On the Simultaneous Solution and Optimization of Large-Scale Engineering Systems

by

L. T. Biegler

EDRC-06-26-87 3

ON THE SIMULTANEOUS SOLUTION AND OPTIMIZATION OF LARGE-SCALE ENGINEERING SYSTEMS

L. T. Biegter
Department of Chemical Engineering
Carnegie-Mellon University
Pittsburgh, PA 15213

The application of recently developed process optimization strategics is motivated and reviewed from a Newton perspective. Over the past fifteen yean Newton-type methods have been used to solve difficult wm^iatirpu problems. Similar principles and advantages also extend to process optimization. Using the Newton framework we discuss recent improvements to nonlinear pn?tfra"wftff.y algorithms and the application to large-scale systems. The improvements in the optimization strategy also lead as to consider new and more difficult IMOOCSS probwcms* Rcocot results idstms to tftrpc of uicse*

- finMlltMKHWW QPtlfP'T^Tffl** SDd tlCSt IHfrgFUWW

opmniZattioti oi processes ^wni uiscvGotutt cotutioii models

- sensitivity of optimal flowsheets

will be reviewed. Finally future directions will be riwrasysri thai address the solution of brgcr aiid nxxc difficult ()pumizaiJoo problems.

L Introduction

Over the past five years the application of the Successive Quadratic Programming (SQP) nonlinear programming algorithm by a number of researchers [7,13,43,44] has led to renewed interest in flowsheet optimization as an efficient tool for process design and operation. The key concept to its effectiveness is SQP*s ability to converge equality constraints as part of the optimization problem. Through the application of Newun-like strategies for process simulation [14,53,60], limc-consuming solution procedures for convergence of nested recycle and calculation loops, could now be substituted by simultaneous methods that easily account for the interactions among different loops. Process optimization can be seen as representing the integration of another calculation loop; accounting for its "interaction" with the rest of the process represents the contribution of the SQP strategies. In a recent review paper [4], parallels were drawn between the simultaneous simulation and optimization problem and a number of studies were summarized for both problems. Here we use the well-known properties of the Newton method and more recently discovered properties of SQP to review the work of the last five years from a different perspective. Moreover, an understanding of SQP's convergence properties also allows us to analyze newly developed decomposition methods for large-scale problems and develop guidelines for their further improvement.

For flowsheet optimization with conventional process simulators, most recent studies [15,43,44,46] report some experimentation to determine tuning parameters to improve problem specific aspects or optimization, such as variable scaling, flowsheet initialization and the use of partial flowsheet convergence. This experimentation may always be needed for better performance on specific problems, but it is our hope that better theoretical understanding will eliminate many of the "tuning" parameters and lead to important generalizations for flowsheet optimization. As an example of this, recent developments in partial convergence algorithms will be highlighted.

The capability to do process optimization quickly allows us to formulate and solve many problems that could not be considered before. Here, we expand the simultaneous concept to consider problems with a wider scope. For example, in parallel with algorithms for process optimization, very powerful strategies have been developed for heat integration [39,48]. Unfortunately, these only apply to problems with fixed flows and temperatures, e.g., after the process has been "optimized". This paper *wmamrirr** recent results for *simultaneous* heat integration and process optimrniion using process simulators. This simultaneous concept is also extended to process flowsheets with differential equation models, such as reactor units. As self-contained modules these are often expensive to evaluate and are usually not part of a flowsheet optimization study. On the other hand, interaction between the reactor and the rest of the flowsheet is frequently the most significant aspect of an optimfiation problem. Here we outline a simultaneous approach using orthogonal collocation on finite elements. and summarize its performance on a small problem.

In addition to extending the simultaneous concept to more difficult optimization problems, one is often interested in the *sensitivity* of the optimal solution to parametric and, possibly, modelling changes. As with Newton-type methods for simulation, much of this sensitivity information is already available from the optimization algorithm. Therefore, it is straightforward to quickly analyze which variables are locally sensitive to nominally assigned parameters, and for which others there will be little change.

The last section on future issues deals primarily with the solution of larger and more difficult optimization problems. To this end, development and evaluation of large-scale decomposition strategies seem to be an important issue. Also the incorporation of problem structure to improve performance will be a key area in handling more difficult problems. The brief review in this paper shows that the past five years have demonstrated the feasibility and effectiveness of these simultaneous concepts. Future research will help tailor these simultaneous concepts to develop more powerful algorithms for more difficult problems.

2. Newton Derivation of SQP

Consider for simplicity an equality constrained optimization problem of the form:

$$\begin{array}{ll} \text{Min} & F(z) \\ z & \\ \text{s.t.} & h(z) \end{array} \qquad (1)$$

The first order necessary optimalily conditions for this problem ar;; straightforward:

$$V_2L(z\ .\ v^*)\ -V_zF(z^*)\ +\ V_2h(z)v^*\ -\ 0$$

$$h(«^*>\ -\ 0 \qquad (2)$$

and a Newton method applied 10 these conditions leads to:

$$\begin{bmatrix}\nabla_{zz}L(z^{i}, v^{1}) & V^\wedge td^{1})j\ fAx \\ \nabla_{z}h^{T}(z^{i}) & 0\ JU\end{bmatrix} -$$

$$- \begin{pmatrix}\nabla_{z} L(z^{i}, v^{1}) \\ h(z^{i})\end{pmatrix} \qquad (3)$$

Here L(z,v) is the Lagrange function and v is the vector of Lagrange multipliers for h. At the optimum, (z\v*) second order necessary conditions require that the projection of V ^ z * . v*) into the null space of V_s h(z*) be positive sen*4eifinite. Therefore, for some region around the optimum, the Newton step is equivalent to solving the following quadratic program:

$$\bullet\underset{d}{in} \quad V_2F(*^{i})^{T}d + \tfrac{1}{2}d^{T}\nabla_{zz}L(z^{i},v^{i})d \qquad (4)$$

$$a.c. \quad hCz^{1}) + \nabla h(z^{i})^{T}d - 0$$

where d is the search dkectJon *in* z.

With this quadratic programming equivalence the Newton concept is I also easily extended to include inequality constraints. g(z) *˙ a In addition to computing the Newton step, by including inequality constraints, the quadratic program win identify the correct active set and satisfy additional complementarity conditions given by the Kuhn-Tocker conditions. For most engineering problems, however, gradient evaluation requires ?**fa*r*ti* work and second derivative information is usually too expensive to be calculated. It is therefore desirable to replace V_n L(z,u,v) by a suitable positive definite approximation matrix, B\ derived from differences in gradients from point 10 point. ! The following quadratic program (QP) is then solved ac each iteration, i, to determine the "Newton step" foroptimrration.

$$\bullet\underset{d}{in} \quad \nabla_{z}F(z^{i})^{T}d + JdVd$$

$$s.t. \quad g(z^{i}) + \nabla g(z^{i})^{T}d * 0 \qquad (5)$$

$$hCz^{1}) + \nabla h(z^{i})^{T}d - 0$$

In parallel with Newton's method, a number of properties can be staled to characterize the performance of the optimization algorithm. Thfsc can be summarized by the following categorici.

Local Convergence Properties

For solving nonlinear. equations. h(z) = 0, Newton's method converges at a Q-quadratic rate in the neighborhood of the solution, i.e:

$$||z^{i+1} - ,*|| - (Kll.^{1} - z*||^{2}) \qquad (6)$$

Quasi-Newton methods, on the other hand, are known to converge at a superlinear rate, and because they do not require derivatives at each iteration, can have significant advantages over Newton's method. For the SQP optimization algorithm. Garcia-Patoniares and Mangasarian [32] showed that if the actual Hessian matrix, $V*_a$ L(z,u,v), is used in the QP. convergence is. of course, also quadratic in z and v. On the other hand, if a positive definite quasi-Newton approximation is substituted for the Hessian, convergence rates are expected to be lower. Han [351, and later Boggs, et al [11], showed that if the actual Hessian is positive definite at the solution (a very restrictive assumption) then SQP converges at a Q-superiinear rate. A simpler proof of this property was later given by Nocedal and Overton [58]. In general, however, only the *projected* Hessian is positive definite and using a positive definite approximation for the/utf Hessian may lead to slower convergence. Powell [62] showed that by using an update formula that only approximates the projected part of the Hessian accurately, the convergence rate is actually two-step superlinear, i.e:

$$\underset{i\to\infty}{Urn} \|z^{i+1} - z^{*}\| / \|z^{i-1} - z^{*}\| - 0 \qquad (7)$$

This implication b important because it means that for general nonlinear problems only the two step superlinear convergence rate can be proved. Moreover, in a recent study, Nocedal and Overton [58] show that if one updates only the projected Hessian (the smaller positive definite part), convergence OCCUR at the same rate. This may lead to a strong motivation and justification for decomposition strategies applied to large-scale optimization problems.

Of course* rales of convergence are only assymptotic properties and other acton such as appropriate matrix miriilJTation (as for Broydcn's method) and variable scaling will have significant effects on pcifouuanrft. Nevmhclcn, convergence rate wunams an important guideline that can be used to motivate more efficient strategies.

GlobalConvergemce

While Newton's method for nonlinear equations has local quadratic convergence properties, it is not guaranteed to converge from poor starting points. To prove global convergence, a unique Newton step must be calculable at each iteration (the Jacobian must be nonsingular) and a tnTr*-* decrease should be found at each iteration for some roerkftmctk»,e^the2-rjormc/thefuncti^ values [65]. Fortunately, the Newton step gives a descent direction for this merit function and thus one has a guarantee (for smooth functions with no singularities) that a stepsize exists which progresses toward the solution. In the nrighborhood of the solution, one, of course, expects full Newton steps to be taken for fast convergence.

With SQP one can draw similar analogies to Newton's method in order to ensure convergence from poor starting points. Han [36] showed that the QP solution gives a descent direction for the-exact penalty function:

$$P(z,n) - F(x) + n(Eg_{j}(z)_{+} + E|h_{j}(z)|) \qquad (8)$$

$$\text{where} \quad gjU)_{+} - nax(O,gj(z))$$

Ensuring a sufficient decrease of this function will lead to global convergence as long as the QP is solvable at each iteration. Several researchers, however, recognized that this condition may lead to slow convergence in the neighborhood of the optimum (the "Maratos effect" [51]). because full steps were not taken. Chamberlain ct al [12] proposed a modified exact penalty line search (the "watchdog technique") where a sufficient decrease was required over a sequence of iterations. Other merit functions and line search strategics have also been proposed [10, 22, 64] Here, the augmented Lagrange function appears to be a suitable choice because it is difTcrcnúabk and has the global convergence property together with allowing full steps in the neighborhood of the optimum [6, 66]. We have modified this merit function to the following form:

$$L*(z,u,v) - F(z) + u^{T}g(z)_{+} + v^{T}h(z) + \qquad (9)$$

$$\frac{\eta}{2}\, \|g(z)_{+},h(z)\|^{2}$$

and developed an adaptive strategy for choosing ihc penally parameter, *r*\. Compared to the exact penalty and Watchdog ^(cpsi/c proccdyna we lave found ntesttMcgylo consistently give as good as or better performance on literature and process optimization problems [61. Use of this line search function also solves the pathological j examples that cause\ ime exact penalty function to faiL Finally, the augmented Lagrangian st*egy is straightforward to impknient with few adjustable panmeters.

SiAfuiarity

Lastly, we mention that the main drawback to Newton's method, encountering singular lacobian inatrices, sJso hv its parallels in SQP. This occurs when the feasible region is inconsistem and the QP has no solution. As with Newton's method numerous "safeguards" are available that attempt 10 move away from singular points [52,61,65]. In an analogous fashion, safeguards exist for SQP [61 that have the effect of relaxing the feasmfe region and BUS solving an altered QP. However, none of these rHswaffun techniques avoid failure for all problems. Consider, for example, the problem shown in Figure 1. Here, starting at or below point B and applying a relaxation strategy based on the linearized constraints will lend to a zero solution vector (U., collapse of SQP at an infeasible point) at some point between point A and B. This s due to the net that locally, SQP cannot tell whether no solution exists for this problem or it is simply unable to find IL Because of this dilemma, it appears unlikely that any safeguard based on local information can be developed to avoid this problem, other than by exhaustively restarting the problem.

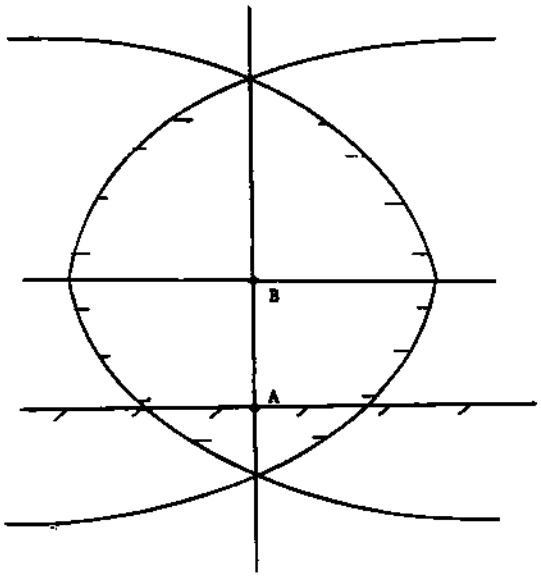


Figure 1: Failure of the QP Step

$$\mathrm{Min}\, x_2$$

$$\mathrm{S.t.}\quad 1. - \bar{X}_1 - x_2^2 \le 0$$

$$1. + X_1 - x_2^2 \le \upsilon$$

$$0.5 + x_2 fc\ \ 0$$

3. SQP Implementation for Process Optimization Problems

Having outlined some of the properties and analogies to Newton's method it is instructive to see how SQP is applied to flowsheeting problems. This section deals with equation oriented and modular qmnlatnrs separately as each has unique characteristics that must be considered for opinurarion. We first begin with modular simulators as these are morecapplicable toecurrent industrialalapplications.

Modular Simulaton

Flowsheeting programs that incorporate process models in black box or modular form are widely used in the process industries [54]. Because solution strategies are specific to process modules, they can be individually tailored for robust and efficient performance. On the other hand, since recycle convergence only uses input-output information from these imMhi'rf. inefficient and often unreliable algorithms based on function values are often applied. On the other hand, superior convergence algorithms, such as Newton's method, require gradient information which can only be obtained by repeated modular perturbation*

Despite this drawback. Newton and quasi-Newion methods are often preferable for simulation, especially if nested recycle loops and design constraints need to considered. In fact, application of these gradient based strategies over function-based algorithms (like successive substitution and Wcgstcin) is the key difference between newer simultaneous modular strategies [14,60] and the traditional sequential modular approach [54,63]. As mentioned in [7], process optimization can be thought of as a simulation proMffm with difficult design constraints, Le. the Kuhn-Tucker conditions. Application of the SQP algorithm to this problem can thus be thought of as application of a quasi-Ncwton method to satisfy these conditions. For flowsheet uprimITHim, however, a number of other issues must be considered that are less important for process simulation. Among these are gridlent *rifajiotirm* problem scaling and the use of partial recycle convergence. These will be motivated first by a simple presentation of theopumizauonfofmiilauon.

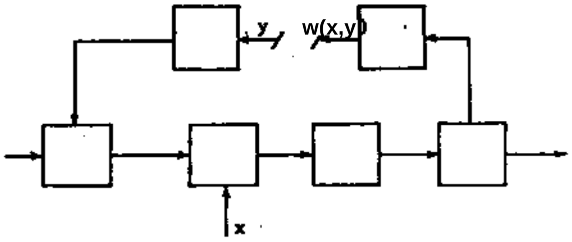


Figure.2: Flowsheet Optimization

$$\mathrm{Min}\ F(x,y)$$

$$\mathrm{s.t.}\ h_t(x_t y) - y - w(x,y)\ »\ 0$$

$$h_d(x,y) - f(x,y) - c = 0$$

$$g(x»y) \le 0$$

Consider the flowsheet shown above in Figure 2. To simulate this process we need to break (or tear) all recycle and calculation loops and define a set of tear variables, y, that converges these loops. Because most of the equations are solved within modules.this problem deals with a much smaller subset of process equations and variables.and can be partitioned as $\backslash(y) = y - w(y)$ » 0 for the recycle loops and h_d » f(y) - c • 0. for additional design specifications. Here c is a constant design specification vector. For optimization, as seen in Figure 2, one also adds decision variables, x, an objective function, F(x,y), and any additional limits on the process or product, g(x,y) · ≤ 0. This leads to the optimization problem:

$$\bullet\mathrm{in}\ \ F(z)$$

$$\mathrm{s.t}\cdot\ g(z)\ *\ 0 \quad \text{where}\quad z^T - [x^T\ \ y^T]$$

$$h(z) - 0 \qquad\qquad h^T = \{h_t^T\ \ h_d^T\} \qquad (10)$$

that was discussed in the previous section. Note, however, that the SQP procedure requires gradients far F, g and h with respect to z at each iteration. Moreover, this iwffflwintftw, either directly or indirectly, determines whether die optimt7aiion problem has converged or not Consequently, efficient determination of accurate gradient information is an important aspect to process optimiration.

Initially, direct loop perturbations (ej. along a path bom x or y to w<x.y) in Figure 2) were used for dm purpose [7.43]. This approach is very easy to implement without much knowledge about the simulator, but it may be inefficient because it requires additional modular perturbations. A study of the perturbation error uto^mH with this approach is PKKitted in Biegkr [3]; an innovative strategy for minimizing this error has been developed and successfully dcnionstrated by KagsJuoto [43]. Later, more accurate gradient strategies were developed in [5, 13]. Here perturbations were done module by module and die resulting partial derivatives were either assembled in a tower triangular form [13] or dinxtiy diainruled using die calculation sequence [5]. In this way analytic Jacobians could be whtrifutrd where available and greater control could be exercised in determining gradient accuracy. However, this strategy is much more difficult to implement on a general purpose simulator. In fact, Chen and Stadtherr [13] cons true ltd an ffPr'f!. demonftnukw sJnwfrFor for this

[illegible]

Together with gradient calculation, scaling of the optimization problem, especially die z vector, can greatly affect algorithmic performance. The purposes of scaling are, first, to transform the problem so that die Hessian is "wefl-approximaied* by a diagonal (usually the identity) matrix, and also to avoid ill-conditioned quadratic programming problems. While this task is often problem specific, several heuristic strategies are available which seem to work well [6, 68,71]. Moreover, ill-conditioning can be monitored "on-line" in order to signal poorly scaled problems.

Finally, process optimizations sometimes benefit from partial recycle convergence applied between QP iterations. This was observed indirectly by Biegler and Hughes [8,9], who noted that feasible variant methods (i.e. where the SQP algorithm was applied but the flowsheet was converged at each iteration) could be faster and more reliable on some process problems. This is especially true if exact penalty line search strategies are used with direct loop perturbations. Chen and Stadtherr [15], on the other hand, observed better performance if additional successive substitution passes were made at the initial point. More recently. Kisala [44], applied a fixed number of Wcgstein posses after die solution of each QP and observed better performance on some problems. Given these findings, one wonders:

- what criterion can be used to determine whether partial convergence will be beneficial and,
- bow can gradient informstion available for optimization be used for efficient partial recycle convergence?

At this stage, little is known about the effects of partial convergence. Therefore, to address die first question we outline a heuristic procedure that makes use of a Kuhn-Tucker convergence lest frequently used in SQP:

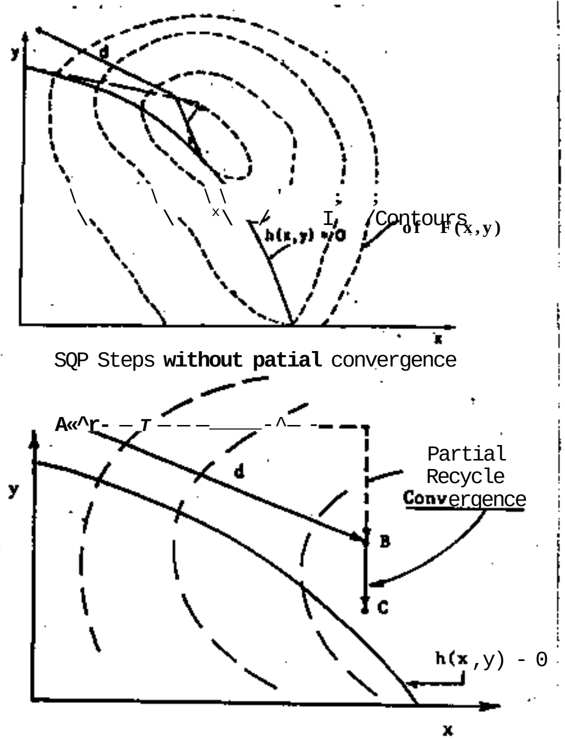
$$|\nabla F(z^i)^T d| + \Sigma\, u_j g_j(z^i)_+ +$$

E|vjhj(s)| * c (ID

vfa«r« c is a tolerance

Note here that die contribution of the last term can be set to zero just by converging the equality constraints. Consequently, if this term ftammatrs die Kuhn-Tucker error at a given iteration, partial recycle convergence (reducing this term to a point where it no longer rtominairi) will reduceejhe overall error. After partial convergence, the Kuhn-Tucker error will then be better "balanced" and the QP should not need any additional help for simultaneous convergence and optimization. Currently this strategy has been implemented using tuning parameters set by the user. Overall, performance, however, does not appear to be very sensitive to their values [46].

In order to perform partial recycle convergence efficiently, it is instructive to consider the two contour pkxs in x and y shown in Figure 3 below. The first shows a sequence of SQP steps along an infeasible path toward the optimum. In the second, partial convergence is performed from point B to point C using gradient information from point A. Note here that the QP b solved (and gradients are only evaluated) at points A and C Consequently, a conventional Broyden method starting with the Jacobian at point B can not be applied here. Instead we derive a nonsquare Broybcn update from point A to point B and then use this Jacobian information for tbe remaining steps [46]. Interestingly, this approach satisfies the same convergence properties that the conventional Broyden method does [21] and performs well in practice (with usually only one or two iterations required for partial convergence).



SQP Steps without patial convergence

Figure 3: Partial Convergence Strategy

Equation Oriented Simulators

For modular simulators the major computational cost is due to calculations performed within modules. Consequently, the SQP algorithm, which numerous computational studies identify [3,6.38,61, 68] as requiring the fewest number of function evaluations for optimization, is the most efficient algorithm for modular optimization. As the size of the problem becomes large (greater than 100 variables, say), die computational effort and overhead required to solve QP's with current, dense quadratic programming implementations becomes significant and. indeed, may be the major computational cost For equation oriented simulators where much of the process is solved and optimized simultaneously, it is therefore necessary to reconsider how large-scale optimization should be performed

Unlike SQP, MINOS [55, 56] was developed as an efficient nonlinear programming problem for targe-scale implementation. Here the nonlinear program is solved as a sequence (injor iterations) of linearly constrained nonlinear programs. The solution of each linearly constrained problem can be found by applying unconstrained algorithms (minor iterations), such as variable metric methods, in the null space of the active linear constraints. Under mild conditions k can be shown [see 561 that major iterations converge to the solution of the nonlinear program at a quadratic rate. Minor iterations generally converge to the solution of the linearly constrained problem at a superlinear rate. Moreover, although this strategy generally requires more function and gradient evaluations than SQP, it lends itself lo efficient sparse implementations, especially with respect to matrix tfldOnZBDOOS TOT OUtl SPBCC QDCS9D0QS* ^^QDBGQDGDujrt 1D6 A&u^v^J'w algorithm often performs well on targe scale optimization problems.

For large scale *process* problems, however, effort required for fiimtw"_ evaluations can still bef»g"_»flnwt and SQP strategies tailored 10 these problems still can be very effective. In the conventional SQP approach die full Hessian of the LagnofB function is approximated through a dense quasi-Newton updating formula. For large scale problems, it therefore seems reasonable to approximate the Hessian matrix in some sparse form and adapt the QP algorithms to take advantage of sparsity in the Hessian and the constraint gradients. Not only wiB this lead to more efficient QP solutions, but die incorporation of a sparse structure will lead to a better approximation of the Hessian and fewer SQP iterations. Lucia and coworkers [45, 74] have successfully demonstrated this approach for ofXtmhration of vapor-liquid process models. Here, derivative information b supplied where calculable and remaining information is approximated through sparse quasi-Newton formulae that are thermodynamically consistent

A second approach comes from realizing that process optimization problems usually have many variables and equations, and relatively few degrees of freedom. Consequently, if SQP can be applied in the null space of the equality roMtrainn. a much smaller QP needs to be solved. To motivate this riivimion, consider an equality constrained optimization problem with an onhonormal null space to V_rh, Z and range space, Y, that satisfy the following relations:

$$
\begin{array}{lll}
Vh^T Z - 0 & Z^T Z = I & YS - \nabla_z h \\
Y^T Z - 0 & Y^T Y = I &
\end{array}
\qquad (12)
$$

Here $\nabla_z h$ e $R^{n \times m}$, $YeR^{1* \times m}$, $ZcR^{n \times (n-m)}$ and SeR*** la son* nonsingulār matrix.

Using this onhonormal basis, one can therefore transform the quadratic programming problem (for eqn. (4) above) from:

$$
\begin{bmatrix} B & \nabla_z h \\ V_z h^T & 0 \end{bmatrix}
\begin{bmatrix} d \\ v \end{bmatrix} = - \begin{bmatrix} V \\ h \end{bmatrix}
\qquad (13)
$$

to:

$$
\begin{bmatrix} Y^T BY & Y^T BZ & S \\ Z^T BY & Z^T BZ & 0 \\ S^T & 0 & 0 \end{bmatrix}
\begin{bmatrix} d_Y \\ d_Z \\ v^{i+1} \end{bmatrix} = - \begin{bmatrix} Y^T V_Z F \\ Z^T \nabla_z F \\ h \end{bmatrix}
\qquad (14)
$$

$$
d - Zd_Z + Yd_y
$$

From the second and third block matrix rows of (14), one sees that the search direction, d, is independent of the Lagrange multipliers and the search direction in the range space, d_Y does not depend on the Hessian, B. Moreover, Gill, Murray and Wright [34] show that by eliminating $Y^T BY$, $Y^T BZ$ and $Z^T BY$ (or setting them to zero) and using only the projected Hessian $Z^T BZ$, one can solve a much smaller problem:

$$
\begin{array}{l}
Z^T BZ d_Z = - Z^T VF \\
S^T d_Y - - h \\
Sv^{1+1} = - Y^T \nabla_z F
\end{array}
\qquad (15)
$$

and still obtain a good estimate of the multipliers. More recently, Nocedal and Overton [58] show that an SQP algorithm based on (15) also converges at a two-step supcriinear rate. They also present a brief numerical study using a family of reduced strategies that compare well to fuU Hessian SQP strategies.

With this background, we can now re-examine recently developed decomposition strategies for process problems in terms of range and null spaces. For sparse implementations an easy to compute null space can be defined by:

$$
. \; . \left[\begin{matrix} -(\nabla_y h^T)^{-1} \nabla_x \\ I \end{matrix} \right. 1 \qquad (16)
$$

Using this definition, Berna ct al [3] extended the strategy of cqn (15) to deal with inequality constraints. Here, solution of the QP is only required in the space of x, but *toe full* Hessian is approximated with update vectors stored in a Cactorizcd form. However, this strategy is difficult to implement and can require excessive storage for problems that require a large number of iterations [40, 41]. The more recent approach by Locke et al [50] is simpler and deals with solving the reduced QP, and updating $Z^T BZ$ directly. Since the full Hessian is generally not positive definite for large-scale problems, both implementations should have the same convergence rate. However, very limited testing of the latter strategy has been reported [49,50,72]. Moreover, no reliable implementation of this reduced Hessian strategy is currently available that addresses the other SQP properties mentioned above.

The previous two sections of this paper focus on properties and characteristics of algorithms that have been developed for simultaneous process simulation and optimization. In the following section we deal with formulations of more difficult problems that the simultaneous strategy now allows us to consider.

4. Applications of Simultaneous Simulation and Optimization

The next three subsections deal with problems thai expand the scope of process optimization. Fust we discuss the simultaneous application of heat integration with process optimization and summarize a case study that demonstrates its effectiveness. Next we show that other time-consuming process models can also be treated with this simultaneous concept, and outline an application to differential equation models. Finally, we present a strategy for assessing the sensitivity of an optimal flowsheet to parametric variations.

Simultaneous Heat Integration and Optimization

In parallel with process optimization, efficient strategies have been developed for synthesizing energy efficient heat exchanger networks. While these strategies are easy to apply [39,48], offer valuable insights for process improvement and can be automated [29, 59], they can only deal with fixed process flows and temperatures. In the context of process optimization, however, the effect of heat integration interacts strongly with the rest of the process and therefore must be included as part of the problem. To see this effect, consider the trade-off curves in Figure 4. Here, application of heat integration *after* optimization serves to reduce the total cost but the overall process conditions (e.g. flows, temperatures, overall raw material conversion) remain the same. On the other hand, if one is able to account for heat integration as a function of process conditions, the dashed curves result in Figure 4. Note that, as has recently been shown [23, 47], this not only results in lower total costs but also leads to better utilization of raw materials through greater overall conversion.

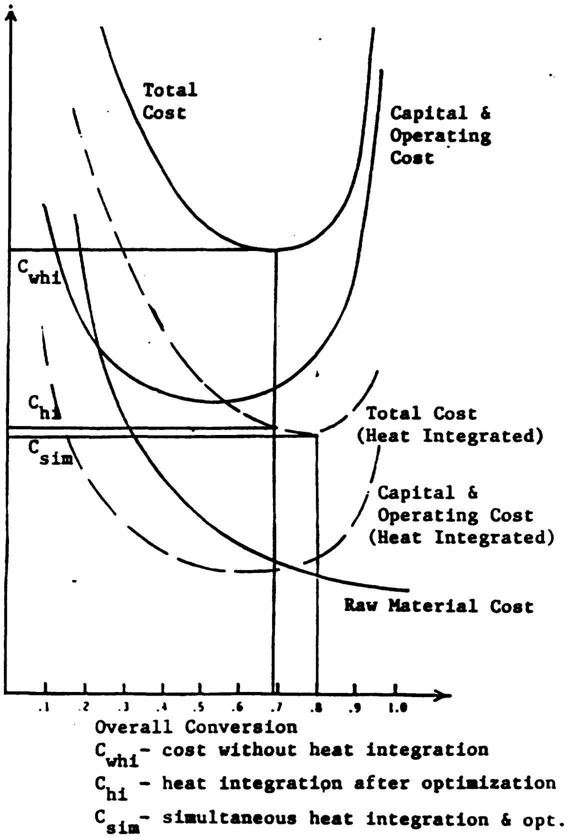


Figure 4: Comparison of Costs with Heat Integration

To account for heat integration within process optimization, Duran and Grossmann [23] include additional constraints and objective function terms that represent conditions for maximum energy recovery. These have recently been incorporated into the process optimization capability [46] implemented on FLOWTRAN [63, 67]. One resulting formulation of the problem is presented in Figure 5, where heat integration is implemented as an additional "cost block". Within this block, one solves for maximum energy recovery, using the equivalent of a linear programming strategy proposed by Duran and Grossmann [23] for multiple utilities. This information is then passed to the optimization problem simply as additional terms in the objective function. This approach now accounts for the interaction of the process with the heat integration network. It should be noted, however, that the straightforward formulation of Figure 5 may lead to nondifferentiable functions. While these have not presented any problems in our current work, this difficulty can be circumvented by using a more complex formulation as presented in [23, 47].

To see the effect of this strategy consider the ammonia process flowsheet presented in Figure 6. As seen by the hot and cold streams (H1, C1, etc.) in Figure 6, this single loop process presents several opportunities for heat integration. The optimization problem is given below and includes nine decision variables such as inlet flash temperatures, compressor pressures and feed flowrates. This process was first optimized on FLOWTRAN without heat integration. Starting from a base profit of \$20.66 x 10^6 /yr, the optimum was found to be \$24.93 x 10^6 /yr after only 5 SQP iterations. (Interestingly, with the partial convergence strategy outlined above, the optimum can be found for this problem after only 3 SQP iterations [46].) Application of heat integration strategies on the "optimized" flowsheet further increases the profit to \$26.86 x 10^6 /yr. Instead of this sequential procedure, we now perform these two tasks simultaneously using the strategy in Figure 5. Here, the optimal profit is \$27.65 x 10^6 /yr or almost \$800,000/yr greater than with the sequential approach. The reason for this is justified by different operating conditions in the optimal flowsheet. Here the overall conversion of hydrogen is 1.5% higher, the main compressor pressure is 41 atm lower and the recycle ratio is 33% higher [47] than with the sequential procedure.

This example illustrates the importance of accounting for interacting subsystems in process optimization. In the next section a similar situation occurs with process reactors in flowsheets. While in this section simultaneous solution was made possible through a new formulation of the optimization problem, the next section discusses the reformulation of the process *model* in order to adapt to simultaneous solution.

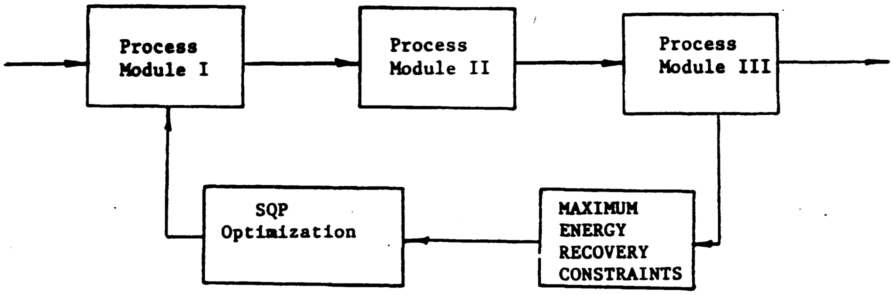


Figure 5: Simultaneous Optimization and Heat Integration Strategy (Implicit Form)

Simultaneous Solution and Optimization of Flowsheets with Differential Equation Models

Process models that require the numerical solution of differential equatkrawithconipltttheniiodyijai^ and/or transport properties can lead ao very tme-consuming calculations. Here the model must be solved for each flowsheet pass and physical properties must be caioilaiedtoeachimegratk»«Kp.lf optimization is then applied and additional gradient mformatioa is required, the computational expense can be protuoajve. in pracuce, many process SMnwanons use simple reactor models based on split frictions or approaches to equilibrium. These do not, however, fully account for ihe imcrsction of the reactor with die rest of the flowsheet. Moreover, since the reactor frequently oeieiuimes me cnaracier oi me IMJWSUUH uvougn icyiaimi steps, recycle structure and possibilities for heat integration, incorporating this interaction into the optimization problem is clearly important.

Consider the following optimiiaiion problem with an ordinary differential equation (ODE) modet

$$
\begin{aligned}
&\bullet\text{in} \quad F(z.C) \\
&\text{s.t.} \quad g(z.C) \; * \; 0 \\
&\qquad h(z,C) \; - \; 0 \qquad\qquad\qquad (17) \\
&\qquad \frac{dC.}{{}^{X}dt} \; - \; f(2,C,t); \; \zeta(0) - \zeta_0
\end{aligned}
$$

Hoe the variables fare functions of the spatial variable L

In order to allow for simultaneous solution of the *ODE* model, recycle equations and the optimization problem, we fist apprfl**maff the differential equation model by a set of algebraic equations. In order to keep the problem small and also to guarantee a stable solution, we apply an implicit, higher order method. Here, perhaps the easiest method to apply is orthogonal collocation [73]. By approximating the polynomial with a set of Lagrange polynomial basis functions, algebraic equations can be written at n points, the shifted roots of an nth degree orthogonal polynomial, t^ These collocation equations are:

$$
\begin{aligned}
&\sum_{j=0}^{n} \zeta_j \frac{dl_j}{dt}(t_i) - f(z,\zeta_i,t_i) - 0 \quad i-1,n \\
&l_j(t) - \prod_{\substack{j=0 \\ \neq i}}^{n} \frac{(t-t_j)}{(t_i-t_j)} , \; t_0 - 0 \qquad (18)
\end{aligned}
$$

Note that these coUocation equations are easily constructed from the differential equations. Values for dlj/dt can be computed once the number of collocation points is known, and can be stored for later use. Because this method requires far fewer points than a conventional ODE solver, fewer equations and property evaluations are required for the reactor module. Moreover, the collocation equations (18) and polynomial coefficients, Cj, can be added directly to the optimization problem and any profile constraints can be incorporated directly as constraints on the coefficients. Consequently, this approach is potentially more powerful and efficient than embedding a conventional ODE solver within the process module.

CuthreU and Biegler [18] recently demonstrated this approach on the simple flowsheet optimization problem shown below in Figure 7.

For this problem both Ihe conversion and temperature profiles could be approximated accurately using only five collocation points. In comparison with direct use of an ODE solver, the number of property evaluations for the reactor module is reduced by a factor of 30. Here the number of iterations for the simultaneous solution was 27 while the formulation with an ODE solver [37] required 22. Consequently, at the expense of dealing with a larger nonlinear piogiain (and its associated QP step), the computational effort for the reactor module and any other similar process modules can be significantly reduced.

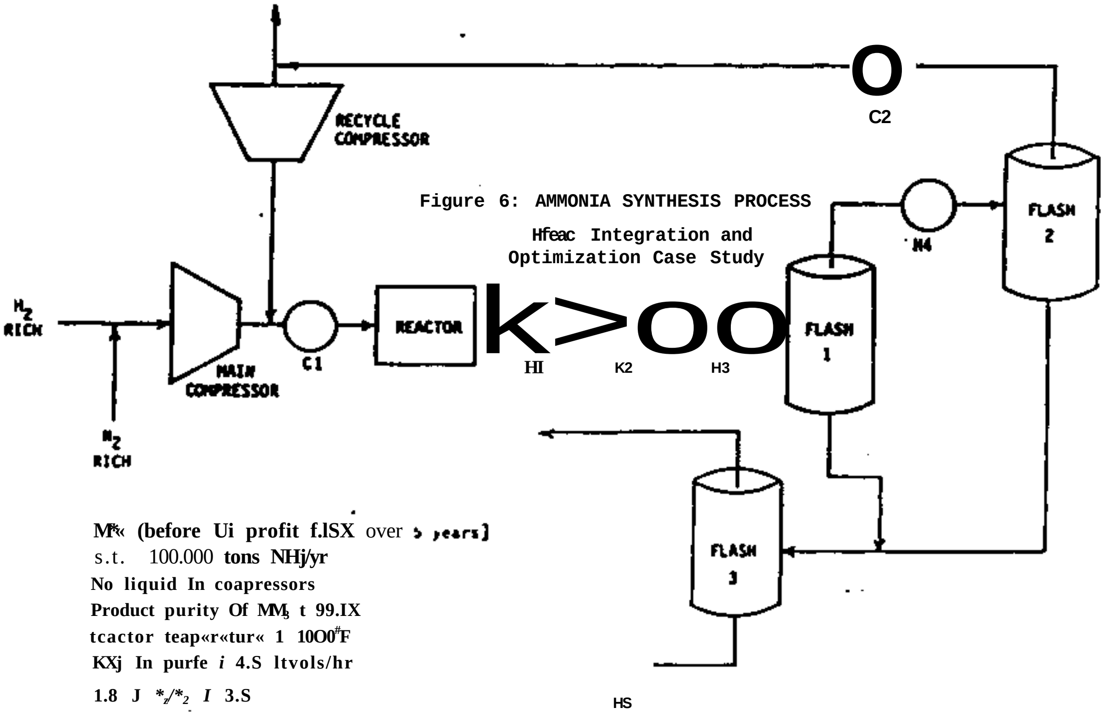


Figure 6: AMMONIA SYNTHESIS PROCESS

Hfeac Integration and Optimization Case Study

M*« (before Ui profit f.lSX over 5 years]
s.t. 100.000 tons NHj/yr
No liquid In coapressors
Product purity Of MM_3 t 99.IX
tcactor teap«r«tur« 1 10O0#F
KXj In purfe *i* 4.S ltvols/hr
1.8 J $*_z/*_2$ *I* 3.S

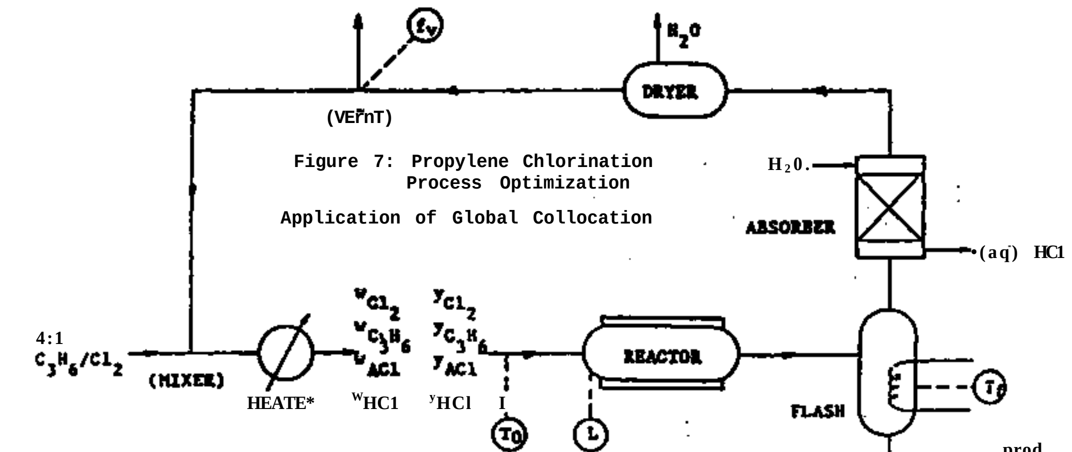


Figure 7: Propylene Chlorination Process Optimization

Application of Global Collocation

RS/sec) • -220.462($/gmole)/$^{fO<l}_{ACl}$* 9.1i3E-2($/m-secX. • $w^{prod}_{Cl_2}$

S.1535E-3($/m-sec-k)$\int_0^L$(T(z)-T$_{sink}$)dz • 9.486E-6(S/J)Q$_{flash}$ w^{prod}_{ACl}

However, for stiff systems of differential equations or problems with steep profiles, a global collocation approach cannot [73] give accurate solutions. Instead, one can successfully formulate this problem using orthogonal collocation on finite elements [2, 33]. Briefly, the oprimi7arion problem can be stated as:

Min F(z,O

g(z»O S 0

h(z,O - 0 i - 1, n; k - 1, NE

$$\Sigma\ \zeta_{jk} \frac{dl_j}{dt} (t_{ik}) - \Delta\alpha_k f(\zeta_{ik}, z) = 0$$

C(t) is continuous and piecewise polynomial of order n+1 (1.)

$A\alpha_k$ 2 0

$$t_{ik} = \sum_{j=1}^{k-1} \Delta\alpha_j + \Delta\alpha_k \tau_i$$

T< - shifted Legendre roots between 0 and 1

$A\alpha_k W_k$ - $A\alpha_{k-1} W_{k-1}$ k-2,NE

where V_k is based on higher derivatives of C(t) in $A\alpha_k$ with respect to tiae.

Compared to the global collocation formulation, the collocation equations in (19) are now written on finite elements, $a_{i(}$ and continuity is imposed at the boundaries of these elements. Also we have included an additional set of conditions that govern the placement of the elements. As formulated in [17, 20], one can show that the element placement equations are necessary and sufficient for minimizing the approximation error. Moreover, this approach has strong stability properties as well as a high approximation order. Ascher and Bader [1] show that this method is equivalent to a fully implicit Runge-Kutta method applied at OfHif"]** roots. Such methods are *algebraically* stable, i.c stable for a wider class of problems than A-stable methods. In fact, because of this property orthogonal collocation on finite elements is desirable not only for differential equation models but also for two point boundary value problems.

To demonstrate the formulation given by (19), consider the reactor optimization problem given in Figure 8 [19. 28]. Here the reaction is exothermic and goes to completion in the reactor. Because it raises steam in the cooling jacket, it may be advantageous to design a reactor that is longer than needed for complete conversion, in order to provide additional heat transfer. This design leads to temperature profiles, with hot spots that arc difficult ID approximate. Applying the finite element formulation leads to the dimcnsionlcss temperature profiles presented in Figures 9 and 10. Note that elements arc closer together in steep regions and farther apart for more gradual ones. Moreover, in Figure 10 a bound of T(t) £ 1.4S is placed on the temperature profile and it can be seen that the hot spot and reactor length shift to accommodate this constraint

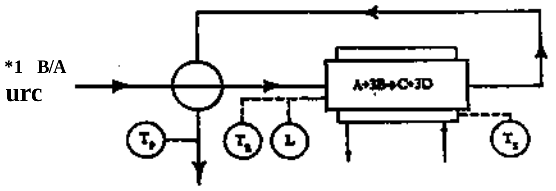


Figure 8: Reactor Optimization with Hot Spot

Min (Reactor Capital Cost - Steam Credit)

The ability to handle difficult differential equation based problems in this way allows us also to consider optimal control problems as well. A number of examples with control profiles have been presented in [17,19] but discussion of these more difficult problems is beyond the scope of this paper. However, it is important to note that the ability to handle complex ODE process models is directly influenced by the ability to do large scale optimization. More will be said about this in the final section. Before discussing this, however, we first turn to ihc topic of postoptimality analysis.

Sensitivity Analysis of Optimal Flowsheets

In the topics covered so far emphasis has been placed on formulating the optimization problem and solving it efficiently. In this section we try to characterize the optimal solution and its sensitivity to parametric and modd changes. For example, a given design or optimization study win require the selection of external parameters, p. at some nominal point. These can include external utility supply temperatures and pressures, composition and Oowntes of the feedstock or even hcai

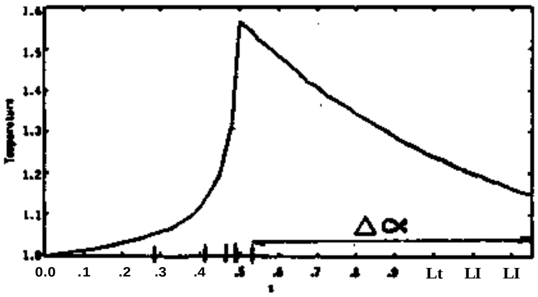


Figure 9: Optimal Temperature Profile
____Collocation Solution
– Solution from ODE Solver

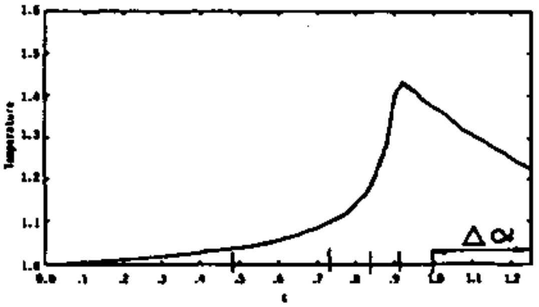


Figure 10: Optimal Temperature Profile with T(t)≤ 1.45

____Collocation Solution
____ODE Solver Solution [37]

transfer coefficients and kinetic pvr&u*. Here we are interested in estimating the change in the optimal sokrica as a fwc*» c/changes in these parameters. Lc.dz*/3 p. To do this we consider ibe parametric cycunujtiunprobiem:

$$\begin{array}{lll} \text{Mia} & F(z, p^{o}) & \\ \text{s.t.} & g(«_f\ p^{o})\ *0 & \\ & h(z.\ p^{\circ})\ -\ 0 & \qquad (20) \end{array}$$

with the associated Kuta-Tocker conditions:

$$\begin{array}{c} VF(z^\circ)\ +yg(x^\circ)u^\circ\ \ +_V h(z^\circ)v^\circ\ -\ 0 \\ h(z^\circ,\ p^\circ)\ -\ 0 \\ g(z\ \cdot\ p\ )\ *\ 0 \qquad (21) \\ u^{To}g(x^{o}._P{}^\circ)\ -\ 0 \\ u^\circ *\ 0 \end{array}$$

Fiacco *{2627]* shows that if strict ccopiemetarity and linear independence of the active constraint *set* holds at the optimum, then the active set will not change for small parametric perturbations. Consequently, we can write the change of the solution, A z. to an C"i pites mal) change in the parameters, Δ p, as:

$$\begin{pmatrix} *V^3 \\ \Delta(\nabla_y L) \\ \Delta g_A \\ Ah \end{pmatrix} *- \begin{pmatrix} {}^7{}_{xx}L* & V_{xy}L* & \nabla_x g_A^* & V_x h- \\ V''' & VyyL'V & \nabla_y g_A^* & 7_y tf \\ *x«A^{T} & \nabla_y g_A^{T} & 0 & 0 \\ V* & \nabla_y h^{T} & 0 & 0 \end{pmatrix} \begin{pmatrix} Ax \\ Ay \\ Au \\ Av \end{pmatrix} \qquad (22)$$

Note that the gradient information in the above matrix equation is available from the optimal solution. Because a qutsi-Ncwion Hessian approximation is used by SQP, however, second derivative informaiioo most stU be evataated Um»gh add^tioiul penurbtlions. Here we lake advantage of dttowposkic* coaccptt ineiitioncd*« section 3 in order to evaluate the actual Hessian projected k a reduced space. As suggested by Edahl [25], btock Gaussian decomposition of the above system leads to:

$$\begin{bmatrix} a \\ b \\ f \\ e \end{bmatrix} = - \begin{bmatrix} I & 0 & E & 0 \\ 0 & I & L & M \\ 0 & 0 & H & Q \\ 0 & 0 & Q^T & 0. \end{bmatrix} \begin{bmatrix} \Delta y^* \\ Av^\circ \\ \Delta x^* \\ Att^\circ \end{bmatrix} \qquad (23)$$

where

I « Identity Matrix

$$E\ «\ -\ A = (\nabla_y h^{T*})^{-1}\ \nabla_x h^{T*}$$

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

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

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

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

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

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

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

and the reduced space system of linear equations is given by:

$$\begin{bmatrix} f \\ e \end{bmatrix} = - \begin{bmatrix} H & Q \\ Q^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x^* \\ \Delta u^* \end{bmatrix} \qquad (24)$$

Now instead of constrocting the enüre Lagrange Hessian by perturbation. H can be constructed by performing simultaneous perturbations in x and y and remámingím the null space ofV^Lc,

$$(Ax_{jt}\ -(Vy'aVS^r'JAXj)\ -\ ^.Ay)$$

The number of flowsheet evaluations required to construct the scQŠuivi(y Buofmjtfi̇oQ is Urns fOtŠiooQ ťinoiia*

$$NFE = (n_x + n_y)(n_x + n_y + 1)/2 \qquad (25)$$

where

V number of dependent (tear) variables
n_x - number of independent (decisíon) variables

to:

$$NFE = (n_z(n_z+1)/2)+3n_z+2 \qquad (26)$$

In addition, sensitivities for y can be evaluated from EA x and a: the only information not available from the reduced equations is the sensitivity of the equality constraint multipliers, v. Normally this information is not as important as for the other variables.

As an illustration of this analysis, consider the simple flash loop optimization in Figure 11. Hoe the solution lies at the tower bounds of the split firactkm and flash pressure. As expected from physical insight, a change in the feed flowraie will not change the optimal values of these decision varibles; this is easily verified using the reduced Hessian sensitivity analysis in eqn (24). As shown in [31, 32], however, substitution of different objective functions wiO make the decision variables more sensitive to parametric variations.

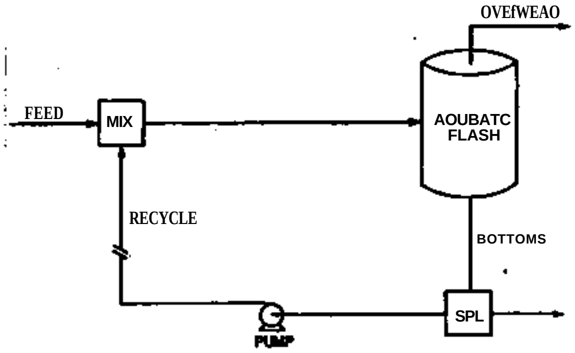


Figure 11: Sensitivity of Optimum

Max (Light Comp. in Ovhd.)

$0.2 \leq$ Split Fraction ≤ 0.8

$69 \leq$ Flash Pressure $\leq$3A5 kPa

Sensitivity of the optimal solution to the choice of process *models* is more difficult to analyze and a thorough derivation is beyond the scope of this paper. To a first approximation, a simple analysis can be made by considering the model "sensitivity" as a Newton step, from the optimal solution obtained with Model I, say, applied to the optimality conditions in the space of Model II. A simple illustration of this concept is shown by the contour plots in Figure 12. Although this problem differs conceptually from parametric sensitivity, the derivation of the Newton step leads to same equations as in (22) and the reduced Hessian strategy in (23) can be applied directly to the calculation of the Newton step [32].

5. Problems for Future Research

While simultaneous simulation and optimization strategies have certainly allowed the solution of difficult process optimization problems, much work still needs to be done in order to solve larger and more complex problems. If anything, the above review has outlined recent work that demonstrates the *feasibility* of these strategies. Although successful implementations of SQP-based strategics have been made on commercial process simulators there still remain some open questions regarding efficient and reliable formulations for different classes of process optimbaiion problems. In this section we briefly discuss a short list of unsolved problems with an eye toward future research.

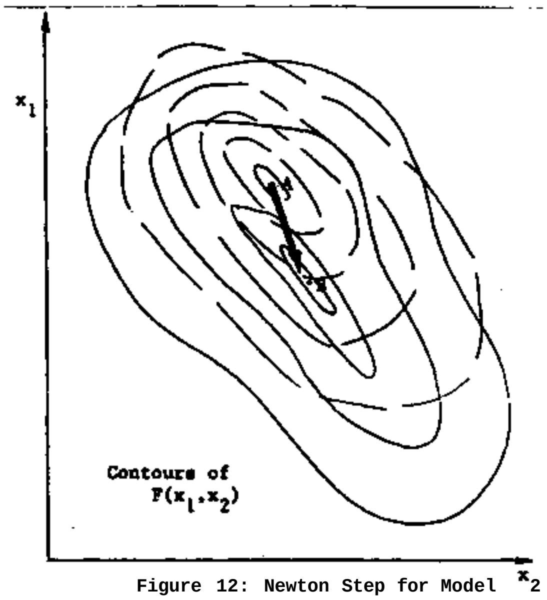


Figure 12: Newton Step for Model "Sensitivity"
———— Model I Space
A- Model I Optimum
———— Model II Space
B- Model II Optimum

SQP algorithms for larger optimization problems.

As mentioned in the first section of this paper, decomposition strategies have been developed but as yet no general purpose, reliable implementations of these are available. Two reasons for this are that these strategies rely on a linearly independent representation of the null space and that the reduced problem is often very highly constrained. For large, sparse problems, representation of the null space depends strongly on the choice of dependent and decision variables. Here it is most important that the Jacobian matrix, $V_y h^T$, remains nonsingular for all points encountered by SQP. And conditioning of this matrix will affect the scaling of the problem and performance of SQP. Also, QP problems after decomposition tend to have few variables but many inequality constraints that derive from bounds on dependent variables. Because of this highly constrained nature, small inaccuracies in constraint gradients may even cause the feasible region in the reduced space to vanish. Finally,with inequality constraints, the QP algorithm needs to solve a more difficult combinatorial problem.

An alternative to this formulation is to consider sparse QP implementations that lead to well conditioned null spaces. One very promising QP strategy was recently developed by Ng and Thompson [57]. For the SQP algorithm, however, one must still determine whether the Hessian matrix should be approximated in the full or reduced space, and how to treat changes in the set of dependent and independent variables from iteration to iteration.

Solution oflarge optimization problemsfor modular simulators

Even with a reliable large-scale SQP strategy, large process optimization problems can still be time-consuming if gradients arc required through modular perturbations. Consider, for example, a flowsheet with two decision variables and fifty tear variables. If the flowsheet is easy to solve with a simple recycle convergence algorithm, it would clearly be better, from a simulation and nonlinear programming standpoint, to treat the flowsheet as a black box and require three converged flowsheets at each SQP iteration (one for function evaluation and two for gradient perturbations), rather than 53 flowsheet passes at each SQP iteration. On the other hand, if the flowsheet has nested or difficult recycles, or design constraints are applied, the black box approach will also be inefficient and less reliable than the simultaneous strategy.

To reduce the work for the simultaneous strategy a number of ideas have been proposed. First, simple models [16] can be substituted for gradient and function evaluation. These could be adjusted during optimization to match the properties of the more complex process models. Jirapongphan et al [42] proposed a very efficient strategy based on this idea, although it was later realized that this approach can lead to suboptimal solutions. More recently, safeguards have been proposed that still allow the use of simple models [31,72] and lead to the optimal solution of the original problem. However, these remedies are at the expense of greater computational effort; it still remains to be seen whether these strategies can be made efficient for general flowsheeting problems.

Other ideas that reduce the size of the optimization problem and the work of gradient calculation include the lumping of stream components, and the selective incorporation of recycle loops into the optimization problem [69]. At present, however, these have only been applied in an ad hoc manner, based on specific information of the process problem. Although these concepts have the potential to improve performance, it seems that their systematic application to general purpose problems will be difficult.

Simultaneous solution and optimization for differential algebraic models

Use of simultaneous strategies for differential equation modules can greatly reduce the effort in solving the process model as well as deal directly with state variable profile bounds in the optimization problem. However, while the potential for simultaneous treatment of differential equation models has been demonstrated, much work still needs to be done. First, reliable and efficient large scale SQP capabilities need to be developed since increasing accuracy with finite elements quickly leads to a large number of variables and equations. On the other hand, because more of the structure is known for collocation equations, it should be possible to develop more reliable decomposition strategies.

More general questions involve the number of finite elements and collocation points required to achieve a given level of accuracy. An encouraging result for answering these questions comes from establishing the relationship between this approach and implicit Runge-Kutta methods [1]. As discussed above, this can lead to better insights regarding stability and error control.

Extension of simultaneous strategies to more difficult optimization problems

The simultaneous heat integration and optimization strategy outlined above is an illustration of combining two separate tasks in order to solve a more general problem. Similarly, mixed integer programming [24, 59] strategies for process synthesis can communicate with process models and often require the results of nonlinear optimization problems. Consequently, it would be useful to combine the simultaneous strategy on process modules with synthesis strategies in an efficient way.

Recent developments in flexibility analysis [70] also require the efficient solution of process optimization problems. While these strategies have been demonstrated on small equation-oriented process problems, they can clearly be formulated to evaluate the flexibility of complex simulation models. Moreover, simultaneous optimization strategies should allow the automation of flexibility analysis as part of a design and optimization study.

6. Conclusions

Simultaneous simulation and optimization strategies for process problems were first developed by Berna et al [3]. Since then a number of different studies have shown the effectiveness and flexibility of this approach to handle time-consuming and difficult optimization problems. In this paper, we summarized some theoretical concepts for this approach by relating convergence properties of the SQP algorithm with Newton's method. This discussion also allowed us to compare recent improvements to SQP on a theoretical basis. Following this, application of SQP to simultaneous strategies was summarized for equation oriented and modular simulators. Here decomposition strategies were highlighted for equations while issues of gradient calculation and partial convergence were stressed for modular simulators.

Once an optimization capability is available, problems that were previously considered out of reach can now be handled. Here we discuss three such problems:

- simultaneous optimization and heat integration [47]
- process optimization with differential equation models [18,19]
- sensitivity analysis of optimal flowsheets [32].

For each topic a small example is presented to illustrate the approach taken and its results. The above discussion shows that these problems can be solved quickly and the results are either novel or noticeably better than those obtained with conventional methods. However, the above formulations for these problems have only recently been developed and much work remains in refining these approaches.

After reviewing some encouraging results obtained over the last five years we conclude with a summary of unsolved problems for simultaneous strategies. These deal mainly with developing strategies for larger problems and formulating approaches to handle more difficult ones. Nevertheless, it should be noted, that development of simultaneous solution and optimization strategies has led to reconsideration of process optimization as a valuable design tool. The work accomplished so far has demonstrated the feasibility and efficiency of these approaches for process problems. It now remains for us to understand and improve these approaches in order to solve more difficult optimization problems.

REFERENCES

1. Ascher, U. and G. Baöer,"Stabfliry of Collocation at Gaussian Points", SIAM J. NOIL Ami., 23. 2, p. 412 (1986)

2. Ascher, U., J. J. Christiansen and RX>. Russell, "A Collocation Solver for Mixed Oder Systems of Boundary Value Problems,"

3. Bern*. TJ., M.H. Locke and A.W. Westerberg. "A New Approach to Optimization of Chemical Processes", AIChE J., 26,1. p. 37. (1960)

4. Biegler, L.T.. -Simultaneous Modular Simulation and Optimization", in Fo—riitio- of Computer Aided Process Design - II. p. 369 Wesierberg and Chien (eds.), CACHE Committee, (1984)

5. Biegler. L.T., "Improved Infeaable Path Optimization for Sequential Modular SbnvJaton. Pirt L The Interface", Comp. and Cbem. Eat>, 9.3. p. 245. (1985)

6. Biegler, L.T. and J.E. Cuthrelc z'Ig i vU c P x h Optimization for Sequential Modular Simulators, Part II: Hie Optimization Algorithm", Comp. and Chem. Eng., 9. 3.Q.257.(1985)

7.Biegler. L.T. and HR. Hughes, "Infeasible Path Optimization of Sequential Modular Simulators" AIChE J.. 28.6. p. 994 (1982)

8.Biegier. L.T. and RJL Hughes. -Optimization of Propylene Chlorinatioa Process: A Case Study Comparison of Four Algorithm*' Comp. and Chem. Eng.. 7.5. p. 645 (1983)

9. Biegler. L.T. and RJL Hughes, "Feasible Path Optimization for Sequential Modular Simulators" Comp. and Chem. Eng.. 9.4. p. 379 (1985)

lO.Boggs, P.T. and J.W. Toile, "A Family of Descent Functions for Constrained Optimization". SIAM J. Num. AnaL, 21.6. p. 1146(1984)

11. Boggs. P. T., Tolle, J. W. and P. Wang. "On the Local Convergence of Quasi-Newton Methods for Constrained Optimization". SIAM J Control and Optimization. 20.161 (1982).

12. Chamberlain, R.M., MJJ>. Powell, C. Lemarechal, and H.C. Pedersen, "The Watchdog Technique for Forcing Convergence in Algorithms for Constrained Optimization". Math. Prog. Study. 16. p.l. (1982)

13. Chen. H-S and M.A. Stadtherr. "A Simultaneous Modular Approach to Process Flowshecting and Optimization. Part I: Theory and Implementation ". AIChE J.. 31. 11. p. 1843.(1985)

14. Chen. H-S and MA Stadtberr, "A Simultaneous Modular Approach to Process Flowsheeting and Optimization. Pan II: Performance on Simulation Problems". AIChE J.. 31. 11. p. 1857.(1985)

15. Chen. H-S and MA. Stadtherr. "A Simultaneous Modular Approach to Process Flowsheeting and Optimization, Part III: Performance on Optimization Problems". AIChE J.. 31.11. p. 1868. (1985)

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

n.Cuthrell. J. E., PhJ). Dissertation. Carnegie Mellon University. (1986).

18. Cuihrell. J.E. and L.T. Biegler. "Simultaneous Solution and Optimization of Process Flowsheets with DifTcrntial Equation Models". Chem. Eng. Res. Des.. 64. p. 341 (1986)

19. CuthreU. J.E. and L.T. Biegler. "On the Optimization of Differential-Algebraic Process Systems", AIChE J.. to appear. (1986)

20. deBoor, C, A Practical Guide to Splines. Springer-Veriag, (1978).

21. Dennis, J£. and JJ. More, "Quasi-Newton Methods. Motivation and Theory", SIAM Review, 19. 1. p. 46 (1977)

22. daPOlo, G. and L. Grippo, "A new class of augmented Lagrangiatts in nonlinear programming". SIAM J. Com. and Opu 17. p. 618 (1979)

23. Duran, M. A. and I. E. Grossmann, "Simultaneous Optimization and Heat Integration of Chemical Processes" AIChE J, 32,123 (1986).

24. Duran, M. A. and I. E. Grossmann, "A Mixed Integer Nonlinear Programming Algorithm for Process Systems Synthesis-, AIChE J. 32.592 (1986).

25.Edahl, R., "Sensitiviry Analysis in Nonlinear Programming - Applications to Comparative Statics and Algorithmic Construction", PhJ> Thesis, Carnegie-Mellon University (1982).

26.Fiacco, A. V., "Sensitivity Analysis for Nonlinear Programming Using Penalty Methods", Mathematical Programming,10.287(19761.

27. Fiacco, A. V., Introdaction to Sensitivity and Stability Analysis hi Nonlinear Programming. Academic Press, New York (1983).

28.Finlayson, B. A^ The Method of Weighted Residuals awl Variation.! Principles, Academic Press, (1972).

29.Floudas. CJL, AJL Ciric, and LE. Grossmann, "Automatic Synthesis of Optimum Heat Exchanger Network Configurations", AIChE J, 32,2, p. 276,(1986)

3O.Ganesh, N. PhJ). Dissertation, Carnegie Mellon University. (1986).

31.Gancsh, N. and L.T. Biegler, "A Reduced Hessian Strategy for Sensitivity Analysis of Optimal Flowsheets", AIChE J. to appear, (1986)

32. Gaicia-Palomares, U.M. and O.L. Mangasarian, "Superlinearly Convergence Quasi-Ncwton Algorithmsfor Nonlinearly Constrained Optimization Problems". Math. Prog.. 11, p. 1.(1976)

33. Gaidini. U A. Scrvida. M. MorbideUi and S. Carra, "Use of Orthogonal Collocation on Finite Elements with Moving Boundaries For Fixed Bed Catalytic Reactor Simulation." *Comp. and Chan. Eng* 9(1). pp. 1-17. (1985).

34. Gill. P. E., Murray. W. and M. H. Wright, Practical Optimization, Academic Press, London (1981).

35. Han, S-P, "Superlinearly Convergence Variable Metric Algorithms for General Nonlinear Programming Problems", Math. Prog^ 11. p. 263 (1976)

36. Han, S-P, "A Globally Convergent Method for Nonlinear Programming-, J. Opt Theo. and Appl.. 22, 3. p. 297 (1977)

37. Hindmarsh, A. C. "LSODE and LSODI, Two new initial value ordinary differential equation solvers," ACM-SIGNUM Newsletter. 15(4). pp. 10-11. (1980).

38. Hock. W. and K. Schittkowski. "Test Examples for Nonlinear Programming Codes", Lecture Notes in Economics #187, Springer (1981)

39. Hohmann, EC, PhD Thesis, University of Southern California, (1971)

40. Hutchison, H.R, DJ. Jackson and W. Morton. The Development of an Equation-Oriented Flowsheet Simulation and Optimization Package-1. The Quasilin Program", Comp. and Chem Eng., 10,1, p. 19 (1986)

41. Hutchison, H.P.. DJ. Jackson and W. Morton. The Development of an Equation-Oriented Flowsheet Simulation and Optimization Package-II. Examples and Results", Comp. and Chem Eng., 10.1. p. 31 (1986)

42. Jirapongphan, S.. Boston, J. R, Brio, H. I. and L. B. Evans, "A Nonlinear Simultaneous Modular Algorithm for Process Flowsheet Optimization", presented at 80th AIChE Meeting. Chicago (1980).

43. Kaijaluoto, S., "Process Optimization by Flowsheet Simulation" Technical Research Center of Finland. Publication #20 (1984)

44.Kisala, TP.. ScD Thais, Mimarhiwcm Institute of Technology. Cambridge, MA (1985)

43. Kumar. A. and A. Lucia. "Distillation Optimization-. Paper 55f at Annual AIChE Meeting. Miami Beach (1986)

46. Lang. Y-D and L.T. Biegler. "A Unified Algorithm for Flowsheet Optimization", Comp. and Chem. Engr., to appear (1986)

47. Lang, Y-D, L.T. Biegler and LE. Grassmann. "Simultaneous Optimization and Heat Integration with Process Simulators" submitted to Comp. and Chem. Engr. (1986)

48. Lumhoff. B. etal. "User Guide on Process Integration for the Efficient Use of Energy". I. Chem. Eng. (1982)

49. Locke. M.H. and A.W. Westerfaerg. "The ASCEND-H System-An Application of a Successive Quadratic Programming Methodology". Comp. and Chem. Eng., 7, p. 615 (1983)

50. Locke, M.H.. RJEdahl and A.W. Wesierberg. "An Improved Successive Quadratic Programming Optimization Algorithm for Engineering Design Problems-, AIChE J. 29.5. p. 871. (1983)

51. Maratos. N.. Ph. D Thesis, University of London (1978).

52. Marquardt, D.W.. "An Algorithm for Least Squares Estimation of Nonlinear Parameter!", SIAM J. AppL MattL, 11. p.431(1963)

53. Meicalfe, S.R. and JJ>. Perkins, "Information Flow in Modular Flowshceung Systems", Trans. L Chem. E., 56. p. 210, (1978)

54. Motard. R.L., M. Shacham and E.M. Rosen, "Steady State Process Simulation". AIChE J. 21,3. p. 417 (1975)

55. Murtagh, B.A. and MA Saundcrs, Math. Prog.. 14. p. 41.(1978)

56. Murtagh, BA., and MA Saunders, "MINOS/Augmented Supplementary User's Manual." Report OR/78/6 (1978). The University of New South Wales, Australia.

57. Ng. S-M., and GX. Thompson. "A Symmetric Dimension Expanding Primal-Dual Simplex Algorithm for Quadratic Programs". GSIA Technical Report. Carnegie-Mellon University (1986)

58.Nocedal, J. and MX. Overton, "Projected Hessian Updating Algorithms for Nonlinearly Constrained Optimization", SIAM J. Num. AnaL, 22.5. p. 821 (1985)

59. Papoulias, SA. and LE. Groiimann, "Structural Optimization in Process Synthesis II - Heat Recovery Networks". Comp and Chem Eng^ 7. p. 707 (1983)

60. Perkins, J.D., "Efficient Solution of Design Problems Using a Sequential Modular Flowsheeting Programme", Comp. and Chem. Eng, 3. p. 375 (1979) (1977)

61. Powell. MJJX. "A Hybrid Method of Nunlinear Equations-, in Numerical Methods for Nunlinear Algebraic Equations, P. Rabinowitz (ed). Gordon and Breach, London (1970)

62. Powell. MJ.D., "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations", Dundee Conf. on Num. Analysis (1977)

63. Rosen, E.M. and A.C. Pauls. "Computer Aided Process Design: The FLOWTRAN System", Comp. and Chem. Eng.. 1.1, p. 11(1977)

64. Sargent, R.W.H.. "On the Parametric Variation of Nonlinear Programmes and the Convergence of Nonlinear Algorithms", presented at 10th Intl. Symp. on Math. Prog.. Montreal (1979)

65. Sargent, R.W.H., "A Review of Methods for Solving Nonlinear Algebraic Equations". Foundations of Computer Aided Chemical Process Design, p. 27. Man and Seider(eds.) (1981)

66.Schiukowski. K.. "The Nonlinear Programming Algorithm of Wilson, Han and Powell with an Augmented Lagrangian Type Line Search Function". Numer.Math..38.p.83(1982)

67. Seader, JJX, WD. Seider and A.C. Pauls. FLOWTRAN Simalatkm - An Introduction. 2nd edition. CACHE Corp. (1977)

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

69.Stadtherr. MA, RJ>. Laroche and R.C. Alkirc, "A Simultaneous Modular Approach lo Electrolytic Process Optimization Using ASPEN". Paper 50c. 1986 National AIChE Meeting. New Orleans. LA (1986)

70. Swancy. R£. and LE. Grossmann, "An Index for Operational Flexibility in Chemical Process Design". AIChE J.. 31.4, p. 621 (1985)

71. Toman. J.A^ "On Scaling Linear Programming Problems". Man. Prog. 4. p. 146 (1975)

72. Trevino-Lozano R. ScD Thesis D[illegible] of Chemical Eng., MIT. Cambridge. Massachusetts (1985).

73. ViUadsen, J. and M. L. Michdsen. Solution of Differential Equation Models by Polynomial Appradmalkm. Prenuce-Hall Inc.. (1978).

74.Wesunan, K. and A. Luda."Low Cost Solutions to Separation Problems by a Hybrid Fixed Point Algorithm", in Foundations of Computer Aided Process Design - II, Westering and Chien (eds.). CACHE Committee, (1984)