

Subscripts

A = active inequality constraint
k = kth element of input parameter vector
1 = optimal solution for model I
2 = optimal solution for model II

Superscripts

T = transpose of a matrix
o = base case optimal solution for parametric sensitivity analysis
j = identifier for KKT multiplier corresponding to jth inequality constraint, g_j

Appendix A

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^o, a feasible solution to $\boldsymbol{P}(p^o)$, the gradients of the binding constraints form a linearly independent set, i.e.,

$$\Sigma u^j \nabla g_j(z^o, p^o) + \Sigma v^i \nabla h_i(z^o, p^o) = 0$$

and $g_j = 0 \Longrightarrow u^j = v^i = 0$ for all i, j.

Strict complementary slackness (SCS)

For z^o, an isolated minimizer of $\boldsymbol{P}(p^o)$, SCS holds if the KKT multipliers u^o, v^o are such that

$$\nabla_z \phi^o + \nabla_z g^o u^o + \nabla_z h^o v^o = 0$$

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

$$u^o \geq 0$$

and $g_j(z^o, p^o) = 0 \Longrightarrow u^{jo} > 0$.
Note: $\phi^o \equiv \phi(z^o, p^o)$; similarly g^o, h^o

Second-order sufficient condition (SOC)

SOC is said to hold for $P(p^o)$ at z^o if:

$$\text{(1) } z^o \text{ is a feasible point of } P(p^o)$$

There exists u^o, v^o such that

$$\text{(2) } \nabla_z\phi^o + \nabla_z g^o u^o + \nabla_z h^o v^o = 0$$

$$u^{o^T} g(z^o, p^o) = 0$$

$$u^o \geq 0$$

and (3) for all nonzero $q \in \mathbb{R}^n$ satisfying $\nabla g_A^T q = 0$ and $\nabla h^T q = 0$ we have

$$q^T \nabla_{ZZ} \text{L}(z^o, u^o, v^o, p^o) q > 0$$

where $\text{L} = \phi + u^T g + v^T h$ (Lagrange function).

Appendix B

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 (see Boggs et al., 1982):

$$Q^{k+1} = Q^k - \frac{Q^k d d^T Q^k}{d^T Q^k d} + \frac{\eta\eta^T}{\eta^T d}$$

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

$$\eta = \nabla_z \text{L}^{k+1} - \nabla_z \text{L}^k \tag{B1}$$

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

$$\lim_{k\to\infty} \frac{\|Z^T(B - Q^k)d\|}{\|d\|} \to 0 \tag{B2}$$

where Z is the null space of $[\nabla_z g_A \; \nabla_z h]^T$.

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

$$\lim_{k\to\infty} \frac{\|Z^T(B - Q^k)ZZ^T d\|}{\|d\|} \to 0 \tag{B3}$$

In either case, one can show that at the limit point Q^k and B differ by ADA^T, where $A \equiv [\nabla_z g_A \nabla_z h]$ 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 $D[\nabla_z g_A^T \nabla_z h^T].^T$

Aside from this, Q is usually initialized arbitrarily (e.g., $Q^o = I$), and from Eq. B1 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^o. 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 C

Reduced Hessian decomposition strategy

Consider the linear system of equations:

$$\begin{bmatrix} \Delta(\nabla_x \text{L}) \\ \Delta(\nabla_y \text{L}) \\ \Delta g_A \\ \Delta h \end{bmatrix} = - \begin{bmatrix} \nabla_{xx}\text{L} & \nabla_{xy}\text{L} & \nabla_x g_A & \nabla_x h \\ \nabla_{yx}\text{L} & \nabla_{yy}\text{L} & \nabla_y g_A & \nabla_y h \\ \nabla_x g_A^T & \nabla_y g_A^T & 0 & 0 \\ \nabla_x h^T & \nabla_y h^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta u \\ \Delta v \end{bmatrix} \tag{C1}$$

Rearranging the system of equations in Eq. C1, we get:

$$\begin{bmatrix} \Delta h \\ \Delta(\nabla_y \text{L}) \\ \Delta(\nabla_x \text{L}) \\ \Delta g_A \end{bmatrix} = - \begin{bmatrix} \nabla_y h^T & 0 & \nabla_x h^T & 0 \\ \nabla_{yy}\text{L} & \nabla_y h & \nabla_{yx}\text{L} & \nabla_y g_A \\ \nabla_{xy}\text{L} & \nabla_x h & \nabla_{xx}\text{L} & \nabla_x g_A \\ \nabla_y g_A^T & 0 & \nabla_x g_A^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta v \\ \Delta x \\ \Delta u \end{bmatrix} \tag{C2}$$

We note that the term $\nabla_y h^T$ is square and nonsingular. Hence on applying a Gaussian elimination to Eq. C2 we get the resulting system (Edahl, 1982):

$$\begin{bmatrix} a \\ b \\ f \\ e \end{bmatrix} = - \begin{bmatrix} I & 0 & E & 0 \\ 0 & I & L & M \\ 0 & 0 & H & Q \\ 0 & 0 & Q^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta v \\ \Delta x \\ \Delta u \end{bmatrix} \tag{C3}$$

where the terms have been defined previously in the section on reduced Hessian evaluation (refer to 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:

$$\begin{bmatrix} f \\ e \end{bmatrix} = - \begin{bmatrix} H & Q \\ Q^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \end{bmatrix} \tag{C4}$$

In Eq. C4 the terms Q, Q^T, and e are readily obtained from the information regarding the constraint gradients at the local minimum from the nominal optimization problem. However the terms H and f need to be derived.

Let us first consider the derivation for f. The term, $\nabla_x h(\nabla_y h)^{-1}\nabla_{yy}\text{L}(\nabla_y h^T)^{-1}\Delta h$, for f can be expanded at the base case optimum (x_1, y_1) as follows:

$$\begin{aligned} &\nabla_x h(\nabla_y h)^{-1}\nabla_{yy}\text{L}(\nabla_y h^T)^{-1}\Delta h \\ &\quad = \nabla_x h(\nabla_y h)^{-1}\nabla_y \text{L}[y_1 + (\nabla_y h^T)^{-1}\Delta h, x_1] \\ &\quad\quad - \nabla_x h(\nabla_y h)^{-1}\nabla_y \text{L}(y_1, x_1) \end{aligned} \tag{C5}$$

Once again, the second term on the righthand side of Eq. C5 is directly obtained from the base case optimization results. However, the first term on the righthand 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, i.e.,

$$\nabla_x h^T \Delta x + \nabla_y h^T \Delta y = 0 \Longrightarrow \Delta y^T = -\Delta x^T \nabla_x h (\nabla_y h)^{-1} \quad \text{(C6)}$$

Next we expand the Lagrange function, L, about $[x_1, y_1 + (\Delta_y h^T)^{-1}\Delta h + \Delta y]$ to first order so that:

$$\begin{aligned} \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h + \Delta y] &= \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h] \\ &\quad + \Delta y^T \nabla_y \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h] \quad \text{(C7)} \end{aligned}$$

We now substitute for Δy^T in Eq. C7 from Eq. C6, which transforms Eq. C7 accordingly as:

$$\begin{aligned} \Delta x^T\{\nabla_x h(\nabla_y h)^{-1}\nabla_y L[y_1 + (\nabla_y h^T)^{-1}\Delta h, x_1]\} &= \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h] - \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h \\ &\quad - (\nabla_y h^T)^{-1}\nabla_x h^T \Delta x] \quad \text{(C8)} \end{aligned}$$

Equation C8 implies that the first term on the righthand side of Eq. C5 can be constructed by introducing perturbations in the decisions alone; it is worth noting that the term $\nabla_{yy}\mathrm{L}$ need no longer be determined explicitly. Similarly, it can be shown that the term $\nabla_{xy}\mathrm{L}(\nabla_y h^T)^{-1}\Delta h$ can be constructed from the relations:

$$\nabla_{xy}\mathrm{L}(\nabla_y h^T)^{-1}\Delta h = \nabla_x \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h] - \nabla_x \mathrm{L}(x_1, y_1)$$

and

$$\begin{aligned} \Delta x^T \nabla_x \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h] &= \mathrm{L}[x_1 + \Delta x, y_1 + (\nabla_y h^T)^{-1}\Delta h] \\ &\quad - \mathrm{L}[x_1, y_1 + (\nabla_y h^T)^{-1}\Delta h] \quad \text{(C9)} \end{aligned}$$

Once again, the term $\nabla_{xy}\mathrm{L}$ 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, $\boldsymbol{H}$. It is seen that $\boldsymbol{H}$ can be derived from:

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}^T \begin{bmatrix} \nabla_{xx}\mathrm{L} & \nabla_{xy}\mathrm{L} \\ \nabla_{yx}\mathrm{L} & \nabla_{yy}\mathrm{L} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

by introducing the Δy perturbations in accordance with the relation $\Delta y = -(\nabla_y h^T)^{-1}\nabla_x h^T \Delta x$ from Eq. C6. Substituting this result and expanding the above product term gives:

$$\begin{bmatrix} \Delta x \\ A\Delta x \end{bmatrix}^T \begin{bmatrix} \nabla_{xx}\mathrm{L} & \nabla_{xy}\mathrm{L} \\ \nabla_{yx}\mathrm{L} & \nabla_{yy}\mathrm{L} \end{bmatrix} \begin{bmatrix} \Delta x \\ A\Delta x \end{bmatrix} = \Delta x^T \boldsymbol{H} \Delta x$$

where

$$\boldsymbol{A} \equiv -(\nabla_y h^T)^{-1}\nabla_x h \quad \text{(C10)}$$

The term $\Delta\psi$ (for $\psi = \nabla_x \mathrm{L}, \nabla_y \mathrm{L}, g_A, h$) is computed by perturbing the parameters and evaluating the respective differentials between the perturbed and base case function values, see Eq. 10.

Appendix D

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. Examples 3 and 4 serve 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 where all g_j are linear and another that involves nonlinear inequality constraints. Note that $\hat{u}$ only appears in $\nabla_{zz}L$, and in the linear case it does not appear at all. Here we can obtain a consistent active set by solving the quadratic program (QP) of Eq. 25. For the nonlinear case, this QP can be solved repeatedly or we can represent changing active sets by binary (0–1) variables, y_j, and rigorously formulate the problem as the mixed-integer nonlinear program (MINLP) shown below:

$$\begin{aligned} \max_{\Delta z, \Delta u, \Delta v, y, \hat{u}} \quad & \Sigma \hat{u}^j \\ \text{s.t.} \quad & \nabla_z L(\bar{z}, u, v) + \nabla_{zz} L(\bar{z}, \hat{u}, v_1)\Delta z = 0 \\ & g(\bar{z}) + \nabla g^T(\bar{z})\Delta z \le 0 \\ & h(\bar{z}) + \nabla h^T(\bar{z})\Delta z = 0 \\ & u = \hat{u} + \Delta u \ge 0; \quad v = v_1 + \Delta v \\ & 0 \le \hat{u} \le u \\ & u^j \le U y_j \\ & g_j(\bar{z}) + \nabla g_j^T(\bar{z})\Delta z \ge U(y_j - 1) \\ & y_j = 0, 1 \end{aligned} \quad (1)$$

and U is an arbitrarily large constant

It can be seen that the constraints are linear in y and Δz. A sufficient condition for a solution is that the QP of Eq. 27 is solvable with $\hat{u} = 0$. This is also a feasible lower bound. The upper bound occurs if $\hat{u} = u$, but this may not be feasible. This problem can be solved by applying the outer approximation approach (Duran and Grossmann, 1986). It is obvious that this formulation gives rise to a combinatorial problem, where the number of nonlinear inequalities is an important factor in the solution of the MINLP formulation. Here, if the number of nonlinear inequality constraints is equal to n, then the solution of the MINLP to obtain a consistent active set may require up to 2^n discrete decisions.

Literature cited

Atherton, R. W., R. B. Schainker, and E. R. Ducot, "On the Statistical Sensitivity Analysis of Models for Chemical Kinetics," *AIChE J.*, **21**, 411 (1975).

Bard, Y. A., *Nonlinear Parameter Estimation,* Academic Press, New York (1974).

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

Biegler, L. T., "Improved Infeasible Path Optimization for Sequential Modular Simulators," *Comp. Chem. Eng.*, **9**, 257 (1985).

Biegler, L. T., and R. R. Hughes, "Infeasible Path Optimization of Sequential Modular Simulators," *AIChE J.*, **28**, 994 (1982).

Biegler, L. T., I. E. Grossmann, and A. W. Westerberg, "A Note on Approximation Techniques Used in Process Optimization," *Comp. Chem. Eng.*, **9**, 201 (1985).

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

Bryan, P. F., and E. A. Grens II, "Efficient Use of Thermodynamic Properties for Phase-Equilibrium Calculations in Solvent Extractions," Int. Solvent Extraction Conf., Denver (1983).

Chimowitz, E. H., T. F. Anderson, S. Macchietto, and L. F. Stutzman, "Local Models for Representing Phase Equilibria in Multicomponent Nonideal Vapor-Liquid and Liquid-Liquid Systems. 1: Thermodynamic Approximation Functions," *Ind. Eng. Chem. Process Des. Dev.*, **22**, 217 (1983).

———, "Local Models for Representing Phase Equilibria in Multicomponent Nonideal Vapor-Liquid and Liquid-Liquid Systems. 2. Application to Process Design," *Ind. Eng. Chem. Process Des. Dev.*, **23**, 609 (1984).

Duran, M. A., and I. E. Grossmann, "A Mixed-Integer Nonlinear Programming Algorithm for Process System Synthesis," *AIChE J.*, **32**, 592 (1986).

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

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

———, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. Academic Press, New York (1983).

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

Gallagher, S. A., and M. A. Kramer, "Parametric Sensitivity Analysis of Chemical Process Flowsheets," AIChE Meet., San Francisco (1984).

Grens II, E. A., "Efficient Use of Thermodynamic Models in Process Calculations," 2nd Int. Conf. Found. Computer-Aided Process Design, Snowmass, CO (1983).

Jirapongphan, S., J. F. Boston, H. I. Britt, and L. B. Evans, "A Nonlinear Simultaneous Modular Algorithm for Process Flowsheet Optimization," 80th AIChE Meet., Chicago (1980), Paper 3b.

Klein, H. H., "Computer Modeling of Chemical Process Reactors," 2nd Int. Conf. Found. Computer-Aided Process Design, Snowmass, CO (1983).

Luenberger, D. G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA (1973).

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

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 (1978).

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

Stadtherr, M. A., and H. S. Chen, "Numerical Techniques for Process Optimization by Successive Quadratic Programming," 3rd Int. Cong. Computers Chem. Eng., Paris (1983).

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

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

Manuscript received May 9, 1986, and revision received Aug. 22, 1986.