Distributed Stream Method for Tray Optimization

Y-D Lang and L. T. Biegler
Dept. of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

Determining the optimum number of trays in a distillation column has been an important process optimization problem for over 50 years, which over the past decade has been addressed successfully as a mixed-integer nonlinear-programming (MINLP) problem. But tools for solving MINLPs are not widespread, especially in connection with detailed simulation models. In a differentiable distribution function (DDF), all streams around a column, except top and bottom products, are directed to all of the column trays and the distributed flow rate of entry or exit streams is directed to a specific tray based on the value of its DDF at that tray. It allows the placement of feeds, sidestreams, and number of trays in the column to be continuous variables in the DDF, and in the optimization problem it eliminates the need for integer variables. Instead, the tray optimization problem is formulated as an NLP using the original MESH model. It is critical to deal with trays that have no liquid or vapor flows. To describe this phenomenon properly in the optimization, complementary constraints are formulated and added to the NLP, by taking advantage of a smoothing algorithm developed by Gopal and Biegler for phase equilibrium. Also, since pressure in each tray affects distillation calculations substantially, the pressure drops across each tray depend on the optimal number of trays and a related smoothing function formulation removes the pressure drop on dry trays. Pressure in the condenser or reboiler is then adjusted accordingly. The results of the methods applied to three distillation problems show significant cost reductions in distillation design.

Introduction

Design and optimization of equilibrium-staged separation (which we term *tray optimization*) has long been regarded as a difficult problem with discrete and continuous decision variables. A number of strategies have been applied for column optimization for over 50 years. In particular, to treat this problem as a simpler continuous variable optimization, several relaxations can be made. These include:

(1) Relaxations in the thermodynamic equations from possibly nonideal equilibrium models to constant relative volatility;

(2) Relaxations of MESH (mass, equilibrium, summation, heat) equations to a continuous form by applying an approximation procedure;

(3) Continuous relaxation of the feed and product-stream locations.

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

Early work for column design and optimization relies on (1) the familiar Fenske, Underwood, and Gilliland shortcut calculations (Holland, 1963), where constant relative volatility is assumed for all trays, mass balances and energy consumption are derived from limiting total and minimum reflux cases, and an interpolation is made between these cases. However, for realistic, nonideal separations this approach produces serious inaccuracies that can lead to misleading column designs.

To deal with nonideal phase equilbrium, a number of studies have focused on the application of approximation schemes (2) to relax the discrete tray-by-tray equations. Brosilow et al. (1968) reformulated these difference equations to a continuous form, and then applied larger finite difference steps to yield a smaller set of equations. A number of studies applied orthogonal collocation for this approximation as well (Cho and Joseph, 1983; Stewart et al., 1985; Wong and Luus, 1980), and this was also applied to finite-element collocation (Srivastava and Joseph, 1987) on multiple column sections.

Following on this work, Stewart et al. (1985) proposed the use of more accurate Hahn polynomials for the orthogonal collocation framework. In addition, they showed the desirable property of equivalence to the tray-by-tray equations when the number of trays equals the number of collocation points. Moreover, in collocation formulations, the tray number and feed location can now be treated as continuous variables. This was first noted by Swartz and Stewart (1986) and applied and refined in a number of studies (Huss and Westerberg, 1996; Seferlis and Hrymak, 1994a,b), both for single-column and multicolumn optimization. Collocation models are especially useful for optimizations with many trays. On the other hand, it may be difficult to choose the level of approximation without trial-and-error calculations.

On the other hand, Sargent and Gaminibandara (1976) were the first to recognize the tray optimization of distillation columns as a mixed-integer nonlinear program. With the development of efficient MINLP strategies (e.g., Viswanathan and Grossmann, 1990), a general MINLP formulation was developed for tray optimization (Viswanathan and Grossmann, 1993). This model forms the basis for numerous extensions to multicolumn systems and azeotropic separations (see Bauer and Stichlmair, 1998; Duennebier and Pantelides, 1999; Smith and Pantelides, 1995). Nevertheless, this MINLP approach can encounter some difficulties in solving the underlying NLP and MILP subproblems, particularly when dealing with the absence of phases and linearization about zero flows. To deal with these issues, Grossmann and coworkers have recently addressed these MINLP problems through generalized disjunctive programming (GDP). In particular, Yeomans and Grossmann (1999) have recently developed a novel tray-by-tray optimization model that only deals with existing two-phase systems, constructs convex envelopes for bilinear constraints, and considers far fewer discrete decisions in the combinatorial search. They have also demonstrated this approach on a number of nonideal column optimizations.

In industrial practice, feed-tray locations and tray numbers are often optimized through finite difference perturbations embedded in continuous optimization solvers on process simulators (with large perturbation steps!). While this approach has led to quite good designs, this practical approach introduces nondifferentiable and discontinuous elements that can lead to failure in the optimization solver. Moreover, such a procedure cannot be implemented on equation-based optimization platforms. Nevertheless, with the widespread use of these process-simulation and -optimization tools, a continuous formulation can still be useful in order to lead to good designs with NLP solvers. As a result, this study considers the following question: Is there an accurate tray-optimization formulation that can be solved using current nonlinear programming algorithms?

Here we present a formulation that can be viewed as combining elements of the formulations in Sargent and Gaminibandara and Viswanathan and Grossmann. In particular, we allow stream locations for the feed, product, and reflux streams to be continuous variables in a differentiable distribution function (DDF). In addition, the formulation must model trays that may not have vapor–liquid equilibrium. We deal with this using complementarity conditions that are approximated using smoothing functions.

In the next section, we discuss the definition of the DDF, its properties, and the distributed stream-tray optimization method (DSTO). In the third section, we present an optimization case study for distillation columns with ideal and nonideal systems. In the fourth section, we show an industrial application. Finally, the fifth section concludes the article and presents directions for future work.

Problem Formulation

Distillation tray optimization considers feed placement and tray number as integer variables. As a result, feed and reflux flow rates are directed to specific points in the column. In this study, we propose a DSTO method where these flow rates can be distributed and directed to every tray according to a DDF. Using this DDF, the locations of the feed, reflux and other sidestreams can now be treated as continuous instead of integer variables. Consequently, this approach allows flow sheets that include distillation units to be optimized by NLP algorithms with optimal placements of streams and number of trays treated as part of this formulation.

The key points of the method include the introduction of a DDF into the distillation model and the formulation of the optimization problem. Also, the ability to handle dry trays through complementarity conditions is critical for a reliable solution of the problem, and we will treat these using a smoothing approach. Both of these features allow us to locate the feed tray and the reflux tray during the optimization. These locations lead to an optimal design, and the trays above the reflux tray will dry up. The implementation of both the distribution function and the smoothing approach are detailed next.

Differentiable distribution function

Let $N_{\max}$ denote an integer that represents the maximum number of trays to be considered in a column, and let $I = \{1, 2, \ldots, N_{\max}\}$ denote the set of integers corresponding to the column trays. In addition, we define N_c and σ as the central value and dispersion factor, respectively, for the distribution function. The DDF used in this study is derived from a discretization of a Gaussian distribution with mean N_c and standard deviation σ. This function offers global support over the column and can also be adjusted to approximate a Dirac delta function. Several other choices for a DDF can be made, as long as they satisfy the properties presented next.

We denote N and a as integer and real variables, respectively, defined by

$$N = \text{int}[N_c], \quad a = N_c - N \tag{1}$$

Now if we let

$$z_i = \frac{(i - N_c)}{\sigma} \quad i \in I \tag{2}$$

$$Q_i = \exp\left(-z_i^2\right) \quad i \in I \tag{3}$$

$$A = \sum_j Q_j \tag{4}$$

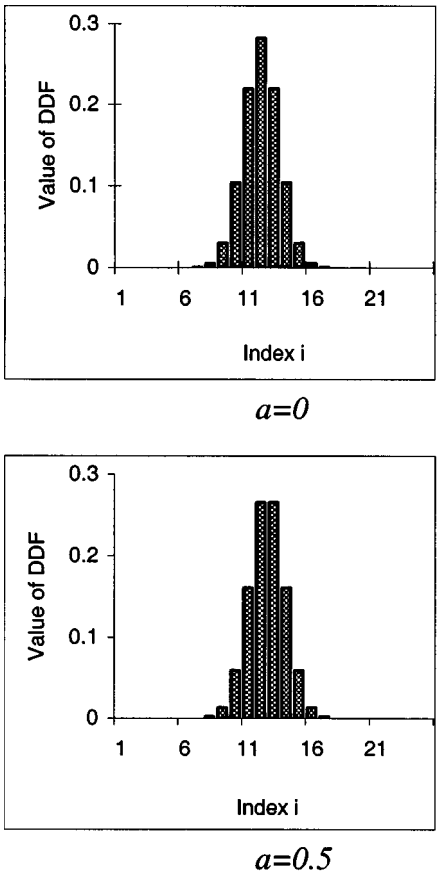


Figure 1. Differentiable distribution function (DDF).

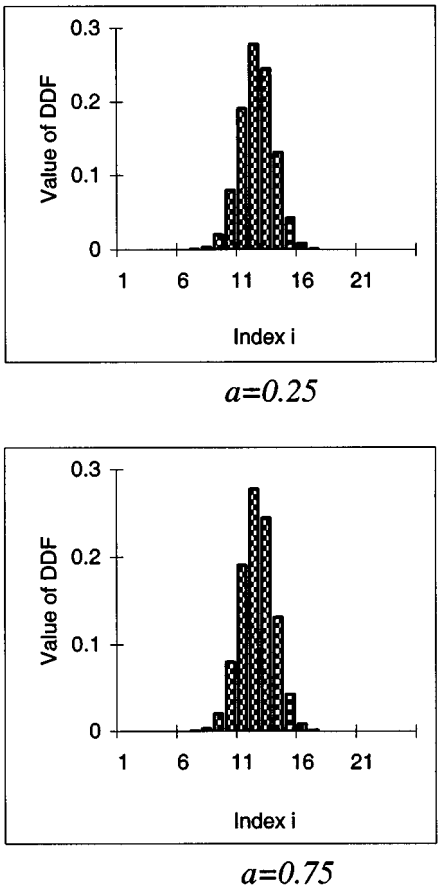


Figure 2. Asymmetry and skewness of the DDF.

then the DDF is defined as

$$d_i = \frac{Q_i}{\sum_j Q_j} = \frac{Q_i}{A} \quad i \in I \tag{5a}$$

Or more clearly and directly, DDF is

$$d_i = \frac{\exp\left[-\left(\frac{i-N_c}{\sigma}\right)^2\right]}{\sum_j \exp\left[-\left(\frac{j-N_c}{\sigma}\right)^2\right]} = \frac{Q_i}{A} \quad i \in I, \quad j \in I \tag{5b}$$

This DDF can be summarized by the following properties:

(P1) The values of DDF are positive everywhere, that is, $d_i \geq 0$, $i \in I$.

(P2) Summation of DDF over I is equal to unity, that is, $\Sigma_j d_j = 1$.

(P3) Symmetry (see Figure 1): If $a = 0$, that is, $N_c = N$, then the values of the distribution function of a series of pairs in set I are symmetric to N

$$d_{N-1} = d_{N+1},\ d_{N-2} = d_{N+2},\ \ldots,\ d_{N-k} = d_{N+k} \tag{6}$$

If $a = 0.5$, that is, $N_c = N + 0.5$, then the values of the DDF of a series of pairs in set I are symmetric to $N + 0.5$

$$d_N = d_{N+1},\ d_{N-1} = d_{N+2},\ \ldots,\ d_{N-k} = d_{N+k+1} \tag{7}$$

(P4) Asymmetry (Figure 2): If $a \neq 0$ and $a \neq 0.5$, then there is no symmetry in set I, that is

$$d_N \neq d_{N+1},\ d_{N-1} \neq d_{N+2},\ \ldots,\ d_{N-k} \neq d_{N+k+1} \tag{8}$$

(P5) The skewness in the asymmetric case depends on the value of a. If $0 < a < 0.5$, skewness is positive as follows, as shown in Figure 2

$$d_N > d_{N+1},\ d_{N-1} > d_{N+2},\ \ldots,\ d_{N-k} > d_{N+k+1} \tag{9}$$

and if $0.5 < a < 1$, skewness is negative as follows

$$d_N < d_{N+1},\ d_{N-1} < d_{N+2},\ \ldots,\ d_{N-k} < d_{N+k+1} \tag{10}$$

(P6) One-sided distribution: If $N_c = 1$ or $N_c = N_{\max}$, the DDF is skewed to one direction only, as shown in Figure 3.

(P7) Effects of σ. The distribution function can be concentrated to a single stage as long as the value of σ is small

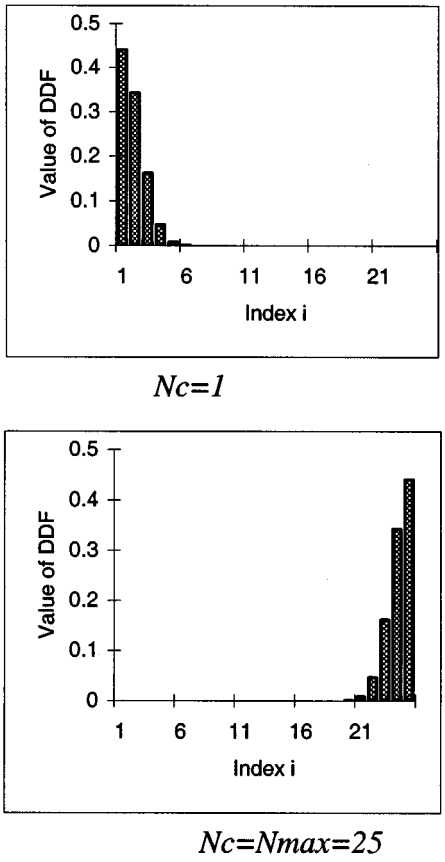


Figure 3. One-sided distribution of the DDF.

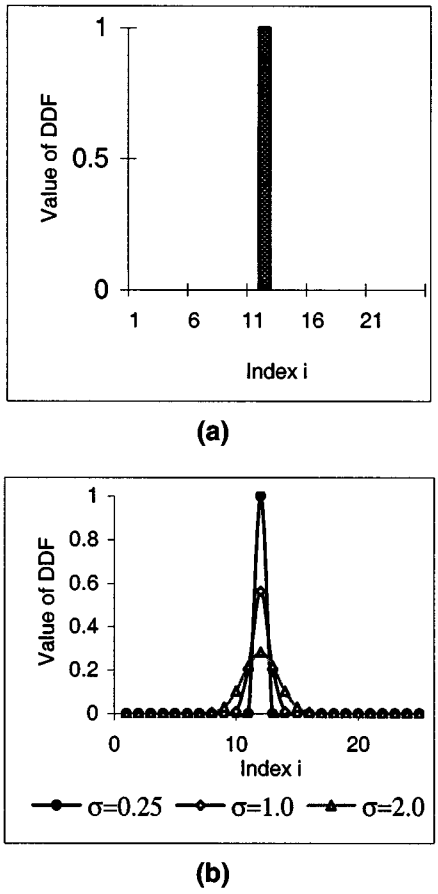


Figure 4. Dispersion of DDF: (a) $\sigma = 0$; (b) different dispersions.

enough, that is

$$d_N = 1 \quad \text{and} \quad d_i = 0 \quad \text{for all} \quad i \neq N, i \in I. \tag{11}$$

Moreover, by manipulating σ, we can get different dispersions of the distribution, from a delta function to a desired flat distribution over the column. Note that, even though i is an integer, the DDF is continuous and differentiable with respect to N_C.

To choose an appropriate value of σ, we need to consider the sensitivity of the DDF with respect to N_c during the optimization. Since this variable represents the locations of feed, sidestreams, and number of trays in the NLP problem, the sensitivity to N_c needs to have a large enough magnitude. The derivative of DDF with respect to N_c is given by

$$GD_i = \frac{\partial d_i}{\partial N_C} = \frac{2Q_i\left[Z_i \cdot A - \sum_j (Q_j \cdot Z_j)\right]}{A^2 \cdot \sigma} \quad i \in I \tag{12}$$

In order to investigate the properties of this derivative function, we calculated the values of GD_N and GD_{N+1} for a from 0.0 to 1.0 in increments of 0.1. Also, σ was considered at five levels—0.25, 0.35, 0.50, 1.00, and 2.00—for each step of a. The results are plotted in Figure 5. Here we focus on the values of derivatives at the neighbors of N_c, namely, GD_N and GD_{N+1}. Because $GD_i < GD_N$ for $i < N$ and $GD_i < GD_{N+1}$ for $i > N+1$, the two derivatives have a significant impact on moving N_c during the optimization. In particular,

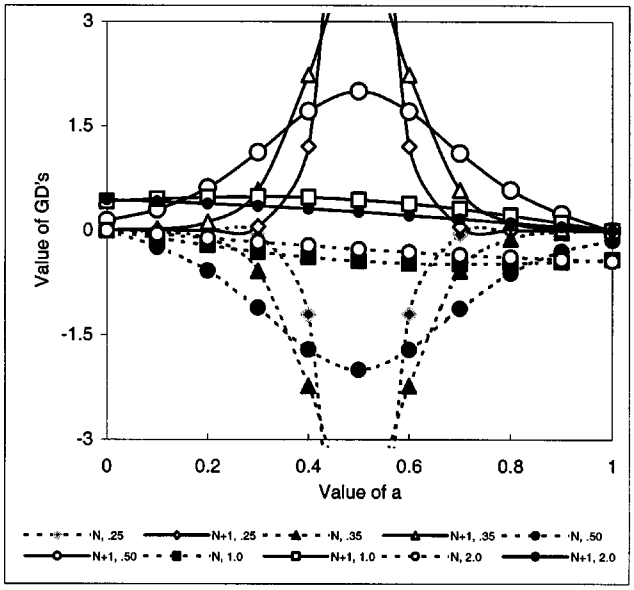


Figure 5. Profiles of gradients of DDF.

the derivatives are related to values of σ and a as follows.

- For $\sigma = 0.25$, the GD change drastically with a. If a is close to zero or one (N_c is close to an integer) all GD's are close to zero and the NLP is insensitive to N_c. On the other hand, for $0.8 > a > 0.2$, GD increases dramatically until it reaches its extreme values ($GD_N = -8.0$ and $GD_{N+1} = 8.0$) at $a = 0.5$.
- For $\sigma > 0.35$, the absolute values of GD are larger than 5.E−3, which we consider a reasonable sensitivity, and the dependence on a decreases with increasing values of σ, as shown in Figure 5.

Therefore from Figure 5, a reasonable sensitivity of NLP to N_c requires $\sigma \geq 0.35$. As seen below, we actually solve the optimization problem by solving a series of NLPs with a decreasing sequence of σ, starting from the previous solution. In this way, the sensitivity effects seen in Figure 5 help to ensure that the effects of the tray location, N_c, are initially considered throughout the column, and are then localized in later stages of the optimization.

Column calculation with DDF

We now consider the column shown in Figure 6. To perform a tray optimization—optimal placement of feed tray(s) and optimal number of the trays—we define the feed and the reflux entering the ith tray as

$$F_i = Ff_i \quad \text{and} \quad R_i = Rg_i \quad i \in I \tag{13}$$

where F and R are the total feed and reflux streams, respectively; F_i and R_i are the flow rates of feed and reflux entering into the ith tray, respectively; f_i and g_i are the values of DDF for the feed and reflux, respectively, that are defined by substituting N_f, σ_f, and N_t, σ_t (central values and dipersion factors for feed and reflux) for N_c, σ, respectively, in Eqs. 2–5.

We then treat F_i and R_i as normal feeds entering the ith tray and formulate the MESH equations. The properties of DDF and its gradient are especially suited for tray calculation using DSTO. In DSTO, we distribute the flow rates of all streams entering the column among its trays according to the DDF. Each tray gets a fraction of the flow rate of these streams equal to the value of DDF for the stream on that tray. From properties P1 and P2, the fraction of flow rate from any stream on each tray must be greater than or equal to zero, and summation of the fractions of any distributed stream on all trays is unity. Also symmetry, asymmetry, and operational skewness will allow a variety of distributions among more than one tray.

Dry trays and complementarity conditions

The only remaining problem is in the shaded portion in Figure 5. Here the trays between N_t and $N_{\max}$ dry up, since there is (almost) no liquid flowing from them and the MESH equations do not hold. To deal with this, Gopal and Biegler developed an NLP formulation based on complementarity conditions for the flash equations. These model the loss of two phases and show that the overall mass and energy balance for the column is still satisfied (Gopal and Biegler, 1999).

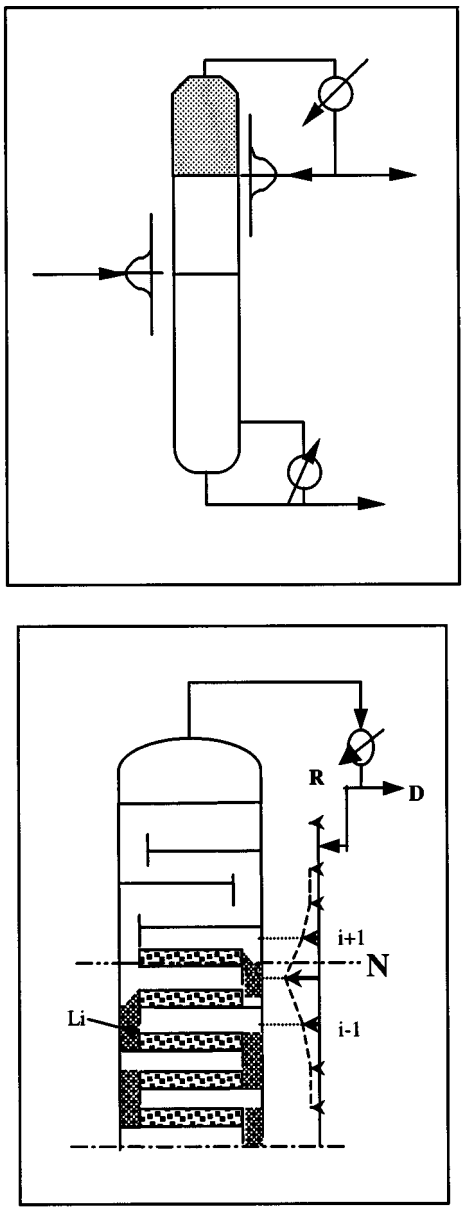


Figure 6. Continuous parameter formulation.

The formulation has been tested on benzene–toluene as well as nonideal (UNIQUAC) five-component systems.

The basic idea here is to modify the MESH equations by modeling the phase-equilibrium equations as

$$y_{ij} - \zeta_i \cdot K_{ij}(T_i, P_i, x_i) \cdot x_{ij} = 0 \tag{14}$$

$$\zeta_i - 1 = s^i_- - s^i_+ \tag{15}$$

$$V_i \cdot s^i_- = 0 \tag{16}$$

$$L_i \cdot s^i_+ = 0 \tag{17}$$

$$s^i_-,\ s^i_+,\ V_i,\ L_i \geq 0 \quad \text{for} \quad i = 1, \ldots, N_{\max} \quad \text{and} \quad j = 1, \ldots, c$$

where y_{ij}, x_{ij} are mole fractions of component j in the vapor and liquid phase, respectively, on the ith tray; V_i and L_i are

the flow rates of vapor and liquid flowing from the ith tray; c is the number of components involved; ζ_i is a corrector for the ith tray; and s^i_-, s^i_+ are slack variables for the ith tray. Their values are relative to the existence of the vapor and liquid phases on the tray according to the complementarity conditions

$$\text{If} \quad \zeta_i > 1 \quad \text{then} \quad s^i_- > 0 \quad \text{and} \quad V_i = 0 \tag{18}$$

$$\text{If} \quad \zeta_i < 1 \quad \text{then} \quad s^i_+ > 0 \quad \text{and} \quad L_i = 0 \tag{19}$$

In our case, for dry trays, the ζ correctors must be less than unity and the slacks, s_+, are greater than zero. Hence, trays with one or two phases can now be modeled directly. Moreover, using an equivalence from Chen and Mangasarian (1996), it can be shown that the complementarity conditions (Eqs. 16 and 17) can be replaced by

$$L_i - \max\left(0,\, L_i - s^i_+\right) = 0 \tag{20}$$

$$V_i - \max\left(0,\, V_i - s^i_-\right) = 0 \tag{21}$$

In order to make Eqs. 20 and 21 differentiable, we replace the "*max*" operator by a smoothing function described in Gopal and Biegler (1999) and given by

$$\max(0,\, x) = \frac{x + \sqrt{x^2 + \alpha^2}}{2} \tag{22}$$

Here α is a small positive smoothing parameter whose value depends on values of x and the tolerance of NLP. (Typically we used values ranging from 10^{-6} to 10^{-8}.)

The formulation in the article is presented for at most a single liquid phase on each tray, although, as shown in Gopal and Biegler (1999), it can be extended to deal with local conditions for two liquid phases. It should be cautioned that Eqs. 14–17 are only necessary conditions for a local minimization of Gibbs free energy. Consequently, solutions obtained with this formulation for nonideal systems, especially with more than one liquid phase, should be checked carefully with a phase stability analysis.

Pressure correction

When we use the complementarity relations to optimize a column, it is also clear that the liquid flows and holdups on the dried trays are zero. Consequently, the pressure drops on each tray must also be corrected. Because equilibrium temperature on each tray is strongly related to its pressure, this correction will affect the temperatures and equilibrium compositions of all phases on each tray, which will heavily impact on the optimal solution of the column.

In distillation calculations pressure drop across each tray is based on the tray hydraulics, or user specification. Nevertheless, it is reasonable to assume that if the liquid flow rate leaving a tray is zero, then pressure drop across it is also zero. With this definition, the pressure on the ith tray can be calculated as follows

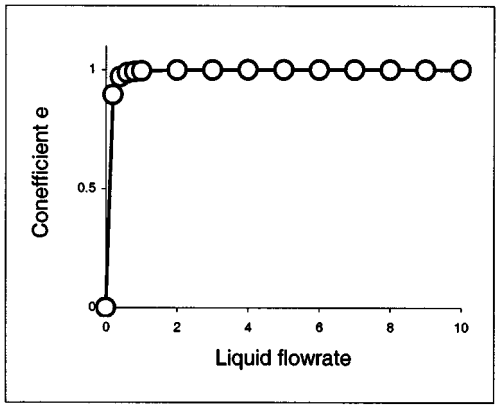


Figure 7. Smoothed coefficient for Eq. 25.

$$\begin{aligned} pd_i &= \varphi(L_i, V_i, p_i, T_i, x_i) \\ p_i &= p_{i-1} + e_i \cdot pd_i, \end{aligned} \qquad i = 2, \ldots, N-1 \tag{23}$$

where pd_i is the pressure drop across the ith tray and e_i is a coefficient determined by the liquid flow rate leaving the ith tray, or

$$e_i = \begin{cases} 1 & \text{if} \quad L_i > 0 \\ 0 & \text{if} \quad L_i = 0 \end{cases} \qquad i = 2, \ldots, N-1 \tag{24}$$

Since this conditional function is discontinuous and nondifferentiable, we use a smoothed function to approximate it

$$e_i = \frac{L_i}{\sqrt{L_i^2 + \beta^2}} \tag{25}$$

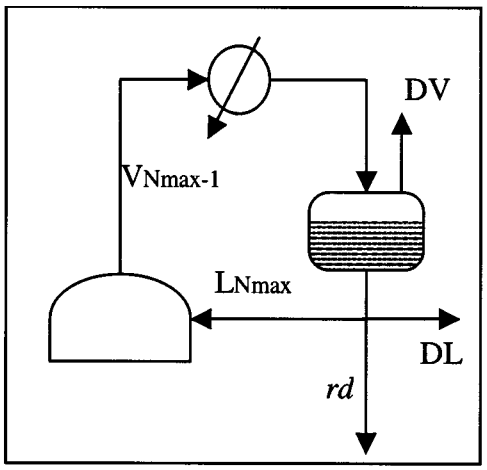
where β is a smoothing parameter (we use 10^{-3} as a typical value) and Eq. 25 serves a similar function as Eq. 22, and is continuous for $L_i \geq 0$. This is shown in Figure 7 for $\beta = 0.1$.

For the optimal number of trays, the total pressure-drop reduction is calculated accurately, and the pressure-drop calculations follow with the column pressure specified either at the top or bottom, depending on the designer's preference.

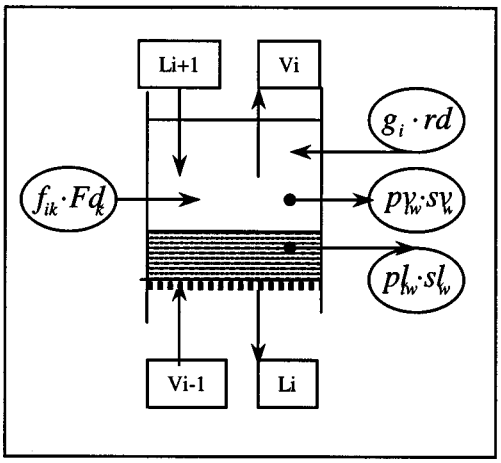
DSTO model for a general distillation column

In order to summarize the equations to simulate a general distillation column, a detailed formulation is described below, starting with variable definitions, illustrated in Figure 8, and followed by the NLP formulation.

Let $N_{\max}$ denote the original number of trays in the distillation column with $I = \{1, 2, \ldots, N_{\max}\}$ the index set of the trays. Here we define $REB = \{1\}$, $CON = \{N_{\max}\}$, and $COL = \{2, 3, \ldots, N_{\max} - 1\}$ as subsets corresponding to the bottom tray (or reboiler), the top tray (or condenser), and those between the reboiler and condenser, respectively. Let m denote the number of feeds, and $K = \{1, 2, \ldots, m\}$ is the corresponding index set of the feeds. Let c denote the number of components in the feeds and $J = \{1, 2, \ldots, c\}$ denote the corresponding index set of the components. In addition, the variables Fd_k, xf_k, shf_k denote the flow rate, mole fraction, and spe-



(a)



(b)

Figure 8. Mass balance for (a) the condenser; (b) the *i*th tray.

cific enthalpy of feeds, respectively. We define sp as the number of sidestream products, and let $W = \{1, 2, \ldots, sp\}$ denote the corresponding index set of those products. The variables sl_w and sv_w denote the liquid and vapor production rate of sidestream product w, and sx_w and sy_w denote the composition in the liquid and vapor phases of sidestream product w.

Moreover, p_i denotes the pressure prevailing on the ith tray and lpd, pd_{con} denote pressure drops across a tray and condenser, respectively. It is assumed that either $p_{\text{reb}} = p_1$, or $p_{\text{con}} = p_{N_{\max+1}}$ and lpd, pd_{con} are given by

$$p_{i+1} = p_i - lpd, \quad p_{\text{con}} = p_{N_{\max}} - pd_{\text{con}} \quad (\text{when } p_{\text{reb}} \text{ is fixed}),$$

or

$$p_N = p_{\text{con}} + pd_{\text{con}}, \quad p_i = p_{i+1} + lpd \quad (\text{when } p_{\text{con}} \text{ is fixed}).$$

The variables L_i, x_i, hl_i and fl_{ij} denote the molar flow rate, composition, the specific enthalpy, and the fugacity of component j, respectively, of the liquid leaving ith tray. Also, V_i, y_i, hv_i, and fv_{ij} denote the corresponding quantities for the vapor, and T_i is the temperature prevailing on tray i. Then

$$fl_{ij} = fl_{ij}(T_i, p_i, x_{i1}, x_{i2}, \ldots, x_{ic})$$

$$fv_{ij} = fv_{ij}(T_i, p_i, y_{i1}, y_{i2}, \ldots, y_{ic})$$

$$hl_i = hl_i(T_i, p_i, x_{i1}, x_{i2}, \ldots, x_{ic})$$

$$hv_i = hv_i(T_i, p_i, y_{i1}, y_{i2}, \ldots, y_{ic}) \quad i \in I$$

The variables, D and B denote the top and bottom product rates, respectively; DL and DV are the liquid and vapor product rates from the top, respectively; and R is the reflux ratio, with R_{total} as the total reflux from the condenser. Let rdf (typically set to a small number, say 0.01) and rd denote the fraction and amount in the total reflux stream, R_{total}, respectively, that enters every tray with

$$L_{N_{\max}} = (1 - rdf) R_{\text{total}}$$
$$rd = rdf \cdot R_{\text{total}}$$
$$L_{N_{\max}} + rd = R_{\text{total}}$$
$$D = DL + DV$$
$$L_1 = B$$

Let f_{ik}, g_i, pl_{iw}, and pv_{iw} denote the feed and reflux fractions that enter tray i and the liquid and vapor sidestream product fractions extracted from tray i, respectively, defined by the DDF. Also, N_{fk}, N_t, N_{pw}, σ_k, σ_r, σ_w^l, and σ_w^v denote the central positions and distribution factors in their corresponding DDF, respectively.

Finally, α and β are small positive smoothing parameters and s_+^i and s_-^i are the slack variables for liquid and vapor on the ith tray, respectively; ζ_i denotes the variable for the ith tray, and e_i denotes the coefficient of pressure drop for the ith tray. Then, if $\zeta_i < 1$, then $s_+^i > 0$, $s_-^i = 0$, $L_i = 0$, and $e_i = 0$; or if $\zeta_i > 1$, then $s_-^i > 0$, $s_+^i = 0$, $V_i = 0$, and $e_i = 1$. Finally, for the two-phase case, if L_i and V_i are nonzero, we have $s_+^i = 0$, $s_-^i = 0$, $\zeta_i = 1$, and $e_i = 1$, $i \in I$. Also, if cooling and heat duties are considered, we let Q_C and Q_H denote the duty of cooling in condenser and heat in the reboiler, respectively.

The modeling equations are as follows. Here we also allow for feeds to enter the reboiler

$$\text{Min } F(N_t, N_f, R, D, B) \tag{26}$$

subject to:

- Phase equilibrium error

$$\sum_j^c x_{ij} - \sum_j^c y_{ij} = 0 \quad i \in I \tag{27}$$

- Definition of reflux

$$R_{\text{total}} = R \cdot D, \quad L_{N_{\max}} = (1 - rdf) R_{\text{total}} \quad rd = rdf \cdot R_{\text{total}}. \tag{28}$$

- Definitions of DDF for all streams

$$f_{ik} = \frac{\exp\left[-\left(\frac{i - Nf_k}{\sigma_k}\right)^2\right]}{\Sigma_j \exp\left[-\left(\frac{j - Nf_k}{\sigma_k}\right)^2\right]} \qquad i \in COL \cup REB, \quad k \in K \tag{29}$$

$$g_i = \frac{\exp\left[-\left(\frac{i - Nt}{\sigma_r}\right)^2\right]}{\Sigma_j \exp\left[-\left(\frac{j - Nt}{\sigma_r}\right)^2\right]} \qquad i \in COL \tag{30}$$

$$pl_{iw} = \frac{\exp\left[-\left(\frac{i - Np_w}{\sigma_w^l}\right)^2\right]}{\Sigma_j \exp\left[-\left(\frac{j - Np_w}{\sigma_w^l}\right)^2\right]} \qquad i \in COL, \quad w \in W \tag{31}$$

$$pv_{iw} = \frac{\exp\left[-\left(\frac{i - Np_w}{\sigma_w^v}\right)^2\right]}{\Sigma_j \exp\left[-\left(\frac{j - Np_w}{\sigma_w^v}\right)^2\right]} \qquad i \in COL, \quad w \in W \tag{32}$$

- Total mass balance

$$L_i + rd + DV + DL = V_{i-1} \quad i \in CON \tag{33}$$

$$L_i + V_i + \sum_w (pl_{iw} \cdot sl_w + pv_{iw} \cdot sv_{iw}) = L_{i+1} + V_{i-1} + \sum_k f_{ik} \cdot Fd_k + g_i \cdot rd \quad i \in COL \tag{34}$$

$$B + V_i = L_{i+1} + \sum_k f_{ik} \cdot Fd_k \quad i \in REB \tag{35}$$

- Component mass balance: $\forall j \in J$

$$(L_i + DL + rd) \cdot x_{ij} + DV \cdot y_{ij} = V_{i-1} \cdot y_{i-1\,j} \quad i \in CON \tag{36}$$

$$L_i \cdot x_{ij} + V_i \cdot y_{ij} + \sum_w \left(pl_{iw} \cdot sl_w \cdot x_{ij} + pv_{iw} \cdot sv_{iw} \cdot y_{ij}\right) = L_{i+1} \cdot x_{i+1\,j} + V_{i-1} \cdot y_{i-1\,j} + \sum_k f_{ik} \cdot Fd_k \cdot xf_{dj} + g_i \cdot rd \cdot x_{ij} \quad i \in COL \tag{37}$$

$$B \cdot x_{ij} + V_i \cdot y_{ij} = L_{i+1} \cdot x_{ij} + \sum_k f_{ik} \cdot Fd_k \cdot xf_{dj} \quad i \in REB. \tag{38}$$

- Enthalpy balance

$$(L_i + DL + rd) \cdot hl_i + DV \cdot hv_i = V_{i-1} \cdot hv_{i-1} + Q_C \quad i \in CON \tag{39}$$

$$L_i \cdot hl_i + V_i \cdot hv_i + \sum_w (pl_{iw} \cdot sl_w \cdot hl_i + pv_{iw} \cdot sv_{iw} \cdot hv_i) = L_{i+1} \cdot hl_{i+1} + V_{i-1} \cdot hv_{i-1} + \sum_k f_{ik} \cdot Fd_k \cdot shf_k + g_i \cdot rd \cdot hl_r \quad i \in COL \tag{40}$$

$$B \cdot hl_i + V_i \cdot hv_i = L_{i+1} \cdot hl_{i+1} + \sum_k f_{ik} \cdot Fd_k \cdot shf_k + Q_H \quad i \in REB. \tag{41}$$

- Pressure correction

$$e_i = \frac{L_i}{\sqrt{L_i^2 + \beta_2}} \quad i \in COL \tag{42}$$

$$p_{i+1} = p_i + e_i \cdot pd \quad i \in COL \tag{43}$$

- Smoothed complementary conditions

$$1 - \zeta_i = s_+^i - s_-^i \qquad i \in I \tag{44}$$

$$L_i + s_+^i = \sqrt{\left(L_i - s_+^i\right)^2 + \alpha^2} \quad i \in I \tag{45}$$

$$V_i + s_-^i = \sqrt{\left(V_i - s_-^i\right)^2 + \alpha^2} \quad i \in I. \tag{46}$$

- Modified phase equilibrium

$$fv_{ij} = \zeta_i \cdot fl_{ij} \quad i \in I, \quad j \in J \tag{47}$$

$\left[\text{or, more specifically, } y_{ij} = \zeta_i K_{ij}(T_i, P_i, x_i) x_{ij},\ i \in I,\ j \in J\right]$

- Specification for sidestream product

$$\frac{\Sigma_i pl_{iw} \cdot sl_w \cdot x_{ij}}{sl_w} = sx_w \quad w \in W \tag{48}$$

$$\frac{\Sigma_i pv_{iw} \cdot sv_w \cdot y_{ij}}{sv_w} = sy_w \quad w \in W \tag{49}$$

with

$$L_i, V_i, s_+, s_-, B, D, R \geq 0$$

NLP formulation and strategy of DSTO

For a tray optimization, there might be different interests in the design and operation of a column or a process flow sheet involving distillation separations. In some cases, the capital investment is the main concern, while in others the operating cost or the product quality are more important. To handle all these cases, we formulated the objective function of NLP to consist of these three factors, with weights for tray

optimization. Given these weights, a DSTO NLP problem can now be solved by any NLP solver. The differentiability of DDF and the properties of its gradient mentioned before are important for DSTO to reach an optimal solution, and the strategy for solving the NLP (Eqs. 26–49) is outlined as follows.

(1) Define the DDF for all streams entering or leaving the columns except the top and bottom products. Each stream has its own central value and dispersion factor.

(2) Provide an initialization for the NLP using any suitable scheme, given all σ's, α's, β's, and initial central values in DDF for each distributed stream.

(3) Solve the NLP using a suitable solver. If the problem is converged and the distribution is narrow enough, then stop.

(4) Otherwise, decrease the dispersion factors σ's, (and α's, β's if necessary) and go to 2.

Since the MESH equations (Eqs. 33–41) are very similar to most distillation models, we note that in step 2 the method allows for any initialization strategy generally used for these models. Once a feasible point is obtained from Eqs. 33–41, 48, and 49, one can calculate the additional variables (ζ, e_i, s_+, s_-, etc.) by direct assignment in Eqs. 42–47. In this study, the optimization formulation (Eqs. 26–49) is implemented in an equation-based environment and all cases are initialized with a feasible point as well as a sufficiently large number of trays and reflux ratio. In the next two sections we consider a number of examples that demonstrate the effectiveness of this NLP formulation for tray optimization. The following section considers the optimization of simple columns, while a coupled column system is optimized in the fourth section.

Tray Optimization for Simple Columns

In order to test the behavior of DSTO, we launched a series of designed experiments with two examples, both simple single-feed columns. The first is used to separate 70% benzene and 30% toluene in the feed. The second is a column dealing with a mixture of five components whose thermodynamic properties are described by UNIQUAC.

Optimal feed placement

In the experiments with the benzene/toluene system we first consider a feed-tray optimization with the reflux ratio

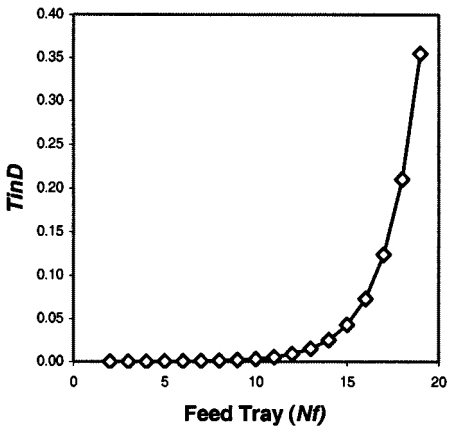


Figure 9. *TinD* depending on feed placement.

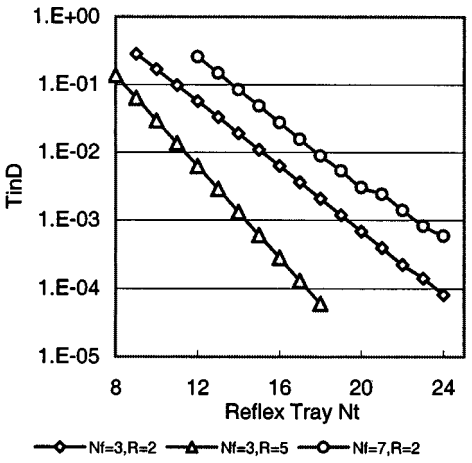


Figure 10. Dependency of *TinD* on N_t.

and total number of trays fixed at 2.0 and 25, respectively. In order to investigate the relationship between N_f and the impurity of toluene in distillate, *TinD*, we ran a series of simulations with a sequence of feed trays from 2 to 19, and *TinD* is plotted vs. N_f in Figure 9. The plot reveals a monotonic relationship. It is noted that *TinD* is very shallow and insensitive to N_f when N_f is below 10 and the optimal N_f for minimizing *TinD* is from 2 to 9. Here, the DSTO NLP formulation works consistently well with this prediction. (For simplicity, we kept $\sigma = 1.0$.) In fact, when the initial value of N_f is above 11, the optimizer with DSTO always gets an optimal N_f around 9 and the value of objective function, *TinD*, is about 0.002. If we amplify the sensitivity of *TinD* to N_f by multiplying the objective by more than 100, the optimal N_f is about 2.7 and *TinD* is $1.2\text{E}-4$, which is close to the global optimum of $N_f = 2$. According to this experiment, we can conclude that the method is able to find optimal feed placement N_f^*.

Minimizing number of total trays

We now investigate how *TinD* depends on the number of total trays, N_t. To predict the results we consider three different combinations of N_f and reflux ratio R, as shown in Figure 10 and Table 1. The relationship of *TinD* with N_t is always monotonically decreasing for the different cases shown by the simulation results in Figure 10. It can be seen that the larger reflux leads to a smaller N_t for the same *TinD* and N_f (compare set 1 and set 2), and we can use these relationships to predict the optimal N_t. Also, because of the low values of *TinD*, it is not sensitive to N_t as N_t increases.

For each of these sets, ($N_f = 3$, $R = 2$), ($N_f = 3$, $R = 5$), and ($N_f = 7$, $R = 2$), we minimize the total number of trays,

Table 1. Optimal Results of the Method

Set	N_f	R	N_t^*
Set 1	3	2	19.38
Set 2	3	5	14.46
Set 3	7	2	21.55

N_t, subject to $TinD \leq 0.001$ and present the results in Table 1. Here the optimal solutions coincide very well with the simulated results in Figure 10. It is worth indicating that since N_t^* is less than the maximum number of 25 trays and there were some dried trays that were treated by the smoothed complementary condition. As a result, we are confident that the method will find the minimum number of trays subject to product constraints.

Results for a comprehensive objective function

An experiment was designed to test the capability of the method to handle complex trade-offs for tray optimization. We consider a composite objective function with three weight factors, quality of distillate represented by *TinD*, reflux ratio R, and number of trays N_t, of the form:

$$\min \text{obj} = w_q \cdot TinD + w_r \cdot R + w_n \cdot N_t$$

Changing the weights represents a different emphasis on the three factors. If lower operation cost is preferred, w_r should be relatively large; if capital investment is emphasized, w_n should be increased. The data for a series of experiments with a variety of weight combinations are listed in Table 2, along with optimal values for N_t, N_f, and R.

In Table 2, three groups of weights are considered. In group I-A, the weight of N_t, w_n, is incremented, with the other two weights constant. That is equivalent to emphasizing the number of trays to reduce investment cost. With increasing w_n, the optimal feed tray, N_f, moves to its lower bound, N_t reduces correspondingly until it reaches its lower bound, and the reflux ratio R and *TinD* increase accordingly.

In group I-B, w_r, the weight factor for R, is incremented, while w_n and w_q are fixed at 0.001 and 1, respectively. This means we want to give priority to reducing operating costs. As shown in Table 2, the optimal R decreases with increasing w_r and hits its lower bound when w_r goes above 0.01. Meanwhile N_t increases to about 18. It should be noted that the optimizer did not raise N_t further to reduce the value of *TinD*, which remains at a value of about 0.003. The reason for this is a tiny sensitivity of *TinD* to N_t at this point.

Finally, in group I-C, we exaggerate the purity of the distillate and give the minimization of *TinD* the highest priority. Interestingly, N_t increases up to 18, but remains less than its upper bound 20. N_f hits its lower bound in order to reduce *TinD*, while R does not increase much. As seen in Table 2, all of these reduced *TinD* to a trace amount. Also when w_n and w_r are much larger, N_t and R hit their lower bound, as shown in last two experiments in I-C. To compensate, the impurity in the distillate, *TinD*, becomes significantly larger.

Each problem in Table 2 has 451 equations with 459 variables. The CPU time needed to solve those problems with CONOPT in GAMS ranges from less than one second to less than one minute on a 350-MHz Pentium II computer, depending on whether the starting point is an arbitrary feasible point or a previous solution from Table 2.

Experiment with five-component separation using UNIQUAC

A column with a more complex mixture is now considered to investigate the performance of DSTO. A mixture of 15% methanol, 40% acetone, 5% methyl acetate, 20% benzene, and 20% chloroform is fed into a column. As in the first example, an objective function is formulated in order to test the method

$$\min \text{obj} = w_M \cdot x_M + w_C \cdot x_C + w_r \cdot R + w_n \cdot N_t$$

where x_M = liquid methanol mol fraction in the bottom; x_C = liquid chloroform mol fraction in the distillate; R = reflux ratio; N_t = total number of the trays in the column; N_f = location of the feed tray; w_M = weight factor of x_m; w_C = weight factor of x_c; w_r = weight factor of R; and w_n = weight factor of N_t.

The maximum number of trays is set to 16, including the reboiler. Among the five components, methanol is lightest and chloroform is heaviest, and we use them to represent impu-

Table 2. Experimental Results of Example 1

Group	No.	w_q	w_r	w_n	Optimal Results			
					TinD	N_f	N_t	R
I-A	1	1	0.01	0.001	0.004135	3.453112	16.50446	2*
	2	1	0.01	0.01	0.010332	2.461897	12.59008	2.847300
	3	1	0.01	0.1	0.056773	2*	8.562828	5.535150
	4	1	0.01	1	0.094481	2*	8*	7.111361
	5	1	0.01	10	0.094069	2*	8*	7.152549
I-B	6	1	0.001	0.001	0.001235	4.534728	16.52286	3.751030
	7	1	0.01	0.001	0.004135	3.453112	16.50446	2*
	8	1	0.1	0.001	0.002826	3.524570	17.48536	2*
	9	1	1	0.001	0.002846	4.483547	17.55760	2*
	10	1	10	0.001	0.002607	4.516653	18.44883	2*
I-C	11	1	0	0	0.000675	4.537108	17.46627	3.482471
	12	10	0	0	0.000382	4.166245	15.47988	6.069465
	13	100	0	0	0.000000	2*	18.03341	6.980814
	14	1	1	1	0.458452	2.864134	8*	2*
	15	1	10	1	0.458452	2.859752	8*	2*

*Variable hits its lower bound.

rity in products from the bottom and the top, respectively. Here, if we wish to reduce the impurity in the bottom product, we can argue qualitatively that the total number of trays and the number of trays in the stripping section must increase, that is, the feed tray moves up. Also if the impurity in the distillate dominates, we expect the reflux ratio R to increase. In each group, we only changed one of the weights and kept the rest constant. The experimental data and results are listed in Table 3.

In group II-A, the weights of x_m are increased from 1 to 10,000. Here the reflux ratio and N_t change little, while the location of the feed tray moves up approximately three trays. This results in reducing x_m from 0.0278% to 0.0271% and increasing the impurity of distillate, x_C, from 2.47% to 3.71%. Since x_m is already small, it is not easy to reduce it much further by increasing R. On other hand, since N_t is close to its upper bound, 16, there is no additional room for it to move up and N_t changes negligibly. In group II-B, the emphasis is on reducing the heaviest impurity in the distillate. When w_C is increased, R increases dramatically. As a result, x_c is reduced to 0.90% for $w_C = 1{,}000$, a 62% reduction compared with 2.41% when $w_C = 1$. In group II-C, the weight of the reflux ratio is increased. Here R hits its lower bound at 2.0 when w_r is greater than or equal to 0.1. In group II-D, the changing weight emphasizes the total number of trays in the column. With increasing w_n, N_t is reduced. Correspondingly, the feed tray moves down a little and x_c and x_m are changed as the reflux ratio is adjusted accordingly.

There are 1,186 equality constraints with 1,193 variables in this example. We used previous solutions to initialize the 20 NLPs solved in Table 3 with CONOPT in GAMS. The first solution spent 661 CPU seconds. Then the CPU time for solving the remaining problems varied from about 3 seconds to 170 CPU seconds on a 350-MHz Pentium II computer, depending on whether the starting point is a feasible point or a previous solution from Table 3.

From the experiments with the two case studies, we can conclude that the DSTO to optimize distillation columns is practical in implementation and the optimal results are also verified by qualitative arguments.

Coupled Column Optimization

In the preceding section, we showed through designed experiments that the method works well with simple individual columns. In this section, we continue our tests with more complicated systems and show that the method can be used in flow sheet optimization as well. We consider a mixture of components A, B, and C, which represent nitrogen, argon, and oxygen in an air separation. Here two columns are coupled in a recycle loop, as shown in Figure 11, and our objective is to minimize the number of trays for fixed feeds with the following quality and product rate specifications satisfied:

- 5,000 lbmol/h of liquid oxygen at 99.6% purity
- 15,000 lbmol/h of gaseous nitrogen with less than 1 ppm oxygen
- 200 lbmol/h of liquid argon at less than 1% oxygen
- Waste stream must be less than 12% of the total feed streams.

The first column, ARGCOL, has a single feed and no sidestreams, while the second, LPCOL, has six feeds including reflux and recycle streams and two sidestreams. The DSTO method is used to formulate a general nonlinear programming problem in GAMS using CONOPT as the solver. Here equilibrium data are calculated using a two-suffix Margules model. This was used instead of the usual equation-of-state model due to space limitations within the GAMS modeling environment.

Optimal results with all specifications satisfied showed that there should be 57 trays in the ARGCOL and 73.5 trays (at

Table 3. Experimental Results of Example 2

						Optimal Solutions				
Group	No.	w_M	w_C	w_r	w_n	N_f	N_t	R	X_m in B	X_c in C
II-A	1	1	1	0.01	0.001	7.918	15.716	2.948	2.781×10^{-4}	2.471×10^{-2}
	2	10	1	0.01	0.001	7.896	15.731	2.948	2.781×10^{-4}	2.467×10^{-2}
	3	100	1	0.01	0.001	7.874	15.764	2.945	2.781×10^{-4}	2.444×10^{-2}
	4	1,000	1	0.01	0.001	8.060	15.859	2.895	2.778×10^{-4}	2.488×10^{-2}
	5	10,000	1	0.01	0.001	11.002	16**	2.827	2.705×10^{-4}	3.712×10^{-2}
II-B	6	100	0.1	0.01	0.001	11.356	15.722	2*	2.736×10^{-4}	5.223×10^{-2}
	7	100	1	0.01	0.001	7.915	15.818	2.969	2.780×10^{-4}	2.417×10^{-2}
	8	100	10	0.01	0.001	7.433	16**	6.665	2.869×10^{-4}	1.196×10^{-2}
	9	100	100	0.01	0.001	8.609	16**	14.657	2.942×10^{-4}	9.124×10^{-3}
	10	100	1,000	0.01	0.001	8.827	16**	15.448	2.945×10^{-4}	9.034×10^{-3}
II-C	11	100	1	0.001	0.001	7.050	15.561	4.711	2.838×10^{-4}	1.547×10^{-2}
	12	100	1	0.01	0.001	7.901	15.721	2.998	2.781×10^{-4}	2.407×10^{-2}
	13	100	1	0.1	0.001	8.802	15.901	2*	2.785×10^{-4}	4.067×10^{-2}
	14	100	1	1	0.001	8.815	15.885	2*	2.784×10^{-4}	4.070×10^{-2}
	15	100	1	10	0.001	8.829	15.876	2*	2.784×10^{-4}	4.073×10^{-2}
II-D	16	100	1	0.01	0.0001	7.235	15.572	2.889	2.820×10^{-4}	2.480×10^{-2}
	17	100	1	0.01	0.001	7.256	14.937	2.930	2.833×10^{-4}	2.646×10^{-2}
	18	100	1	0.01	0.01	7.388	14.444	3.111	2.836×10^{-4}	2.678×10^{-2}
	19	100	1	0.01	0.1	7.388	14.444	3.111	2.836×10^{-4}	2.677×10^{-2}
	20	100	1	0.01	1	6.864	14.441	3.756	2.842×10^{-4}	2.205×10^{-2}

*Variable hits its lower bound.
**Variable hits its upper bound.

Table 4. Optimal Results of Coupled Columns from ASPEN and DSTO

		N_t^*			N_f^*				
ARGCOL	ASPEN	53			1				
	DSTO	57			1				
		N_t^*	**S2**	**S5**	**S7**	**S11**	**S12**	**S17**	**S9**
LPCOL	ASPEN	89	57	28	50	65	60	28	80
	DSTO	74	51	28	48	52	52	28	63

the lower bound) in LPCOL. For comparison, an optimization with ASPEN using more accurate thermal property calculation has been implemented. Referred to as the base case, it has 53 trays in ARGCOL and 89 trays in LPCOL. The results are shown in Table 4 and Figures 12 and 13. For reference, the profiles from ASPEN are plotted as a solid line without data marks. It is obvious that the profiles are either close in temperature profiles or similar in vapor composition on both columns. It is worth noting that some improvements due to DSTO are expected, even with different thermodynamic properties.

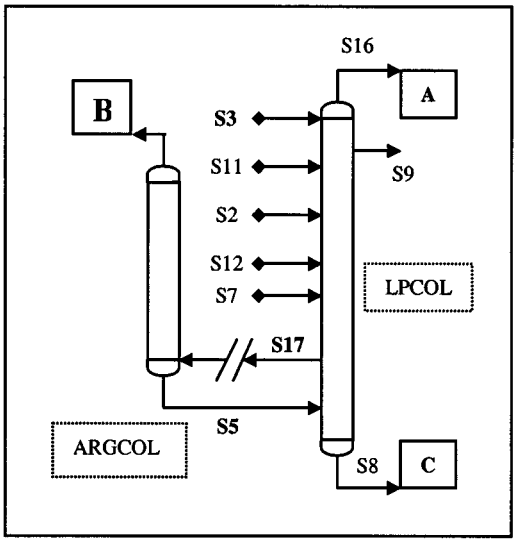


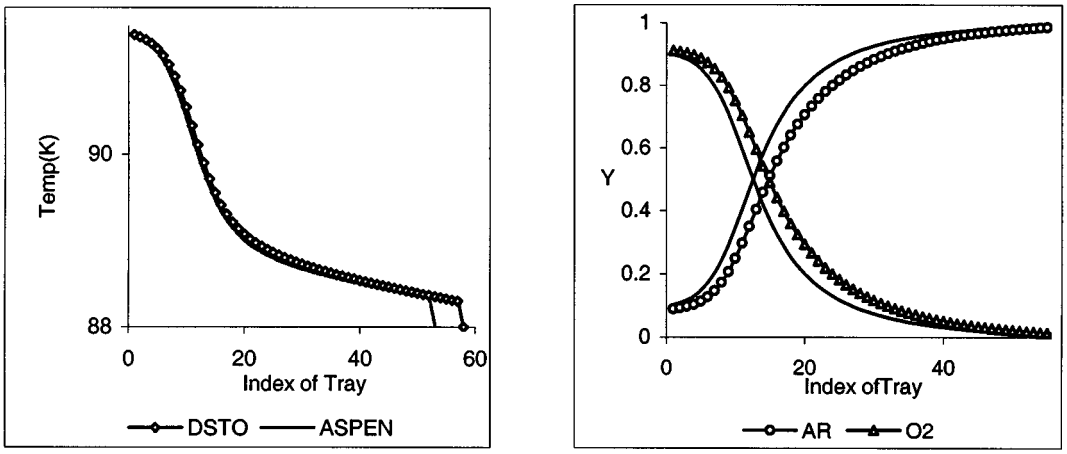
Figure 11. Coupled-column system.

To verify the advantages of DSTO, a simulation with ASPEN using these results has also been implemented. In the ASPEN simulation, the specifications are not satisfied exactly, due to the differences between the available thermodynamic properties (two suffix Margules in GAMS vs. Peng–Robinson in ASPEN). Nevertheless, by adding a few additional trays to the ASPEN simulation, it is easy to satisfy the specifications with a better design than would be obtained with the ASPEN optimizer.

The combined problem has 7,226 equations, 7,249 variables, and 23 degrees of freedom. In order to save CPU time and make the problem easier to converge with CONOPT in GAMS, we split the entire problem into a series of subproblems with the results of the previous subproblem as the starting point for the next one. First we converged the single column, ARGCOL, with fixed feeds, while minimizing its tray count starting from $N_{max} = 85$; this required 86 CPU seconds on a Pentium II 350-MHz computer. Next, we converged the second unit, LPCOL, with fixed feeds, while minimizing its tray count from $N_{max} = 105$; this required 420 CPU seconds. Third, we left the recycle stream open and solved both columns together; this required 36 CPU s. Finally, from this point, we closed the recycle stream and optimized the entire problem with the total number of trays in the two columns as the objective function. The tolerance used here is always 10^{-12}. The last problem was run in three stages in CONOPT by limiting iteration numbers and resuming from the previous point; this step required $(1{,}270 + 2{,}279 + 2{,}395 =)$ 5,944 CPU s.

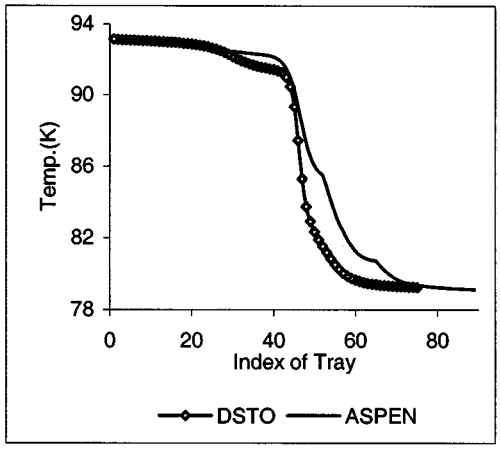
Conclusions and Further Work

This article describes a nonlinear programming formulation for tray optimization. In its original form, tray optimiza-

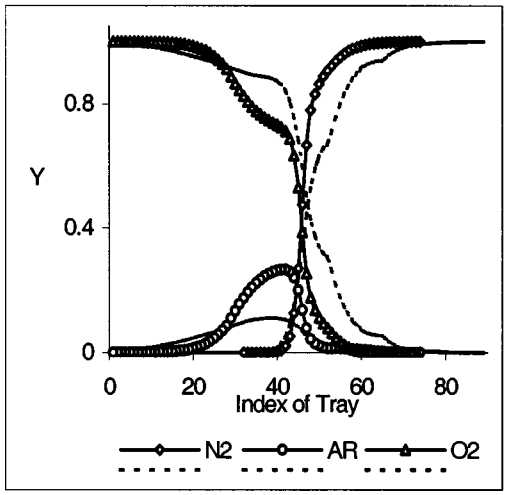


(a) **(b)**

Figure 12. Profile in ARGCOL: (a) temperature; (b) composition in vapor.



(a)



(b)

Figure 13. Profile in LPCOL: (a) temperature; (b) composition in vapor.

tion contains both integer and continuous variables, and can be solved with MINLP algorithms. Instead, we replace the integer variables with continuous ones through the use of a DDF, whose variables are central values of this distribution. As a result, the placement of feed and product trays is obtained via the DDF, and conventional NLP algorithms can now be applied. As a result, the total number of trays is obtained from the distribution of the reflux stream in the column. In this study, we also analyze some properties of the DDF along with the adjustable parameters for this distribution.

A crucial element of this NLP formulation is the ability to model dry trays, especially above the reflux stream. These trays are described by including complementarity relationships for vapor–liquid equilibrium and the loss of one of the phases. In addition, pressure-drop relations need to be modified for trays without liquid holdup. In this study, both of these features rely on smoothing formulations to model discrete events and decisions using continuous variables.

While these smoothing formulations can lead to slight changes in the optimal solution (which can be removed through asymptotic adjustment of the smoothing parameters), the resulting NLP model does include MESH equations and physical property information common to rigorous conventional column models. This formulation is applied to both simple and complex columns, and is modeled and solved in GAMS. A number of examples, drawn from the literature and industry, are considered as NLP problems and compared to optimization results from competing methods, including the Aspen/Plus optimizer. The results show that accurate tray optimizations can be performed efficiently using NLP solvers. Moreover, this NLP formulation can also be applied to extend distillation models in modular simulators; it is fully compatible with specialized initialization strategies for these models.

Finally, future work will deal with the refinement of the NLP formulation and the solution strategy to deal with adjustment of dispersion factors during the optimization step. In particular, a more rigorous treatment of complementarity conditions based on barrier methods (Luo et al., 1996) will be incorporated into this formulation. Also, larger applications for multiple columns and complex flow sheets will be considered.

Acknowledgments

Partial financial support from Air Products and Chemicals, Inc. and Aspen Technology, Inc. is gratefully acknowledged. Also, we gratefully acknowledge the assistance of Claudia Walter.

Literature Cited

Bauer, M. H., and J. Stichlmair, "Design and Economic Optimization of Azeotropic Distillation Processes Using Mixed Integer Nonlinear Programming," *Comput. Chem. Eng.*, **22**, 1271 (1998).

Brosilow, C., R. Tanner, and H. Tureff, "Optimization of Staged Countercurrent Processes," AIChE Meeting, St. Louis, MO (1968).

Chen, C., and O. L. Mangasarian, "A Class of Smoothing Functions for Nonlinear and Mixed Complementarity Problems," *Comp. Opt. Appl.*, **5**(2), 97 (1996).

Cho, Y., and B. Joseph, "Reduced Order Steady State and Dynamic Models for Separation Processes," *AIChE J.*, **29**, 261 (1983).

Duennebier, G., and C. Pantelides, "Optimal Design of Thermally Coupled Distillation Columns," *Ind. Eng. Chem. Res.*, **38**, 162 (1999).

Gopal, V., and L. T. Biegler, "Smoothing Methods for the Treatment of Complementarity Conditions and Nested Discontinuities," *AIChE J.*, **45**, 1535 (1999).

Holland, C., *Multicomponent Distillation*, Prentice Hall, Englewood Cliffs, NJ (1963).

Huss, R. S., and A. W. Westerberg, "Collocation Methods for Distillation Design," *Ind. Eng. Chem. Res.*, **35**, 1603 (1996).

Luo, Z.-Q., J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge Univ. Press, Cambridge (1996).

Sargent, R. W. H., and K. Gaminibandara, "Optimum Design of Plate Distillation Columns," *Optimization in Action*, L. Dixon, ed., Academic Press, London, p. 267 (1976).

Seferlis, P., and A. Hrymak, "Optimization of Distillation Units using Collocation Models," *AIChE J.*, **40**, 813 (1994a).

Seferlis, P., and A. Hrymak, "Adaptive Collocation of Finite Element Models for the Optimization of Multi-Stage Distillation Units," *Chem. Eng. Sci.*, **49**, 1369 (1994b).

Smith, E. M. B., and C. Pantelides, "Design of Reaction/Separation Networks Using Detailed Models," *Comput. Chem. Eng.*, **S19**, 83 (1995).

Srivastava, R. K., and B. Joseph, "Reduced Order Models for Staged Separation Columns, IV. Treatment of Columns with Multiple Feeds and Sidestreams," *Comput. Chem. Eng.*, **11**, 159 (1987).

Stewart, W. E., K. Levien, and M. Morari, "Simulation of Fractionation by Orthogonal Collocation," *Chem. Eng. Sci.*, **43**, 409 (1985).

Swartz, C. L., and W. E. Stewart, "A Collocation Approach for Distillation Column Design," *AIChE J.*, **32**, 1832 (1986).
Viswanathan, J., and I. E. Grossmann, "A Combined Penalty Function and Outer Approximation Method for MINLP Optimization," *Comput. Chem. Eng.*, **14**, 769 (1990).
Viswanathan, J., and I. E. Grossmann, "An Alternate MINLP Model for Finding the Number of Trays Required for a Specified Separation Objective," *Comput. Chem. Eng.*, **17**, 949 (1993).
Wong, K. T., and R. Luus, "Model Reduction of High Order Multistage Systems by the Method of Orthogonal Collocation," *Can. J. Chem. Eng.*, **58**, 382 (1980).
Yeomans, H., and I. E. Grossmann, "Disjunctive Programming Models for the Optimal Design of Distillation Columns and Separation Sequences," *I&EC Res.*, **35**(2), 1637 (2000).

Manuscript received Oct. 26, 2000, and revision received Sept. 10, 2001.