Relaxation Strategy For The Structural Optimization Of Process Flowsheets

by

G. R. Kocis, I. E. Grossmann

EDRC-06-01-86 }

September 1986

RELAXATION STRATEGY FOR THE STRUCTURAL OPTIMIZATION OF PROCESS FLOWSHEETS

Gary R. Kocis and Ignacio E. Grossmann

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

July, 1986

Paper Presented at the Annual AIChE Meeting, Miami, November 1986.

ABSTRACT

This paper presents an Equality Relaxation variant to the Outer-Approximation algorithm for solving mixed-integer nonlinear programming (MINLP) problems that arise in structural optimization of process flowsheets. The proposed algorithm has the important capability of being able to explicitly handle nonlinear equations within MINLP formulations that have linear integer variables and linear/nonlinear continuous variables. It is shown that through the explicit treatment of nonlinear equations, the proposed algorithm avoids computational difficulties (e.g. singularities, destruction of sparsity) that are experienced with algebraic or numerical elimination schemes. Also, theoretical properties of the Equality-Relaxation algorithm are discussed, and its performance is demonstrated with a planning problem and a flowsheet synthesis problem. Finally, a simple procedure for structural sensitivity analysis is presented.

Introduction

Process synthesis can be defined as the selection, arrangement, and operation of processing units so as to create an optimal scheme. This task is combinatorial and open-ended in nature and has received a great deal of attention over the past twenty years. An excellent review of research activities in this area can be found in Nishida, Stephanopoulos, and Westerberg (1981).

Because the synthesis problem is open-ended, it has motivated the development of quite different approaches. Synthesis methods currently include thermodynamic targets (Linnhoff,1981), heuristic (Douglas, 1985) and evolutionary methods (Stephanopoulos and Westerberg, 1976), and optimization techniques (Grossmann, 1985). This paper will address the structural flowsheet optimization problem that arises in the latter approach.

In order to formulate the synthesis problem as a mathematical programming problem, a superstructure must be postulated which includes many alternative designs from which the optimal process will be selected. Superstructures can be developed systematically for homogeneous processes (e.g. heat exchanger networks, separation sequences) while for heterogeneous processes the superstructure is specified by the user based on a preliminary screening of alternatives, possibly through the application of heuristic rules and/or thermodynamic targets. In order to determine the optimal process flowsheet, simultaneous structural and parameter optimization of the superstructure is required. In general, this leads to a mixed-integer optimization problem (see Grossmann, 1985).

Most of the previous work on process synthesis that is based on the mixed-integer optimization approach has relied on the use of mixed-integer linear programming (MILP) formulations (e.g. see Papoulias and Grossmann, 1983; Andrecovich and Westerberg, 1985; Hillenbrand, 1984; Floudas et al, 1986; Shelton and Grossmann, 1986). Although these formulations have proved to be quite powerful, they have the limitation that nonlinearities cannot be treated explicitly, and hence they must be approximated through the discretization of operating conditions (see Papoulias and Grossmann, 1983). The need for the explicit handling of the nonlinearities in the synthesis problem motivates the use of mixed-integer nonlinear programming (MINLP).

MINLP problems, however, are much more difficult to solve than MILP problems for which Branch and Bound methods perform reasonably well.

It should be noted that for process synthesis applications, the MINLP problems have a special structure in which the 0-1 variables appear linearly and the continuous variables appear linearly and nonlinearly (see Duran and Grossmann, 1986a). Current alternatives for solving these MINLP problems include Branch and Bound (Beale, 1977; Garfinkel and Nemhauser, 1972; Little, 1963) , Generalized Benders Decomposition (GBD) (Benders, 1962; Geoffrion, 1972), and Outer-Approximation (OA) method (Duran and Grossmann 1986a; 1986b). Since in structural flowsheet optimization problems the majority of the computational effort is consumed by the solution of a sequence of nonlinear programming problems (NLP), a reasonable measure of the efficiency of the above cited algorithms is the number of NLP subproblems that they must solve.

Duran and Grossmann proposed the Outer-Approximation (OA) algorithm with the objective of reducing the number of NLP subproblems that must be solved. Although this algorithm is an efficient method for solving MINLP problems, its main limitation is that it can handle only linear equality and nonlinear / linear inequality constraints. However, the formulation of the process synthesis problem will typically contain many nonlinear equations which describe the performance of process units. In the OA algorithm, nonlinear equations must be eliminated algebraically or numerically. This is usually a nontrivial task which can lead to theoretical and numerical difficulties as will be discussed later in the paper.

The main objective of this paper is to present a new variant to the OA algorithm that can explicitly handle nonlinear equality constraints. The proposed algorithm relies on an equality relaxation strategy which has the advantage of not requiring the selection of decision variables, nor the algebraic or numerical elimination of the equations. As will be shown, this algorithm is well suited to solve MINLP problems that arise in structural flowsheet optimization. Basic properties of this algorithm are discussed and its application is illustrated with three example problems. Also, it will be shown how to perform structural sensitivity analysis with the proposed method.

Problem Formulation

The structural flowsheet optimization problem for process synthesis can be formulated as a MINLP of the following form:

$$
\begin{aligned}
z &= \min_{x,y}\ \mathbf{c}^T \mathbf{y} + f(\mathbf{x}) \\
\text{s.t.}\quad & \mathbf{h}(\mathbf{x}) = 0 \\
& \mathbf{g}(\mathbf{x}) \le 0 \\
& \mathbf{A}\,\mathbf{x} = \mathbf{a} \qquad\qquad (P) \\
& \mathbf{B}\,\mathbf{y} + \mathbf{C}\,\mathbf{x} \le \mathbf{d} \\
& \mathbf{x} \in X = \{\mathbf{x} \mid \mathbf{x} \in R^n,\ \mathbf{x}^L \le \mathbf{x} \le \mathbf{x}^U\} \\
& \mathbf{y} \in Y = \{\mathbf{y} \mid \mathbf{y} \in \{0,1\}^m,\ \mathbf{E}\,\mathbf{y} \le \mathbf{e}\}
\end{aligned}
$$

where **x** is the vector of continuous variables specified in the compact set X, and y is the vector of 0-1 variables which must satisfy pure integer constraints **Ey≤e**. f(x), **h(x)**, and **g(x)** represent nonlinear functions involved in the objective function, equations, and inequalities of a process, respectively. Finally, **Ax=a** represents the subset of linear equations, while **By+Cx≤d** represents linear inequalities that involve the continuous and integer variables.

In the context of the synthesis problem the continuous variables **x** include flows, pressures, temperatures and sizes, while the binary variables **y** represent the potential existence of units which are embedded in the superstructure. The equations h(x)=0 and **Ax=a** correspond to material and energy balances and design equations. Process specifications are represented by **g(x)≤0** and by lower and upper bounds on the variables in **x**. Logical constraints that must hold for a flowsheet configuration to be selected from within the superstructure are represented by **By+Cx≤d** and **Ey**≤e. The cost function involves fixed cost charges in the term $\mathbf{c}^T\mathbf{y}$ for the investment , while revenues, operating costs, and size dependent costs for the investment are included in the function f(**x**).

In order to solve problem (P) Duran and Grossmann proposed the Outer-Approximation (OA) algorithm. The method is similar to Generalized Benders

Decomposition (GBD) in that an alternating sequence of NLP subproblems and **MILP** master problems are solved. The main difference however, is the formulation of the master problem. In contrast to the master of GBD which is based on a dual representation, the master of OA is based on a primal representation. The master includes the linear constraints from the original MINLP, as well as linear approximations to the nonlinear functions derived at each NLP subproblem solution. This master problem provides an increasingly good approximation to the MINLP, and therefore can predict strong lower bounds on the optimal MINLP objective function value. In this way, convergence is usually achieved in few iterations, but at the expense of solving a larger master problem than in the GBD method. Since nonlinear programming problems are often more difficult to solve than mixed-integer programming problems, the overall result is expected to be a decrease in computational expense. Duran and Grossmann (1986b) proved that when the nonlinear functions in the MINLP are convex, the lower bound predicted from OA is always greater or equal to that of GBD. They also demonstrated computational savings in solving several test problems.

The basic idea behind both GBD and OA, as applied to the flowsheet synthesis problem, is illustrated in the iterative procedure shown in Figure 1. The NLP subproblem step is nothing other than the optimization of the continuous variables for a particular flowsheet. The MILP master problem selects, from among the remaining alternatives embedded in the superstructure, a flowsheet configuration whose lower bound lies below the current best flowsheet objective function value. When such a configuration cannot be found, the MILP master problem provides the termination criterion for the search.

When the nonlinear equations in problem (P) are absent or eliminated, the Outer-Approximation algorithm as suggested by Duran and Grossmann (1986b) can be applied directly to solve this problem. However, flowsheet synthesis problems typically involve a large number of nonlinear equations. Since these constraints cannot be explicitly treated in the OA algorithm, they must be eliminated from problem (P). One option is to algebraically eliminate the nonlinear equations so as to leave only nonlinear inequalities and linear constraints. This task, however, is only possible for certain specific problems; for instance, the design of batch processes

(Vaselenak, 1985) and the synthesis of gas pipelines (Duran and Grossmann, 1986a). Clearly algebraic elimination is not practical for typical flowsheet synthesis problems which involve a large number of nonlinear equations.

The other option is to perform a numerical elimination through linearizations of the nonlinear equations at each iteration of the OA algorithm. However, depending on the choice of decision variables, computational difficulties are often encountered due to singularities. Furthermore, there is loss of sparsity in the master problem (see Appendix C). Although there is a reduction in the number of continuous variables in this elimination scheme, the matrix of coefficients of the original inequalities and of the lower and upper bounds on the continuous variables x becomes very dense. In fact, Duran (1984) reported numerical difficulties when attempting to solve a large-scale MINLP with numerical elimination.

The above cited limitations in the OA algorithm have motivated the development of a new Equality-Relaxation algorithm (OA/ER) which attempts to preserve the computational efficiency of the former method, while having the capability to handle explicitly nonlinear equality constraints.

Master Problem

The proposed Equality-Relaxation algorithm is similar to the Outer-Approximation algorithm in that the procedure involves the solution of a sequence of NLP and MILP problems (see Figure 1). The NLP problems correspond to the original MINLP, problem (P), with the vector of binary variables temporarily fixed. Physically this step is simply the parameter optimization of a particular flowsheet in the space of the continuous variables. The master problem is a primal approximation to the original MINLP problem formulation. The MILP master problem is intended to be a relaxation of problem (P), meaning that the linear objective function approximation will underestimate the nonlinear function and the linear constraints will overestimate the nonlinear feasible region so as to provide a lower bound to the MINLP global optimum. In the OA/ER algorithm a new master problem will be defined in a way that nonlinear equations in the MINLP formulation can be handled explicitly.

At iteration k, the vector y in problem (P) is fixed at y^k (which satisfies the restriction y G Y) yielding the following problem:

$$
\begin{aligned}
& z(y^k) = \min_x \; c^T y^k + fix) \\
s.t.\quad & h(x).= 0 \\
& g(x) \;<\; o \qquad\qquad \text{(NLPD} \\
& Ax=a \\
& Cx \leq d - B\,y^k \\
& xGX
\end{aligned}
$$

Assume that the solution of (NLP1) has x^k as the optimal vector of continuous variables. The associated optimal lagrange multipliers for the nonlinear equations $h_i(x) = 0$, i=1,2...r, will be denoted by X_i^k. It is convenient to define the diagonal matrix 1*, (rXr) where r is the number of nonlinear equations, having diagonal elements t_{ii}^k given by:

$$
t_{ii}^k = \begin{cases} -1 & if\; X_i^k < 0 \\ +1 & if\; X_i^k > 0 \\ 0 & if\; X_i^k = 0 \end{cases} \qquad l=1,2,...r \qquad (1)
$$

The matrix 1* will be denoted as the direction matrix because as shown below, it defines a valid direction for the relaxation of the nonlinear equalities into inequalities. Relaxing the nonlinear equations h(x)=0 in (NLP1), through the premultipiication by this matrix, yields the relaxed problem:

$$
\begin{aligned}
& z(y^k) = \min_x \; c^T y^k + fix) \\
s.t.\quad & T''h(x) \leq 0 \\
& g(x) \;\wedge\; 0 \qquad\qquad \text{(NLP2)} \\
& Ax=a \\
& Cx \leq d - B\,y'' \\
& xGX
\end{aligned}
$$

Problems (NLP1) and (NLP2) can be shown to be equivalent if certain conditions are imposed upon the nonlinear functions f(x), g(x), and $T^k h(x)$.

From the Kuhn-Tucker conditions for problem (NLP1), the stationary conditions of the lagrangian and the sign restrictions on the multipliers yield (Bazaraa and Shetty, 1979)

$$\nabla f(x^k) + (\mu^k)^T \nabla g(x^k) + (\lambda^k)^T \nabla h(x^k) + (\alpha^k)^T A + (\beta^k)^T C = 0$$

$$\mu^k \geq 0 , \quad \beta^k \geq 0 \qquad (2)$$

where λ, α, μ, and β are vectors of lagrange multipliers for constraints h(x)=0, Ax=a, g(x)≤0, and By+Cx≤0, respectively. If in addition the following conditions on the nonlinear functions are satisfied at any feasible point x^k:

- f(x) is pseudoconvex,
- $g_i(x)$ is quasiconvex for $i \in I_{EQ}$, $I_{EQ}=\{i: g_i(x^k) = 0\}$, $I_{IN}=\{i: g_i(x^k) < 0\}$
- $h_i(x)$ is quasiconvex for $i \in J_+$, $J_+=\{i: \lambda_i^k > 0\}$
- $h_i(x)$ is quasiconcave for $i \in J_-$, $J_-=\{i: \lambda_i^k < 0\}$

then the Kuhn-Tucker conditions at x^k are sufficient for the global solution of (NLP1). Note that at x^k the nonlinear inequalities have been partitioned into active (EQ) and inactive (IN) inequalities, and the nonlinear equations into those with positive (J_+) and negative (J_-) multipliers.

Similarly, the Kuhn-Tucker conditions for the formulation (NLP2) yield

$$\nabla f(\bar{x}) + (\bar{\mu})^T \nabla g(\bar{x}) + (\bar{\lambda})^T T^k \nabla h(\bar{x}) + (\bar{\alpha})^T A + (\bar{\beta})^T C = 0$$

$$\bar{\mu} \geq 0 , \quad \bar{\beta} \geq 0 , \quad \bar{\lambda} \geq 0 \qquad (3)$$

where $\bar{\lambda}_T$ is the vector of multipliers for the constraint set $T^k h(x) \leq 0$ and $\bar{x}$ is the optimal solution to (NLP2). Sufficient conditions for a feasible point $\bar{x}$ to be the global optimum to (NLP2) are

- $f(x)$ is pseudoconvex,
- $g_i(x)$ is quasiconvex for $i \in I^r_{EQ}$, $I^r_{EQ}=\{i: g_i(\bar{x}) = 0\}$, $I^r_{IN}=\{i: g_i(\bar{x}) < 0\}$
- $T^k\, h(x)$ is quasiconvex.

Under the assumed class of functions stated above, both (NLP1) and (NLP2) have a unique local solution. Furthermore, h_i quasiconvex for $i \in J_+$ and h_i quasiconcave for $i \in J_-$, implies that $T^k h(x) \le 0$ is quasiconvex. Since (2) and (3) are uniquely defined, $\lambda^k = T^k \bar{\lambda}_r$, $x^k = \bar{x}$, $I_{EQ}=I^r_{EQ}$, $I_{IN}=I^r_{IN}$, and thus the conditions for optimality in (2) and (3) are equivalent. This then implies that the (NLP1) formulation is equivalent to the (NLP2) formulation.

Based on the equivalence of (NLP1) and (NLP2), an approximation to problem (P) can be derived from (NLP2) by freeing y, and by replacing the nonlinear functions $f(x)$, $g(x)$, and $T^k h(x)$ with linear approximations. Following a similar reasoning as in the OA algorithm (Duran and Grossmann, 1986b) one can then define a master problem that will yield a valid lower bound to the optimum solution of problem (P). This can be achieved if the linear approximations of the nonlinear functions are given either by function linearizations or by linear underestimators that will overestimate the feasible region and underestimate the objective function (see Appendix B). These outer-approximations at the point x^k must satisfy the following conditions for all $x \in X$:

$$
\begin{aligned}
&(w^k)^T x - w^k_o \le f(x) \\
&T^k\, h(x) \le 0 \Rightarrow T^k\, [R^k x - r^k] \le 0 \qquad (4)\\
&S^k_{IN}\, x - s^k_{IN} \le g_{IN}(x) \\
&g_{EQ}(x) \le 0 \Rightarrow S^k_{EQ}\, x - s^k_{EQ} \le 0
\end{aligned}
$$

where $g_{IN}(x)$ and $g_{EQ}(x)$ correspond to the inactive and active nonlinear inequalities, respectively. If the matrices S^k_{EQ} and S^k_{IN} and the vectors s^k_{EQ} and s^k_{IN} are combined to yield the matrix S^k and vector s^k, then from (4) the MILP approximation to problem (P) at x^k that yields a lower bound z^k_L is given by:

$$
\begin{aligned}
z_L^k = \min_{x,y}\ & c^T y + (w^k)^T x - w^k_o \\
s.t.\ \ & T^k R^k x \le T^k r^k \\
& S^k x \le s^k \\
& A x = a \qquad\qquad (5) \\
& By + Cx \le d \\
& Ey \le e \\
& x \in X \\
& y \in \{0,1\}^m
\end{aligned}
$$

This problem can be generalized by considering K points for the approximations in (4). The MILP master problem with equality relaxation for problem (P) is then given in final form as:

$$
\begin{aligned}
z_L^K = \min_{x,y,\mu}\ & c^T y + \mu \\
s.t.\ \ & \left.\begin{aligned} & (w^k)^T x - \mu \le w^k_o \\ & T^k R^k x \le T^k r^k \\ & S^k x \le s^k \end{aligned}\right\} \quad k = 1, 2 \ldots K \\
& A x = a \qquad\qquad (M^K) \\
& By + Cx \le d \\
& Ey \le e \\
& z_L^{K-1} \le c^T y + \mu \le z_U \\
& x \in X \\
& y \in \{0,1\}^m \cap \{\text{integer cuts}\}^{K-1} \\
& \mu \in R^1
\end{aligned}
$$

where z_L^K is the predicted lower bound at iteration K, μ is the largest linear approximation to the nonlinear objective function, and the integer cuts correspond to constraints which eliminate the assignments of binary variables analyzed at the previous K-1 iterations (see Duran and Grossmann, 1986b). Also, z_y is the current best estimate to the optimal solution to MINLP (P) which is given by the minimum of all NLP subproblems that have been solved. The lower bound z_L^{K1} of the constraint on c^Ty+// is introduced to expedite the solution of the MILP; the upper bound z_U to produce infeasibility as the termination criterion ($z_L^*>z_U$).

Equality-Relaxation Algorithm

Based on the master problem (M^K) presented in the previous section, the following algorithm can be stated to explicitly handle nonlinear equations. For simplicity in the presentation, it is assumed that the NLP subproblems in Step 2 have a feasible solution.

Step J[Select initial binary assignment y^1, set K=1.
Initialize lower and upper bounds, z_L^o =-oo ^=00.

Step 2 Solve (NLP1) for fixed y^K yielding $z(y^K)$, x^K, and X\
If $z(y^K)<z_u$, then set $y^*=y^K$, $x'=x^K$, and $z_y=z(y^K)$.
Define the matrix T* as in (1).

Step 3 Derive at x^K the linear approximations in (4) for f(x), h(x), and g(x), and set up the master program given by problem (M^K).

Step 4 Solve the master program (M^K):
[a] If a feasible solution y^{K+1} exists with objective value z_L^*;
set K=K+1, go to **Step 2.**

[b] If no feasible solution exists, stop.
Optimal solution is z_y at y\x\

It should be noted that the main advantage of this algorithm is that no elimination of equations is required to set up the master problem. Thus, the original sparsity of the MINLP is preserved in the MILP master problem, and difficulties associated with the selection of decision variables are avoided.

The algorithm as stated above assumes that the subproblems (NLP1) will have a feasible solution at step 2. When this is not the case, the simplest option is to resolve the master problem with an additional integer cut for the infeasible binary assignment. The other option is to add this integer cut and derive linear approximations at the infeasible point to set up the master problem in step 3. In this case however, the infeasible point must satisfy the nonlinear equations in the NLP subproblem in order to obtain the multipliers for the equality relaxation. To determine this point, the linear constraints and nonlinear inequalities of the NLP can be relaxed through the introduction of slack variables which would be added to the objective function so as to minimize the infeasibility.

Theoretical Properties of the OA/ER Algorithm

The relaxation of nonlinear equations to inequalities is based on the sign of the respective lagrange multipliers (matrix T^k) at the solution to the NLP subproblems. The master problem (M^k) is intended to provide a relaxed representation of the original MINLP problem which can be solved efficiently (i.e. a MILP problem). Further, the MILP is expected to overestimate the feasible region and underestimate the objective function (for a minimization problem) while providing a close approximation to the original problem.

The master problem formulation in (M^K) however, was stated without specifying the type of linear approximation to the nonlinear functions. The choice of the linear replacements must be based on the nature of the functions f(x), g(x), and $T^K h(x)$ in order to rigorously guarantee convergence to the global optimal solution of the MINLP. The following is a formal classification of problems that provides sufficient conditions for the type of linear approximations that can be used in the master problem to insure convergence to the MINLP global optimum (see Appendix B for details):

- Class 1 f(x) and $g_{IN}(x)$ are convex, $g_{EQ}(x)$ and $t_{ii}^k h_i(x)$ are quasiconvex $i=1,2,..r$; $k=1,2,..K$.

  For this class the NLP subproblems clearly exhibit a unique local solution (Bazaraa and Shetty, 1979). Moreover, first-order linearizations for f(x), h(x), and active g(x) at the point x^k provide supporting hyperplanes that satisfy (4), and hence insure rigorous lower bounds in the master problem.

Therefore, convergence to the global optimum solution can be guaranteed for this class. It should also be noted that for this class of problems the multipliers in equation (1) must remain invariant in sign for all k € K due to the assumed quasiconvexity in the relaxed equalities T^hM^O.

- <u>Class 2</u> f(x) and $g_{IN}(x)$ are convex or quasiconcave and g_{EQ} and $t_{ii}^{k}h_{i}(x)$ quasiconvex or quasiconcave i=1,2,..r; k=1,2..K.

  For this class valid linearizations for the master problem are obtained if first-order linearizations are used for convex / quasiconvex functions and linear-underestimators to replace quasiconcave functions. Although rigorous lower bounds will be provided by the master problem, the global optimum can only be guaranteed if the NLP subproblems exhibit a unique solution. This is not necessarily true for the class of functions assumed here.

- <u>Class 3</u> f(x), g(x), and $t_{ii}^{k}h_{i}(x)$ are undetermined.

  For this class unique solutions in the NLP subproblems are not guaranteed, and no linearization scheme can provide rigorous lower bounds. Therefore, global optimality for this class of functions cannot be insured.

Clearly, Class 1 problems are most favorable since first order linearizations can be derived easily, especially since gradient information is typically required by the NLP solution technique. The derivation of linear underestimate's in the Class 2 procedure is a less trivial task (see Duran, 1984) and a rigorous approach to Class 3 problems is currently unavailable.

A special subset of Class 1 problems is the case where f(x), g(x), and h(x) are strictly linear meaning that the original MINLP is actually a MILP problem. For problems of this type, the OA/ER algorithm would terminate in only two iterations (one iteration to find the solution and a second one to confirm it). One would therefore expect that the algorithm should perform well on "mostly linear" MINLP problems (problems in which many of the constraints and variables are linear). Since the formulation of structural flowsheet optimization problems often exhibits this characteristic, the Equality-Relaxation algorithm can be expected to be an efficient solution method for these problems.

It should also be noted from the above classification that rigorous solutions can only be guaranteed for Class 1 problems. For Class 2 problems rigorous solutions can also be guaranteed provided the NLP subproblems have a unique solution.

Unfortunately, most flowsheet synthesis problems will lie in Class 3. However, as will be shown later in this paper, numerical evidence indicates that applying the OA/ER algorithm to these problems as if they were of Class 1 will very often produce the global optimum solution. This would suggest that it would be worthwhile to solve Class 3 problems in two phases. The first one being through the procedure for Class 1 to provide a very good estimate of the MINLP solution. The second phase would attempt to verify this as the global optimal solution. This scheme will be explored in a future paper.

Example 1

In order to compare the properties of the proposed Equality-Relaxation algorithm with the Outer-Approximation algorithm involving algebraic elimination of equations (Duran and Grossmann, 1986b), consider the following small example:

$$
\begin{aligned}
\min\ & z = -y + 2x_1 + x_2 \\
\text{s.t.}\ & x_1 - 2\exp(-x_2) = 0 \\
& -x_1 + x_2 + y \leq 0 \qquad\qquad (P1)\\
& 0.5 \leq x_1 \leq 1.4 \\
& y \in \{0, 1\}
\end{aligned}
$$

The NLPs at the two integer values have the following solutions:

1. $y=0 \quad z=2.558 \quad x^T=(0.853, 0.853)$
2. $y=1 \quad z=2.124 \quad x^T=(1.375, 0.375)$

Thus, the second solution is the global MINLP optimum.

To examine the performance of OA/ER algorithm assume that the initial point is given by y=0. The solution to (P1) at y=0 is:

$$z=2.558 \qquad x^T=(0.853,0.853) \qquad X^1 = -1.619 < 0$$

The lagrange multiplier for the nonlinear equation is negative, therefore the direction matrix (1 by 1 matrix since only one nonlinear equation appears in problem) is given by $T^1 = -1$. Relaxing the nonlinear equality to inequality in (P1) yields:

$$T^1 \; h(x) \leq 0 \quad \Rightarrow$$

$$-1 \; \{ X_1 - 2 \exp(-x_2) \} \leq 0 \quad \Rightarrow$$

$$\text{or} \qquad 2 \exp(-x_2) - x_i \leq 0 \tag{6}$$

Note that the nonlinear constraint in (6) is convex, and thus the linearization at x^T=(0.853 , 0.853) will be a rigorous approximation for the master problem so that* the global optimal solution will be found. As an additional point, if y=1 is used as the initial point, the lagrange multiplier is again negative so that the inequality t_n h^x) will be the same as above. Since in this example all the functions are convex, the problem belongs to Class 1, and therefore the OA/ER algorithm will determine the global optimum solution of problem (PI).

To illustrate the difficulties which can arise when eliminating nonlinear equations in the Outer-Approximation algorithm, examine the resulting MINLP formulations upon elimination of the nonlinear constraint $x_i - 2 \exp(-x_2) = 0$ in (P1).

If x_i is eliminated then the reduced MINLP is:

$$\min \; z = -y + 4 \; exptrx_2) + x_2$$

$$\text{s.t.} \quad -2 \; \exp(-x_2) + x_2 + y \; £ \; 0 \tag{PA}$$

$$0.357 \leq x_2 \leq 1.386$$

$$/GO \; , \; 1$$

Note that the first inequality is nonconvex due to the exponential term with a negative coefficient.

If instead x_2 is eliminated then the problem formulation is given by:

$$
\begin{aligned}
& \min z = -y + 2x_1 - 2\ln(x_1) + \ln(2) \\
& s.t. \ -x_1 - \ln(x_1) + \ln(2) + y \le 0 \qquad (PB) \\
& 0.5 \le x_1 \le 1.4 \\
& y \in 0 \ , \ 1
\end{aligned}
$$

In this case the objective function and the first inequality are convex.

If problem (PA) (nonconvex formulation) is solved using the Outer-Approximation algorithm the following difficulty occurs. Assume the y=0 is the starting point. Linearizing the MINLP problem at the solution to NLP(y=0) yields the following master problem.

$$
\begin{aligned}
& z_L^1 = \min \mu \\
& s.t. \ 1.853\, x_2 + y \le 1.580 \qquad (7) \\
& -0.705\, x_2 - \mu - y \le -3.159 \\
& 0.357 \le x_2 \le 1.386
\end{aligned}
$$

For y=1, the first inequality reduces to $x_2 \le 0.313$ which is infeasible since x_2 has a lower bound of 0.357. Therefore, since the nonconvexity has cut off part of the feasible region and the master has no feasible region (for y=1), the global optimal solution will not be found. But if x_2 is eliminated, then the formulation is convex as seen in (PB), and in this case the OA algorithm will converge to the global optimal solution regardless of the initial value of y.

It is clear then, from this example, that the choice of decision and state variables can have a definite effect when implementing the Outer-Approximation algorithm by Duran and Grossmann (1986b). Clearly, the great advantage in the Equality-Relaxation variant is that there is no need to select decision variables nor to eliminate equations which as shown above can destroy the convexity of a problem.

Example 2

A small planning problem will be presented to compare the computational performance of the Equality-Relaxation algorithm with that of Generalized Benders Decomposition and to provide further insight into the former algorithm. A summary of the main steps in the GBD method is included in Appendix A.

Figure 2.a shows the selected superstructure which contains several alternatives for producing product C from raw materials A and/or B (e.g. build plant I only, plants I and III, plants I and II, none, etc.). The objective is to maximize profit given that there is an upper bound on the production of C. Data for this problem are given in Table 1. The model for the superstructure can be formulated as a MINLP problem as shown below:

$$\mathit{min}\ z = 3.5\, y^{1} + y'' + 1.5\, y^{m} + 7.0\, b' + b'' + 1.2\, A^{111} + 1.8\, a - 11.0\, c$$

$$\begin{aligned}
s.t.\quad & b'' - /M1 + a^{M}) = 0 \\
& b^{nx} - 1.2\ /I7{<}1 + a^{m}) = 0 \\
& c - 0.9\, A = 0 \\
& -b + b^{l} + b'' + l,''' = 0 \\
& a - a^{1\,1} - a''' = 0 \qquad\qquad (P2) \\
& b - 5/ \ \wedge\ 0 \\
& a'' - 5\, y^{II} \le 0 \\
& a^{1}'' - 5y^{m} \le 0 \\
& c\ '\le\ 1 \\
& b'' \le 5 \\
& V^{K}, y''\ V''\ G\ \{0.1\}^{3} \\
& a,\ a'',\ a^{m},\ b.\ b\backslash\ b''.\ b''\backslash\ c{*}0
\end{aligned}$$

The above formulation contains 3 binary variables, 8 continuous variables, 2 nonlinear equations, 3 linear equations, 3 logical constraints, and 2 upper bounds.

The problem was solved with OA/ER using function linearizations in the master problem and with the GBD algorithm. Two different starting points were used and in both cases the optimal solution found corresponds to y={ 1,0,1} with objective function z=-1.92 (10^3\$/hr). The optimal solution is shown in Figure 2.b. Table 2 shows the lower bounds predicted by the master problems of GBD and OA/ER; the * denotes convergence to the optimal solution.

As can be seen, for both starting points, GBD required five iterations whereas OA/ER required only two. Also note, as shown in Table 2 and Figure 3, that the OA/ER lower bounds are significantly tighter than the GBD lower bounds. Finally, it is interesting to see that the performance of the algorithms was unchanged by starting with the optimal point y=(1,0,1).

Since there are only two nonlinear functions in the superstructure formulation, it is worthwhile to examine the linearization and relaxation procedure used to derive the master problem in OA/ER. The two starting points behave similarly so only $y^{1=}$(1,1,0) will be presented in detail.

Referring back to the algorithm description (page 10), the first step is to choose the initial point, y^1=(1,1,0). By solving the NLP in step 2, the optimal objective function, $z(y^1)$, was -1.766 (10^3\$/hr) and the lagrange multipliers for the two nonlinear equations (X_1 and X_2) were 5.47 and 5.27.

In step 3 the master (MILP) is formulated, but first the direction matrix T^1 must be defined. Since both multipliers are positive, the diagonal elements are $t_{i\,1} = t_{22} = 1$ yielding T^1 h(x) ^ 0 as follows:

$$b^{u} - /nil + a^{11})\ \pounds\ 0 \tag{8}$$

$$b^{l}{}'' - 1.2\ Ind + a^{111})\ \pounds\ 0$$

Therefore since the relaxed constraint set is convex, problem (P2) belongs to Class i. Thus, the use of function linearizations in the master problem will guarantee the global optimal solution. Figure 4 shows the nonlinear constraint b^n-ln(1+a^M)=0 from the MINLP formulation and the resulting linearized inequalities which appeared in the master problem at iterations 1 and 2.

Example 3

The proposed Equality-Relaxation algorithm will be applied to the following structural flowsheet optimization problem. The superstructure is shown in Figure 5 and the problem data is given in Table 3. It should be noted that the superstructure of this example contains 16 basic alternative process schemes, each involving a number of continuous decisions variables that must be optimized.

As can be seen in Figure 5, there are several alternatives to produce product C from chemicals A and B. Two different feedstocks are available, both containing reactants A and B and inert material D; feedstock F2 is more expensive and contains less inerts than F1. The feed enters the process at low pressure and must be compressed to a higher pressure where reaction is feasible; either single-stage or two-stage compression with intermediate cooling can be selected. Unconverted raw materials are recycled since the reactor conversion per pass is low. The recycle stream must also be compressed because of pressure drops and the freedom in selecting the pressure for the flash separation. Again a choice between single and two-stage compression is available for the recycle stream. Another important choice for the recycle is the purge rate which avoids the accumulation of inert D, and has an impact on the overall conversion.

The exothermic gas phase reaction of A with B to produce C takes place adiabatically and is favored by high pressure and low temperature. A choice between a less expensive reactor of low conversion (R1) and an expensive reactor of high conversion reactor (R2) must be made. The reactor effluent stream is then sent to a flash separator where the lighter reactant and inert materials can be separated from the heavier product C. The bottom stream is the product stream and must contain at least 95% C; the market demand for this product is 86,400 kg-moles/day. The top stream is recycled with a portion purged to avoid inert build-up in the recycle loop.

The MINLP formulation of the superstructure of Figure 5 can be modelled with 4 binary variables (see Figure 5), 128 continuous variables (81 nonlinear and 47 linear), 111 equations (56 nonlinear and 55 linear), and 12 linear inequalities (8 logical constraints and 4 design specifications). The reactor was modelled with a simple correlation for conversion, the compressors assuming adiabatic compression, and the flash with the ideal model. The formulation of equations, inequalities, and objective function has been done according to the modelling techniques developed by Kocis and Grossmann (1986). These techniques have the effect of tightening the MINLP formulation and of providing greater consistency with the OA/ER algorithm. The logical constraints for existence of units were defined in terms of splits fractions with a lower bound of 6=0.01. This means that 99% of a stream is sent to the existent unit and 1% to the nonexistent process unit. In this way, it is possible to find the optimal operating conditions that nonexistent units and streams would take on if they existed in the flowsheet configuration being optimized.

The MINLP problem was solved with OA/ER assuming that the problem belongs to class 1; thus, function linearizations were used for the linear approximations in the master problem. The global optimum is y*={ 1,0,1,0} with objective function (-profit), z*=-2211.3 (10^3\$/year). The optimal process configuration shown in Figure 6 consists of the more expensive feedstock, single-stage feed and recycle compression, and high conversion reactor. Table 4 shows some of the relevant continuous variables of the optimal flowsheet.

Four starting points were investigated and each resulted in convergence to the global optimal solution in just two iterations. This means that only 2 of the 16 flowsheets in the superstructure had to be optimized. NLP subproblems were solved using the computer code MINOS / AUGMENTED (Murtagh and Saunders, 1980) on a DEC-20. The MILP master problems were solved using the package MPSX on an IBM-3083. Details of predicted lower bounds, initial points, and CPU times are given in Table 5. It is interesting to note that in the first three starting points the predicted lower bound from the master problem did not underestimate the optimal nonlinear objective function. This is possibly due to the problem being of class 3, while it is assumed to be of class 1 and where first order linearizations in the MILP master problem fail to satisfy (4). However, in all the cases the sign of the lagrange multipliers for the relaxed equalities remained unchanged.

It is important to note that the master problems of the first three initial points predicted the optimal assignment of binary variables in the first iteration. This means that each master provided a very good approximation of the original problem. Also, earlier in the paper, the larger OA/ER master problem (as compared to the master of GBD which contains only one inequality for each NLP subproblem solved) was justified on the basis that the MILP master problems are easier to solve than NLP subproblems. Table 5 clearly indicates that, for this example, this is the case since the NLP solution comprised approximately 93% of the total CPU time. Thus, a reduction in the number on NLPs which need to be solved will outweigh the additional CPU time required to solve a larger MILP master problem.

Structural Sensitivity Analysis

From the results presented in the examples 2 and 3 an important question that arises is the sensitivity of the optimal solution of the MINLP problem. The sensitivity of the binary variables is particularly relevant because it can provide some useful information on the sensitivity of the optimal flowsheet structure. As will be shown below, this information can be easily obtained from the results of the NLP subproblems.

Through the calculation of partial derivatives of the objective function value (z) with respect to the binary variables (y), one can estimate the effect that replacing or deleting a unit from the current flowsheet will have on the objective function. There are two means of extracting this information from the NLP solution. If a reduced gradient algorithm (e.g. MINOS/AUGMENTED) is used to solve the NLP subproblem with y^k constrained at a fixed value, then the sensitivity of a nonbasic binary variable is simply given by its reduced gradient $(\partial z/\partial y_i)_{y^k,x^k}$. If the binary variable is basic or a different NLP algorithm is used the sensitivity can be obtained through the lagrangian,

$$L(x^k, y, \mu^k) = c^T y + f(x^k) + (\mu^k)^T (C\, x^k - d + B\, y). \tag{9}$$

where x^k is the optimal solution of the NLP subproblem for fixed y^k and μ^k is the optimal Kuhn-Tucker multiplier associated with the mixed-integer inequalities. Note that (9) is identical to the cut used in the GBD method (see step 3, Appendix A). From (9) it clearly follows that

$$(\partial z/\partial y_i)_{y^k,x^k} = (\partial L/\partial y_i)_{y^k,x^k} = c_i + (\mu^k)^T b_i \tag{10}$$

where b_i is the i^{th} column of the matrix B. Thus, the change in the objective function value which occurs when the binary variable y_i is changed from y_i^* to $\bar{y}_i$ (e.g. from 0 to 1 or vice versa) can be estimated as follows:

$$\Delta z_i = (\partial z/\partial y_i)_{y^k,x^k} \, |\bar{y}_i - y_i^k| \qquad (11)$$

where $(\partial z/\partial y_i)_{y^k,x^k}$ is either given directly by the reduced gradient or otherwise calculated from equation (10). It should be noted that the predicted change Δz_i in general provides an overestimation of the change in the objective function value incurred when the binary variable y_i^* is switched to $\bar{y}_i$.

This structural sensitivity analysis procedure was applied to example 3 at the solution of two different NLP problems: y^1=(0,0,1,1) which is suboptimal with $z(y^1)$=-1561.2, and y^2=(1,0,1,0) which is the optimal value of the binary variables. Table 6 shows the predicted changes Δz_i when each i^{th} component of the binary variables in y^1 and y^2 take on opposite values.

First consider the data for y^1=(0,0,1,1) and notice that Δz_1 and Δz_4 are negative. This indicates that if y_1 is set to one or if y_4 is set to zero that there is potential for decrease in the objective function z. Since the partial derivative in (11) implies that the predicted change in z occurs with a change in only a single y_i, a switch in either y_1 or y_4 should be considered. Because Δz_1 is much larger in magnitude, it is logical to change y_1 from 0 to 1 yielding y=(1,0,1,1). As seen in Table 5 this actually decreases the objective function from -1561.2 to -2205.3 which is quite close to the global optimum solution. Interestingly, changing the values of both y_1 and y_4 to the respective opposing values yields y=(1,0,1,0) which is indeed the optimal binary assignment. It should be noted that the predicted values of Δz_i overestimate by a considerable amount the actual improvement in the objective function. Also, as shown in Table 6, for y^2=(1,0,1,0) all terms Δz_i were positive. Physically this means that sensitivity analysis at $y^{\#}$ predicts an increase in the objective function value for any change in y.

One has to keep in mind that although the above results for example 3 are encouraging, the structural analysis presented here has the limitations common to any sensitivity analysis procedure. Therefore, caution must be exercised in its interpretation.

Conclusions

A method for solving MINLP problems in which the binary variables appear linearly and the continuous variables appear linearly/nonlinearly has been developed with special attention given to the handling of nonlinear equality constraints. The proposed Equality-Relaxation algorithm does not require algebraic or numerical elimination of equations which is necessary with the Outer-Approximation algorithm proposed by Duran and Grossmann (1986b). Thus, difficulties such as possible destruction of convexity, loss of sparsity, and singularities are avoided. Theoretical properties for the Equality-Relaxation algorithm were discussed, as well as sufficient requirements on the nature of the nonlinear functions in the MINLP formulation for which the algorithm is guaranteed to find the global optimal solution.

The computational performance of the OA/ER algorithm was demonstrated on a small planning problem and on a larger flowsheet synthesis problem. Encouraging results were obtained in that in the former the OA/ER performed much more efficiently than the Generalized Benders Decomposition algorithm which is also capable of handling nonlinear equality constraints. In the latter problem, regardless of the initial point used, only two of the sixteen flowsheets embedded in the superstructure had to be optimized to find the global solution. Finally, a simple procedure for structural sensitivity analysis has been presented which can provide useful insights for the problem of structural flowsheet optimization.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support from the National Science Foundation under grant CPE-8351237.

References

[1] Andrecovich, M.J. and Westerberg, A.W.
An MILP Formulation for Heat-Integrated Distillation Sequences Synthesis.
AIChE J 31(9):1461-1474, 1985.

[2] Bazaraa, M.S., and Shetty, C.M.
Nonlinear Programming: Theory and Algorithms.
John Wiley and Sons, 1979.

[3] Beale, E.M.L.
Integer Programming.
The State of the Art in Numerical Analysis.
Academic Press, London, 1977, pages 409-448.

[4] Benders, J.F.
Partitioning Procedures for Solving Mixed-variables Programming Problems.
Numerische Mathematik 4:238-252, 1962.

[5] Douglas, J.M.
A Hierarchical Decision Procedure of Process Synthesis.
AIChE Journal 31(3):353-362, 1985.

[6] Duran, M.A.
A Mixed-integer Nonlinear Programming Approach for the Systematic Synthesis of Engineering Systems.
PhD thesis, Carnegie-Mellon University, 1984.

[7] Duran, M.A., and Grossmann, I.E.
A Mixed-Integer Nonlinear Programming Algorithm for Process Systems Synthesis.
AIChE Journal 32(4):592-606, 1986a.

[8] Duran, M.A., and Grossmann, I.E.
An Outer-Approximation Algorithm for a Class of Mixed-Integer Nonlinear Programs.
Mathematical Programming , 1986b.
in press.

[9] Floudas, C.A., Ciric, A.R., and Grossmann, I.E.
Automatic Synthesis of Optimum Heat Exchange Network Configurations.
AIChE J 32(2):276-290, 1986.

[10] Garfinkel, R.S., and Nemhauser, G.L.
Integer Programming.
John Wiley and Sons, New York, 1972.

[11] Geoffrion, A.M.
Generalized Benders Decomposition.
Journal of Optimization Theory and Applications 10(4), 1972.

[12] Grossmann, I.E.
A Mixed-integer Programming Approach for the Synthesis of Integrated Process Flowsheets.
Computers and Chemical Engineering 9(5):463-482, 1985.

[13] Hillenbrand, J.B.
Studies in the Synthesis of Energy-Efficient Evaporation Systems.
PhD thesis, Carnegie-Mellon University, 1984.

[14] Kocis, G.K., and Grossmann, I.E.
Effective Modeling Techniques for Structural Flowsheet Optimization Problems, manuscript in preparation (1986).

[15] Linnhoff, B.
Entropy in Practical Process Design.
foundations of Computer-Aided Chemical Process Design.
Engineering Foundation, New York, 1981, pages 537-572.

[16] Little, J.D.C., Murty, K.G., Sweeney, D.W., and Karel, C.
An Algorithm for the Traveling Salesman Problem.
Operations Research 11:979-989, 1963.

[17] Mangasarian, O.L.
Nonlinear Programming.
McGraw-Hill, New York, 1969.

[18] Murtagh, B.A. and Saunders, M.A.
A Projected Lag rang i an Algorithm and its Implementation for Sparse Nonlinear Constraints, and MINOSIAUGMENTED User's Manual.
Technical Report SOL 80-1R and SOL 80-14, Systems Optimization Laboratory, Department of Operations Research, Stanford University, June, 1980.

[19] Nishida,N., Stephanopoulos,G.,and Westerberg, A.
A Review of Process Synthesis.
AlChE Journal (27):321-351, 1981.

[20] Papoulias, S. and Grossmann, I.E.
A Structural Optimization Approach in Process Synthesis, I, II, III.
Computers and Chemical Engineering 7(6), 1983.

[21] Shelton, M.R. and Grossmann, I.E.
Optimal Synthesis of Integrated Refrigeration Systems-Part 1: Mixed-Integer Programming Model.
1986.
to appear in Computers and Chemical Engineering (1986).

[22] Stephanopoulos, G. and Westerberg, A.
Studies in Process Synthesis, Part II, Evolutionary Synthesis of Optimal Process Flowsheets.
Chemical Engineering Science 31:195-204, 1976.

[23] Vaselenak, J.A.
Studies in the Retrofit Design and Scheduling of Batch Processes.
PhD thesis, Carnegie-Mellon University, 1985.

APPENDIX A. Generalized Benders Decomposition Algorithm

GBD algorithm as applied to problem (P) and assuming feasible NLP subproblems is presented below:

<u>Step</u> 1 Select initial binary assignment y\ set K=1.
Initialize upper bound z^u=oo.

<u>Step</u> <u>2</u> Solve the NLP subproblem:

$$z(y^K) = \min_x \; c^T y^K + f(x)$$

$$\begin{aligned} s.t.\;\; & h(x) = 0 \\ & g(x) \leq 0 \\ & A\,x=a \\ & C\,x \leq d - B\,y^K \\ & x\,G\,X \end{aligned} \qquad \text{U-D}$$

yielding $z(y^K)$. x^K, and p^K (multipliers for constraints Cx^d-By^K).
If $z(y^k)<z^u$, then set y*=y\ x*=x\ and $z^u=z(y^k)$.

<u>Step</u> <u>3</u> Formulate the pseudo-integer master problem:

$$z_B^K = \min \; /'B$$

$$\begin{aligned} s.t.\;\; & //_B \; Ss \; L(x^k,y,/) \qquad k=\text{-}\backslash,2....K \\ & y\,G\,Y \\ & \mu_B \in R^1 \end{aligned} \qquad \text{U-2)}$$

where L(x\y./>^{k}) = c^Ty • f(x^k) + (/>^{k})^T (C x^k - d + B y).

<u>Step</u> <u>4</u> Solve the master problem (A-2) yielding z^ and y**^{1}.
If z^=z^u stop, solution is y", x*; else set K=K+1, go to Step 2.

APPENDIX B. Derivation of Linear Approximations

The master problem is intended to provide a linear approximation which overestimates the original MINLP feasible region and underestimates the objective function. This will guarantee that the master solution will be a lower bound on the MINLP solution. Deriving linear estimators which have the following property will provide the desired approximation

$$(\mathbf{w}^k)^T \mathbf{x} - w^k{}_o \leq f(\mathbf{x}) \qquad (B-1)$$

$$\mathbf{T}^k\ \mathbf{h}(\mathbf{x}) \ \leq\ 0 \Rightarrow \mathbf{T}^k\ [\mathbf{R}^k\ \mathbf{x}\ -\ \mathbf{r}^k]\ \leq\ 0 \qquad (B-2)$$

$$\mathbf{S}^k_{IN}\,\mathbf{x} - \mathbf{s}^k_{IN} \leq \mathbf{g}_{IN}(\mathbf{x}) \qquad (B-3)$$

$$\mathbf{g}_{EQ}(\mathbf{x})\ \leq\ 0 \Rightarrow \mathbf{S}^k_{EQ}\ \mathbf{x}\ -\ \mathbf{s}^k_{EQ}\ \leq\ 0 \qquad (B-4)$$

If the nonlinear objective function term, f(x), is convex and if the nonlinear constraints $\mathbf{T}^k\mathbf{h}(\mathbf{x})\leq 0$ and $\mathbf{g}_{EQ}(\mathbf{x})\leq 0$ are quasiconvex and $\mathbf{g}_{IN}(\mathbf{x})\leq 0$ are convex, then first-order linearizations at $\mathbf{x}^k$ will satisfy conditions (B-1)-(B-4) as shown below.

A first order Taylor series approximation of f(x) about the point $\mathbf{x}^k$ is:

$$f(\mathbf{x})\ \approx\ f(\mathbf{x}^k)\ +\ \nabla f(\mathbf{x}^k)^T\ [\mathbf{x}\ -\ \mathbf{x}^k] \qquad (B-5)$$

meaning that $\mathbf{w}^k$ and $\mathbf{w}^k_o$ are replaced by $\nabla f(\mathbf{x}^k)$ and $\nabla f(\mathbf{x}^k)^T[\mathbf{x}^k]-f(\mathbf{x}^k)$, respectively. Similarly, the set of nonlinear constraints are approximated by:

$$\mathbf{h}(\mathbf{x})\ \approx\ \mathbf{h}(\mathbf{x}^k)\ +\ \nabla \mathbf{h}(\mathbf{x}^k)^T\ [\mathbf{x}\ -\ \mathbf{x}^k] \qquad (B-6)$$

$$\mathbf{g}(\mathbf{x})\ \approx\ \mathbf{g}(\mathbf{x}^k)\ +\ \nabla \mathbf{g}(\mathbf{x}^k)^T\ [\mathbf{x}\ -\ \mathbf{x}^k]$$

which with (B-2), (B-3), and (B-4) lead to

$$\mathbf{R}^k\ =\ \nabla \mathbf{h}(\mathbf{x}^k)^T \qquad \mathbf{r}^k\ =\ \nabla \mathbf{h}(\mathbf{x}^k)^T\ [\mathbf{x}^k] \qquad (B-7)$$

$$\mathbf{S}^k\ =\ \nabla \mathbf{g}(\mathbf{x}^k)^T \qquad \mathbf{s}^k\ =\ \nabla \mathbf{g}(\mathbf{x}^k)^T\ [\mathbf{x}^k]\ -\ \mathbf{g}(\mathbf{x}^k)$$

A differentiable function f(x) is convex (Bazaraa and Shetty, 1979) if and only if the following condition is satisfied for any $\bar{x}$:

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T [x-\bar{x}]$$

When the coefficients w^k, w^k_o, S^*_{IN}, and s^k_{IN} are selected as stated above, this condition is clearly identical to (B-1) and (B-3) for f(x), $g_{IN}(x)$, respectively.

It was claimed that condition (B-2) and (B-4) would be satisfied by function linearizations for $h(x) \leq 0$ and $g_{EQ}(x) \leq 0$ quasiconvex, a weaker restriction than convexity. As shown in Mangasarian (1969), a differentiable function, h(x) (or $g_{EQ}(x)$) is quasiconvex if and only if one of the following equivalent statements hold:

$$\text{if } h(x) \leq h(\bar{x}) \text{ then } \nabla h(\bar{x})^T [x-\bar{x}] \leq 0$$

$$\text{if } \nabla h(\bar{x})^T [x-\bar{x}] > 0 \text{ then } h(x) > h(\bar{x})$$

Since the point of linearization ($\bar{x} = x^k$) satisfies the nonlinear constraints h(x)=0, then the first condition can be written as:

$$\text{if } h(x) \leq 0 \text{ then } \nabla h(\bar{x})^T [x-\bar{x}] \leq 0$$

This is precisely the condition which must be satisfied by the linear approximation in (B-2) (i.e. that any x which satisfies $h(x) \leq 0$ also satisfies $\nabla h(\bar{x})^T [x-\bar{x}] \leq 0$). A similar reasoning applies for $g_{EQ}(x) \leq 0$ in (B-4).

APPENDIX C. Numerical Elimination and Loss of Sparsity

Given fixed values for the binary variables in the MINLP formulation, the objective of this appendix is to compare the sparsity patterns in the full NLP (NLP problems with no elimination) and reduced NLP problems (NLP problems with numerical elimination of equations). Consider the incidence matrices in Figure C-1 for the continuous variables in the full NLP and in the reduced NLP following numerical elimination- The original NLP has n_V continuous variables, n_I inequalities, n_E equalities, and 2 n_V simple lower and upper bounds. If r is the average number of variables which appear in the equations and inequalities, it is easy to show that the number of nonzero elements in the incidence matrix of the full NLP is given by:

$$N'_{NZ} = r \times (n_E + n_I) + 2 \times n_V$$

By assuming that, due to the numerical elimination, the incidence matrix in the reduced NLP is full except for simple bounds on the n_V-n_E decision variable, the number of nonzero elements is given by:

$$N''_{NZ} = (n_V - n_E) \times (n_E + 2 n_I) + 2 (n_V - n_E)$$

To illustrate the difference, assume that n_V=1050 (the number of continuous variables), n_I=n_E =1000 (number of inequalities and equalities), and r=5. Then the number of nonzero elements in the full NLP will be 12,100 versus 150,100 in the reduced NLP. Thus, in this example the number of nonzero elements is increased by more than **a** factor of ten as a result of the numerical elimination of the equality constraints.

Table 1. Problem Data of Example 2

OBJECTIVE FUNCTION

Investment and Operating Costs

	FIXED COST (10^3\$/hr)	VARIABLE COST (10^3\$/ton product)
Process I	3.5	2.0
Process II	1.0	1.0
Process III	1.5	1.2

Raw Material Costs (10^3\$) 1.8/ton A, \$7.0/ton B

Revenue \$13.0/ton C with maximum market demand = 1 ton/hr

Z = objective function = costs - revenue (10^3\$/hr)

$$= (3.5y' + 2c) + (y'' + b'') + (1.5 y^{1''} + 1.2 Z>^{m}) + 1.8 a + 7.0 A^{1} - 13 c$$

$$= 3.5/ + y'' + 1.5 y^{m} + 1.8 a + 7.0 A^{1} + Z>^{M} + 1.2 b^{m} - 11.0 c$$

MASS BALANCES

Process II: $b^{11} = \ln(1 + a'')$

Process III: $b^{m} = 1.2 \ln(1 + a^{MI})$

Process I: $c = 0.9 b$

Split to II and III: $a = a^{11} \bullet a^{m}$

Mixer before I: $\mathbf{b} = b^{1} + b^{11} + b^{IM}$

VARIABLE BOUNDS

Limit on process II: b^{11} ^ 5

Market demand for c: c ^ 1

Table 2. Lower bounds predicted by GBD and OA/ER in Example 2.

		ITERATION				
		1	2	3	4	5
$y^1=(1,1,0)$	GBD	-27.33	-23.83	-11.85	-2.72	-1.92 *
	OA/ER	-3.71	*			
$y^1=(1,0,1)$	GBD	-16.86	-13.36	-11.85	-2.72	-1.92 *
	OA/ER	-3.98	*			

Table 3. Flowsheet Synthesis Problem **Data**

FEEDSTOCK	COMPOSITION	COSTS
F1	60% A	\$0.026/kg-mole
	25% B	
	15% D (inert)	
F2	65% A	\$0.033/kg-mole
	30% B	
	5% D (inert)	

PRODUCT/BY-PRODUCT		
Product P	£95% C	\$0.25/kg-mole
	(≤86,400 kg-mole/hr)	
Purge P_{BY}		\$0.021/kg-mole

UTILITIES	COSTS
Electricity	\$0.03/kw-hr
Heating (steam)	$\$8.0/10^6$ kJ
Cooling (water)	$\$0.7/10^6$ kJ

DESIGN SPECIFICATIONS	
REACTOR Pressure (MPa)	2.5 ^ PR ^ 15
Temperature (inlet, K)	$423 \leq T^{IN} \leq$, 873
Temperature (outlet, K)	523 ^. T^{OUT} £ 873
FLASH SEPARATION	
Pressure (MPa)	0.15 ≤. PF £ 15
Temperature (K)	300 ≤. TF ≤• 500

Table 4. Continuous Variables in Optimal Flowsheet

FLOWRATES	
F2	307,940 kg-moles/day
P	86,400 kg-moles/day
P_{BV}	53,020 kg-moles/day
REACTOR	
PR (MPa)	8.12
T_{OUT} (K)	557.5
	423.0
Conversion of B (per pass)	24.21%
FLASH SEPARATION	
PF (MPa)	7.614
TF (K)	411.5
Overall Conversion of B	91.21%
Purge Rate	2.53%
UTILITIES	
ELECTRICITY (kW)	
Feed Compressor	227.83
Recycle Compressor	27.97
HEATING (steam, 10^9 kJ/year)	
Product Stream	0.608
Purge Stream	1.816
COOLING (water, 10^9 kJ/year)	23.95

Table 5. OA/ER Results for Flowsheet Synthesis Test Problem.

ITERATION	y^K	NLP $z(y^K)$	MILP z_L^K
1	{0,1,1,0}	-1541.6	-1980.
2	{1,0,1,0}	-2211.3	infeasible
CPU time (seconds): 141.4 (NLP: 134 MILP: 7.4)			
1	{0,0,1,1}	-1561.2	-1979.4
2	{1,0,1,0}	-2211.3	infeasible
CPU time (seconds): 153.8 (NLP: 147 MILP: 6.8)			
1	{1,0,1,1}	-2205.3	-2175.7
2	{1,0,1,0}	-2211.3	infeasible
CPU time (seconds): 119.2 (NLP: 108 MILP: 7.2)			
1	{1,0,1,0}	-2211.3	-2206.3
2	{1,0,1,1}	-2205.3	infeasible
CPU time (seconds): 119.2 (NLP: 108 MILP: 7.2)			

Table 6. Sensitivity Analysis Data for Example 3 Problem

(a) y^1

i	y_i^1	$\bar{y}_i - y_i^1$	$(\partial z/\partial y_i)_{y^1, x^1}$	Δz_i
1	0	1	-2520.7	-2520.7
2	0	1	25.3	25.3
3	1	-1	-1284.3	1284.3
4	1	-1	5.9	-5.9

(b) y^2

i	y_i^2	$\bar{y}_i - y_i^2$	$(\partial z/\partial y_i)_{y^2, x^2}$	Δz_i
1	1	-1	-7079.5	7079.5
2	0	1	25.2	25.2
3	1	-1	-792.6	792.6
4	0	1	6.0	6.0

Figure 1. MINLP Algorithm Schematic

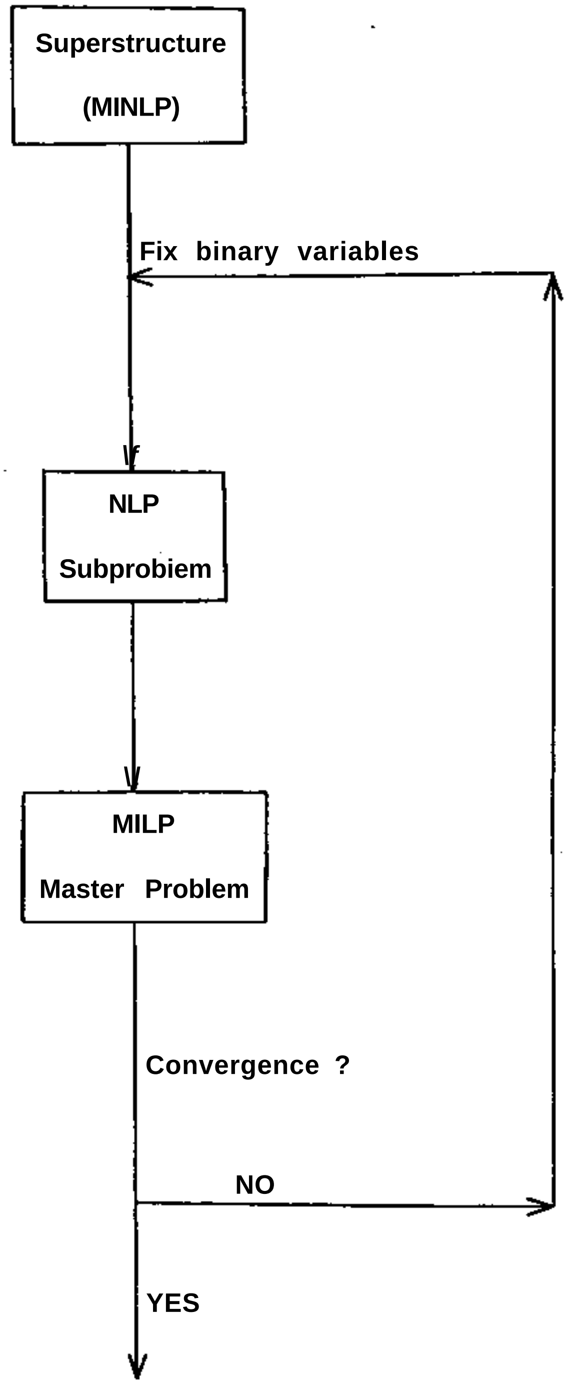
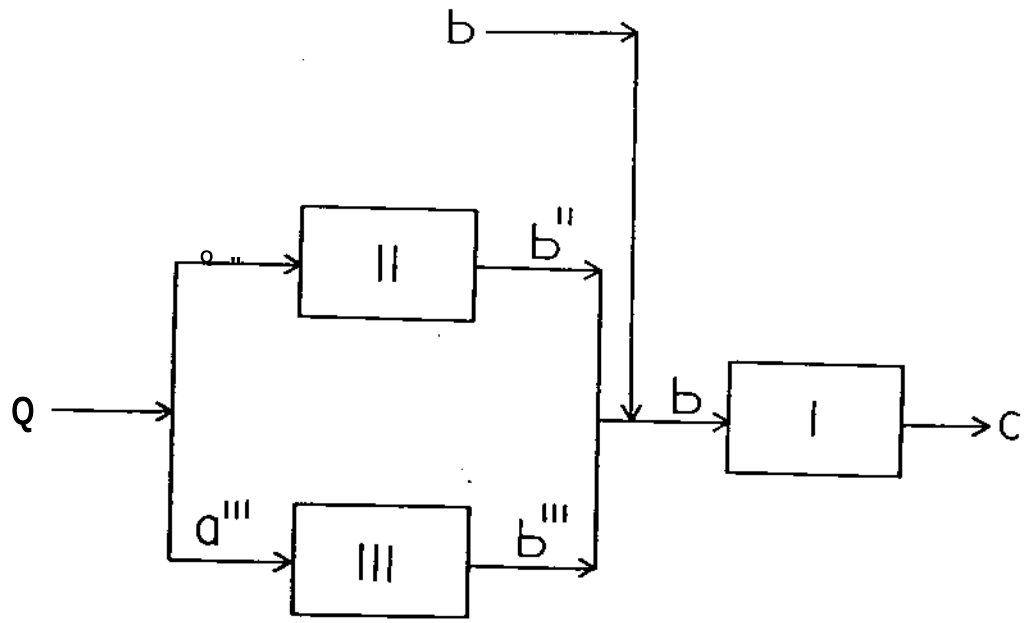


Figure 2. Superstructure and Solution for Problem (P2)

(a) Superstructure



(b) Optimal Process Scheme

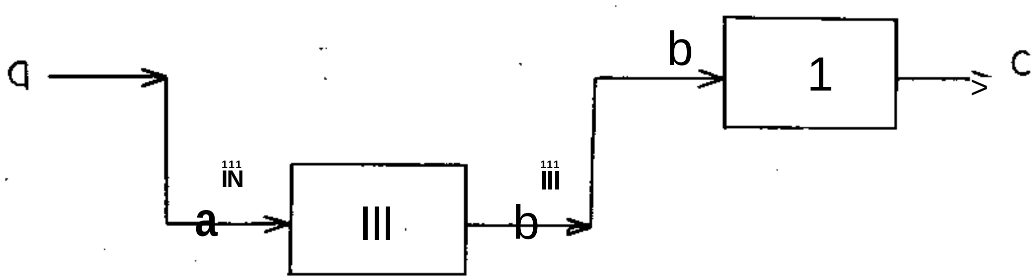


Figure 3. Lower Bounds Predicted by GBD and OA/ER for y^1=(1,1,0)

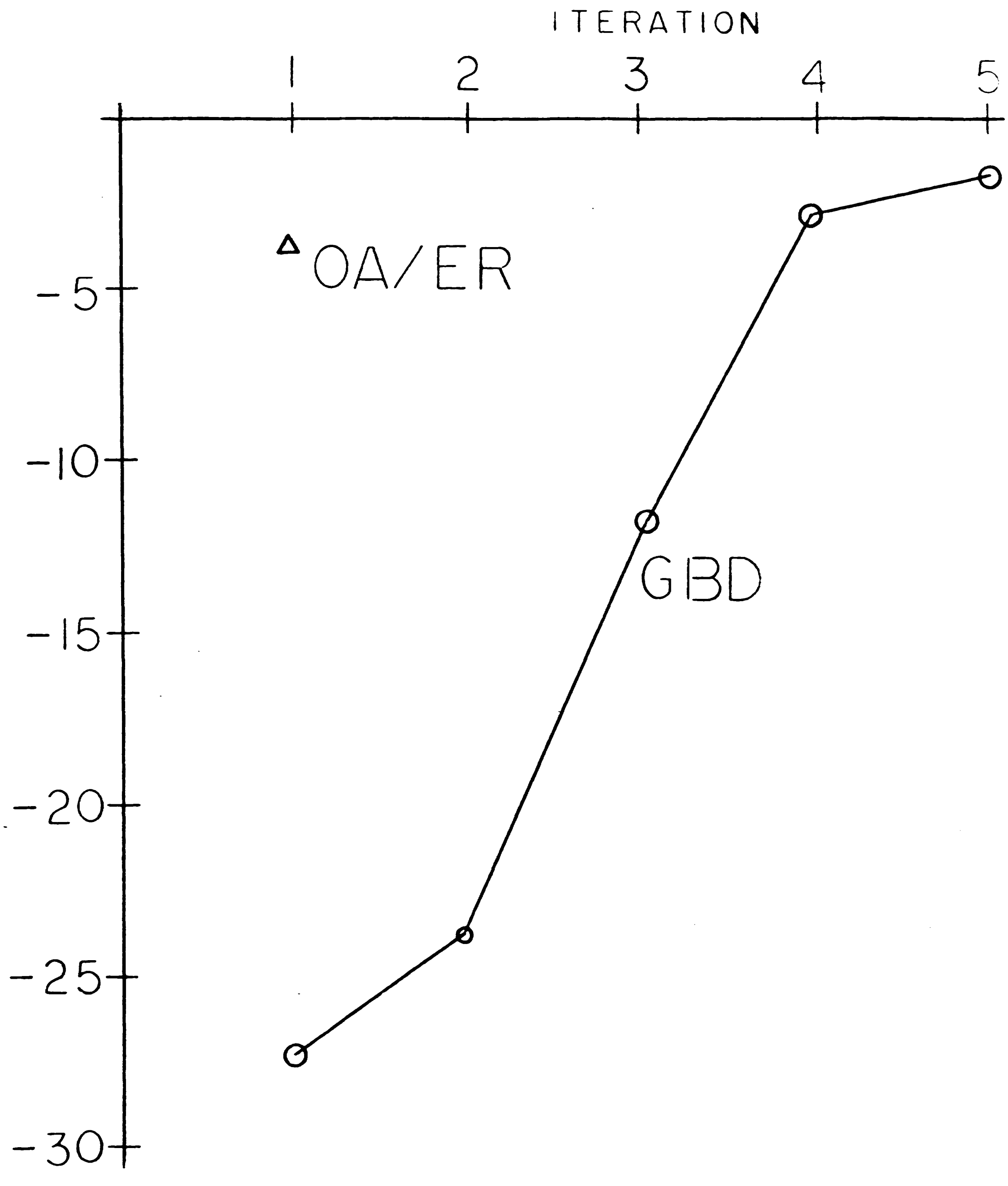


Figure 4. Linear Inequality Approximation to Nonlinear Equation

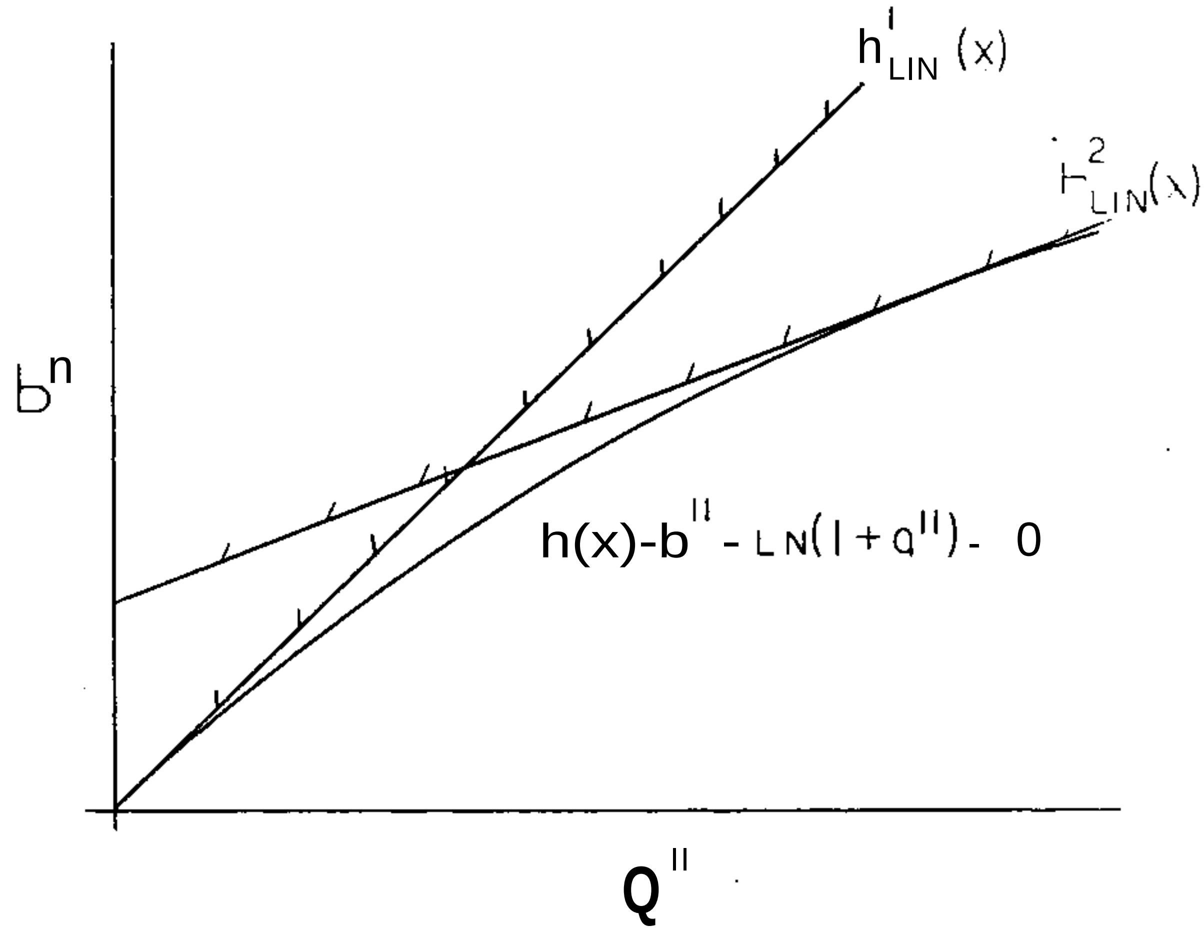


Figure 5. Flowsheet Superstructure

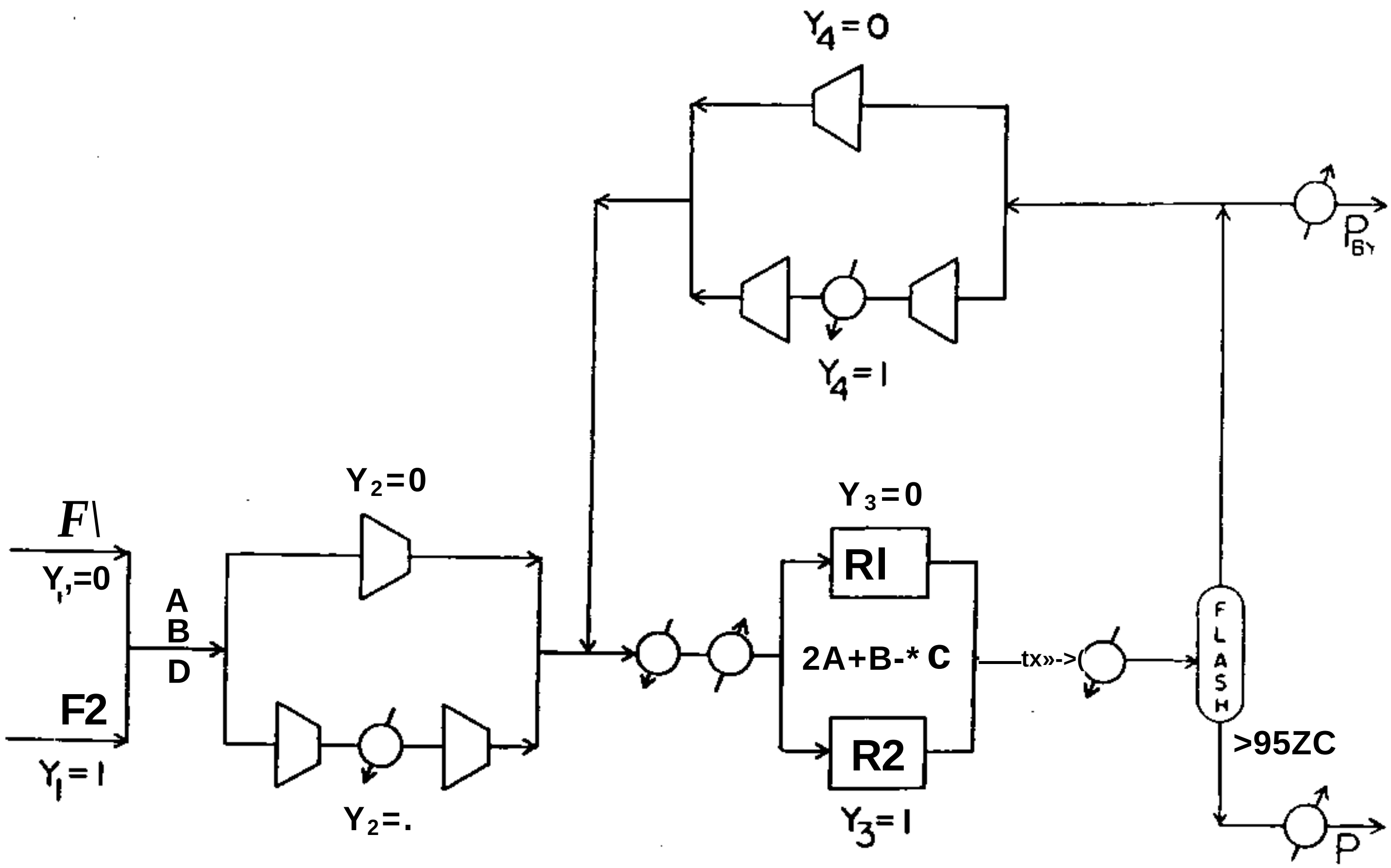


Figure 6. Optimal Process Flowsheet

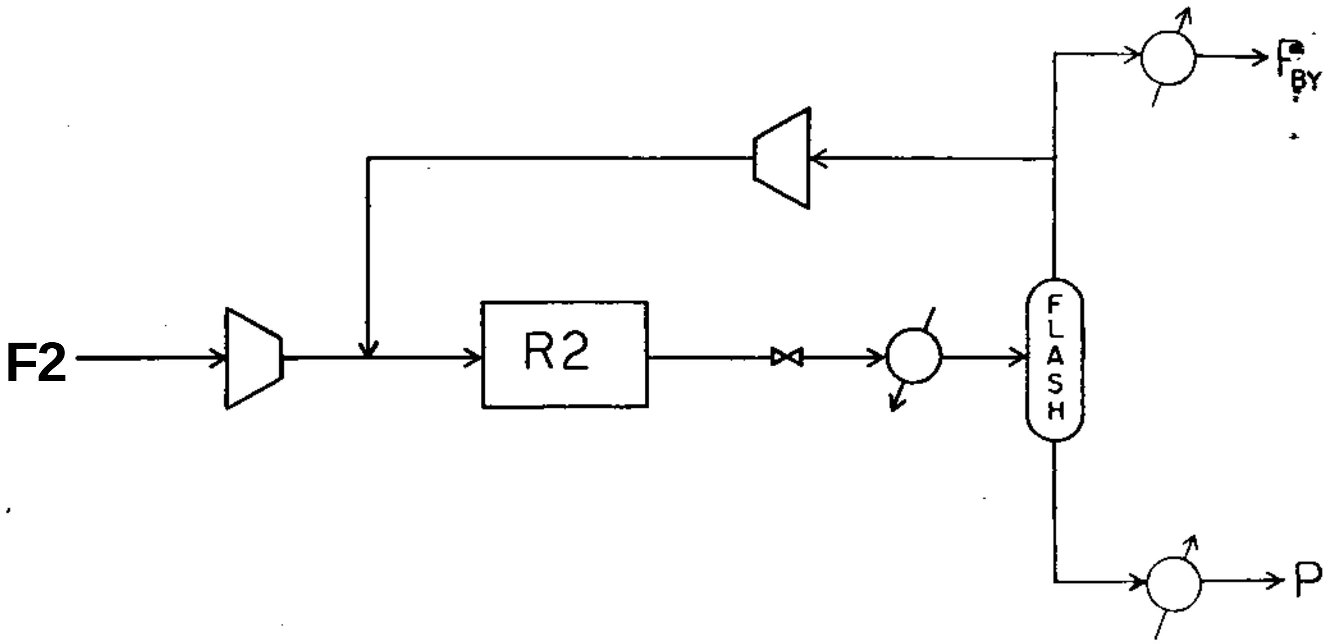


Figure C-1. Incidence Matrices for Full and Reduced NLP