

Towards Computer-Aided Separation Synthesis

Libin Zhang and Andreas A. Linninger

Laboratory for Product and Process Design, Dept of Chemical Engineering,
University of Illinois at Chicago, Chicago, IL, 60607

DOI 10.1002/aic.10689

Published online November 30, 2005 in Wiley InterScience (www.interscience.wiley.com).

A novel computer-aided design methodology is introduced for synthesizing multi-component distillation networks. A single-objective unconstrained design problem formulation circumvents numerical challenges regarding lack of robustness and multiple local minima. A thermodynamically motivated temperature collocation approach drastically reduces the number of state variables occurring in rigorous column equilibrium and component balances. A novel chromosome encompassing all column configurations is proposed to solve the mixed-integer nonlinear programming (MILNP) using a stochastic genetic algorithm (GA) with minimum user input. The general formulation and robust performance of the computer-aided design approach naturally extends to multi-component mixtures. Applications include solutions to optimal sequencing problems for up to five species. The massive parallelism of the approach allows for the generation of complete solution maps classifying the design space into regions of optimality. The method also applies to optimal sequences of azeotropic columns. The temperature collocation approach constitutes a promising new design framework for synthesis problems previously not amenable to computer-aided design and analysis. © 2005 American Institute of Chemical Engineers AIChE J, 52: 1392–1409, 2006

Keywords: separation synthesis, azeotropic distillation, optimum column design

Introduction

Distillative separation processes constitute a significant portion of the total capital investment, energy requirement, and operating expenses of most chemical plants. Hence, the development of systematic approaches for synthesizing optimal separation sequences is of practical importance. The optimal separation synthesis problem can be described as follows: Given a feed mixture of known species, generate the best distillation train to achieve specified purity targets at minimum cost. Its optimal solutions encompass structural decisions like the configuration of columns, succession of intermediate and final products, as well as continuous parameters determining their operating conditions such as column reflux and reboil ratios.

Despite its significance, computer-aided synthesis of column sequences is still not a routine task.

Sargent and Gaminibandara (1976) proposed distillation sequences via superstructure optimization. Heat integration of distillation networks without establishing rigorous column profiles were studied by several groups (e.g., Aggarwal and Floudas, 1990, 1992; Andreovich and Westerberg, 1985; Floudas and Paules, 1988). Other methods determined the feasibility and minimum reflux in sharp separations based on topological concepts (Bausa et al., 1998; Doherty and Malone, 2001; Hauan et al., 1999; Julka and Doherty, 1990, 1993; Wahnschafft et al., 1992; Watzdorf et al., 1999; Westerberg, 2004; Westerberg et al., 2000). Some researchers proposed rigorous mathematical programs for optimal feed location and number of stages (Bauer and Stichlmair, 1995, 1996, 1998; Viswanathan and Grossmann, 1993). Pantelides (Smith and Pantelides, 1995; Dunnebie and Pantelides, 1999) deployed a state task network superstructure with tray-by-tray column models. Optimal heat and mass transfer networks for azeotropic and

Correspondence concerning this article should be addressed to A. A. Linninger at linninge@uic.edu.

reactive separation systems were studied by Pistikopoulos (Ismail et al., 1999, 2001). Shah and Kokossis (2002) developed mixed integer linear programs (MILP) for the synthesis of complex distillation systems. Biegler (e.g. Gopal and Biegler, 1999; Lang and Biegler, 2002) developed smoothing algorithms to eliminate the need for integer variables limiting column design, such as dry trays.

Most existing methods deploy simplified vapor-liquid equilibrium relations (that is, ideal or constant relative volatility) not applicable to many non-ideal and azeotropic mixtures of industrial relevance. Separation network optimization with rigorous tray-by-tray models challenges numerical optimization codes, because the dissimilar mixture of structural decisions and continuous variables introduces discontinuous regions in the search space (Bauer and Stichlmair, 1995, 1996, 1998; Frey et al., 1997). Thus, derivative-based optimization algorithms may fail to locate attractive structural alternatives in separation synthesis. Moreover, the number of possible column sequences and state variables explodes with increasing species numbers. We propose a novel synthesis procedure featuring combinatorial exploration of all relevant structural alternatives with rigorous column profile computation to advance the capacity of computer-aided distillation design. Our novel problem transformation valid for sharp and non-sharp splits extends to general vapor-liquid equilibrium models for any number of species due to a massive problem size reduction. It will be demonstrated how to discover complete separation networks fully automatically with acceptable computational effort.

In this article we introduce a novel feasibility criterion for ascertaining the realizability of arbitrary column specifications based on a thermodynamically motivated transformation of the column tray equations into a dimensionless *bubble point temperature space*. The massive problem size reduction also replaces tray-by-tray vapor-liquid equilibrium and component balance constraints with a scalar metric, such that the resulting unconstrained optimization can be solved robustly by a problem-specific genetic algorithm. Further, we discuss industrially relevant applications including: (i) the optimal separation of mixtures of up to five species, (ii) the creation of complete solution maps for classes of separation problems with varying feed compositions, and (iii) the optimal design of azeotropic distillation sequences. The article closes with a discussion of the results, followed by conclusions.

Methodology

The optimal sequencing problem can be expressed as a constrained mathematical program (Eq. 1) in which d stands for the set of structural and parametric design variables and x are the state variables, such as the tray liquid and vapor concentrations. The equality constraints $h(d, x)$ enforce tray equilibrium, component balances, and the column network connectivity; the inequalities $g(d, x)$ represent the desired final product purity targets and operational limitations. For its numerical solution, this article proposes a novel computational approach combining a problem-specific genetic algorithm (GA) with a rigorous *temperature collocation* distillation design method. A high-level GA constructs different candidate separation configurations by varying the structural as well as key design variables (that is, different product sequences and operating conditions, like reflux ratio, bottoms, and distillate purities, respectively). The assessment of column realizability is

delegated to a feasibility subroutine based on *temperature collocation* on finite elements (Zhang and Linninger, 2004). It ensures the satisfaction of rigorous tray equilibrium and component balances in a *dimensionless bubble point space*, thus eliminating the need for the equality constraints in Eq. 1. The resulting unconstrained problem is searched globally with a stochastic GA. The proposed problem formulation uses the smallest possible design variable set to completely represent the general sequencing problem.

$$\begin{aligned} \min_{d, x} \quad & \text{Cost} \\ \text{s.t.} \quad & h(d, x) = 0 \quad \text{tray balance and connections between columns} \\ & g(d, x) \leq 0 \quad \text{product purities} \end{aligned} \quad (1)$$

Criteria for feasible distillation specifications

The classical *performance* or *rating problem* predicts distillate and bottom compositions for a given feed and known column configuration. An *inverse design problem*, on the other hand, seeks optimal column dimensions and operating conditions for realizing the desired product purities. Solutions to the inverse problem are much harder to converge numerically and may only exist for certain specifications due to thermodynamic separation barriers. Rigorous proofs for the existence of feasible column specifications, in particular for complex column configurations, extractive, or reactive separations, are an open field of research (e.g., Barnicki and Sirola, 2004; Malone and Doherty, 2000; Sirola, 1996; Widago and Seider, 1996).

Recently, a unique distillation problem transformation offers a significant step towards solving *inverse design problems* rigorously (Zhang and Linninger, 2004). According to this novel temperature transformation approach, the realizability of a given separation target (such as three product purity specifications and the desired reflux) can be determined with the help of liquid tray composition profiles expressed in terms of equilibrium tray temperatures as independent variables, as given in Eq. 2 for the rectifying and Eq. 3 for the stripping section.

Continuous rectifying profile equations:

$$\begin{aligned} \frac{\partial x_i^R}{\partial T} = & - \left(x_i^R - \frac{r+1}{r} y_i^R + \frac{1}{r} x_{D,i} \right) \\ & \times \frac{\sum_{j=1}^C \left(\frac{\partial K_j}{\partial T} x_j^R \right)}{\sum_{j=1}^C \left[\left(x_j^R - \frac{r+1}{r} y_j^R + \frac{1}{r} x_{D,j} \right) K_j \right]} \\ = & f_i^R(x^R, r, x_D, K, T) \quad i = 1 \dots C \end{aligned} \quad (2)$$

Continuous stripping profile equations:

$$\frac{\partial x_i^S}{\partial T} = - \left(\frac{s}{s+1} y_i^S - x_i^S + \frac{1}{s+1} x_{B,i} \right)$$

$$\begin{aligned} & \sum_{j=1}^C \left(\frac{\partial K_j}{\partial T} x_j^S \right) \\ \times & \frac{\sum_{j=1}^C \left[\left(\frac{s}{s+1} y_j^S - x_j^S + \frac{1}{s+1} x_{B,j} \right) K_j \right]}{= f_i^S(x^S, s, x_B, K, T) \quad i = 1 \dots C} \quad (3) \end{aligned}$$

The column profile Eqs. 2 and 3 describe the evolution of the liquid tray compositions, x_i^R and x_i^S , as a function of the equilibrium bubble point temperature, T . The stripping profile emerges from the bottom product composition; the rectifying profile starts from the distillate. In order to solve this boundary value problem with unknown intersection, the differential Eqs. 2 and 3 are discretized using orthogonal polynomial approximation. Multiple finite elements, each with a fixed number of collocation nodes, are implemented for high accuracy. The finite elements decompose the temperature range into segments; the nodes of the orthogonal collocation polynomials further divide each element according to a predetermined number of roots. This orthogonal collocation on finite elements (OCFE) method, also known as Spline Collocation (e.g., Villadsen and Michelsen, 1978), converts the system of Eqs. 2 and 3 into a set of nonlinear algebraic equations. These equations are enforced at discrete dimensionless bubble point temperatures $\Theta_{[\alpha, \beta]}$ defined by Eq. 4:

$$\Theta_{[\alpha,\beta]} = \frac{T - T_D}{T_R - T_D} \quad (4)$$

The discretized profile equations are solved for the unknown compositions, $x_i(\Theta_{[\alpha,\beta]})$, according to Eqs. 5 and 6. Rectifying profiles discretized with OCFE:

$$A^R(\Theta_{[\alpha,\beta]})x_{i[\alpha,\beta]}^R - f^R(x_{i[\alpha,\beta]}^R, r, x_D, K, \Theta_{[\alpha,\beta]}) = 0 \quad (5)$$

Stripping profiles discretized with OCFE:

$$A^S(\Theta_{[\alpha,\beta]})x_{i[\alpha,\beta]}^S - f^S(x_{i[\alpha,\beta]}^S, s, x_B, K, \Theta_{[\alpha,\beta]}) = 0 \quad (6)$$

$\Theta_{[\alpha,\beta]}$ refers to the collocation node (i.e., the dimensionless bubble point temperature defined between the bubble point temperatures corresponding to the distillate and bottom compositions, respectively, T_D and T_B) in element α at nodes β . A^R and A^S are the gradient coefficient matrices of the orthogonal polynomials at collocation nodes, $\Theta_{[\alpha,\beta]}$. f^R and f^S are the nonlinear right-hand sides of Eqs. 2 and 3 evaluated at nodes $\Theta_{[\alpha,\beta]}$. $x_{i,[\alpha,\beta]}^R$ and $x_{i,[\alpha,\beta]}^S$ represent the unknown weights (i.e., composition of species i) at node, $\Theta_{[\alpha,\beta]}$.

According to this *temperature collocation* approach, whose theoretical background is discussed in detail elsewhere (Zhang and Linninger, 2004), all liquid composition trajectories originate in bottom and top products and terminate in stationary pinch points, as illustrated in Figure 1. Feasible column profiles exist only within a temperature window delineated by the bubble points corresponding to product compositions and the stable nodes of the pinch equations. A design specification is

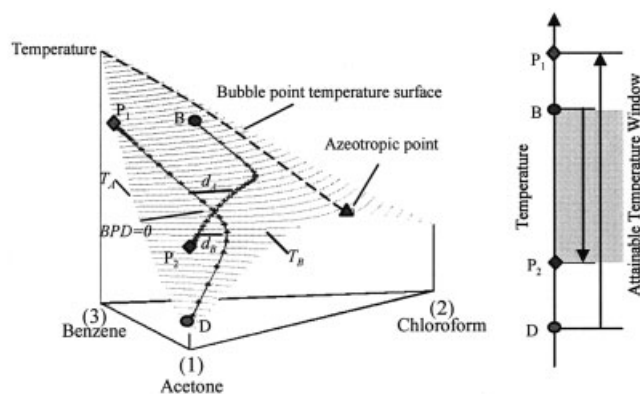


Figure 1. Feasible liquid composition profiles in the bubble point temperature space (left) and attainable temperature window (right).

termed *feasible* or *realizable* with simple distillation if stripping and rectifying profiles intersect. The bubble point distance is the Euclidean difference between a stripping and a rectifying composition profile for a specific bubble point temperature. The minimum *bubble point distance (BPD)* is a precise scalar quantity suitable for formulating a rigorous numerical feasibility criterion. Thus, column specifications leading to zero *BPD* are *feasible*; infeasible designs have non-zero *BPD*, as expressed in Eq. 7:

$$\text{if } \min_{dT} BPD = 0 \rightarrow d \text{ is feasible} \quad (7)$$

In higher dimensions, exact intersection of trajectories is numerically more difficult to attain than in the two dimensional case, since the smallest error may cause the two profiles to miss each other. The use of the minimum *BPD* constitutes a numerically more robust intersection criterion than the exact equality of concentrations at the intersection. Moreover, small *BPD* function values also indicate closeness to a feasible design, so that the *BPD* metric serves as a better guide in the search space than a Boolean hit-or-miss intersection criterion. Temperature collocation offers significant benefits over the regular tray-by-tray models: (i) its independent variable range is finite; (ii) the profile computations expressed in terms of equilibrium tray temperatures are not hampered by singular regions (i.e., saddle and pinch points) requiring infinite numbers of trays to overcome; and (iii) temperature collocation on finite elements drastically was shown to reduce the problem size by two orders of magnitude.

The efficiency of the novel thermodynamically motivated problem transformation has been demonstrated for a wide variety of sharp and sloppy separation problems (Zhang and Linninger, 2004). Apart from its robustness, its decisive advantage is its computational speed. Figure 2 documents a computational result for executing 10,000 design experiments for separating a non-ideal mixture with a high-boiling azeotrope. This computer simulation required only 1,272 sec on a Pentium II 450 Hz PC. The temperature collocation algorithm converges rapidly for both feasible and infeasible specs. For infeasible specifications, the length of the non-zero *BPD* can be

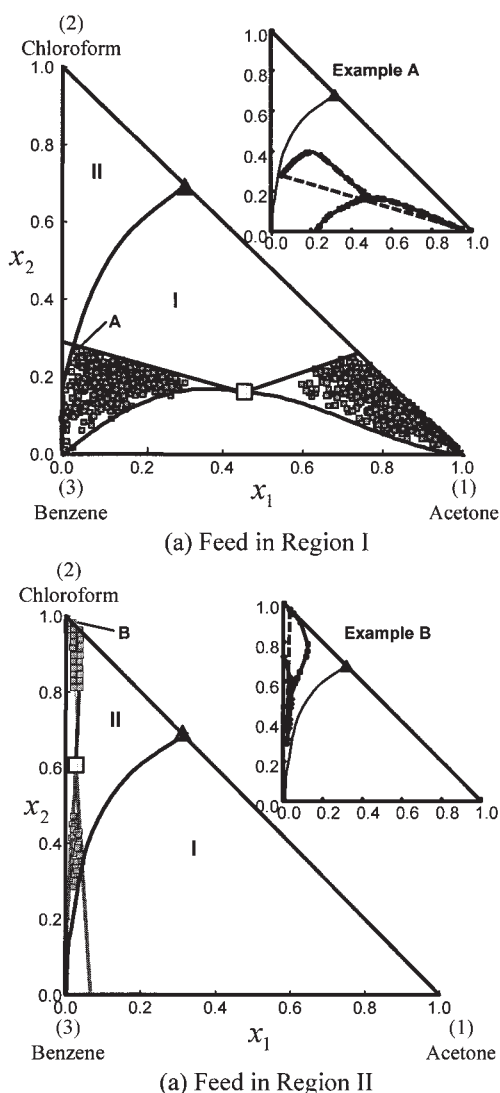


Figure 2. 10000 random design experiments for assessing the feasibility of an azeotropic mixture with two distillation regions (■: feed composition; ▲: azeotrope).

(a) Feed in Region I, (b) Feed in Region II.

interpreted as the distance to a realizable design. Hence, the bubble point distance offers a rigorous metric for the closeness to a desired separation target. The sequel will show that the *BPD* is also a key element for constructing a robust stochastic search algorithm in a massively size-reduced problem space.

Searching structures and key specifications

Genetic Algorithms (GA) (e.g., Gen and Cheng, 2000; Goldberg, 1989; Lewin, 1998; Lewin et al., 1998; Michalewicz, 1992) aim at searching unconstrained optimization problems for their global optimum. Wang et al. (1998) used GA-based optimization for heat integration of distillation systems. Lebreiro and Acevedo (2004) combined a GA with a modular flowsheet simulator to find optimal operating conditions of

fixed distillation trains. However, GAs have not been used successfully to generate rigorous column sequences.

GAs maximize a given performance or *fitness* function by advancing sets of candidate solutions called a *population*. The features (that is, design variables) of each candidate solution are encoded in its chromosome. In each cycle (iteration), all members of the current population participate in a random selection process that picks “winners” with probability proportional to their fitness. Pairs of “winners” are submitted to a mathematical recombination process of their traits (i.e., chromosome) known as *crossover*. Crossover produces two new candidate solutions called *children* or *offspring*, whose chromosomes inherit features of both parents. In addition, the offsprings’ traits may be altered by a random transformation called *mutation*. This iteration loop of crossover and mutation is repeated until a pre-specified generation count is reached or the fitness stagnates.

Remarkably, repeated recombination cycles improve the population’s performance, at least on average, thus solving the original optimization problem without ever computing gradient information. In separation synthesis problems with structural and parametric decision variables, analytical evaluation of the gradient information and higher derivatives of the design objectives and constraints is cumbersome. In column sequencing, this difficulty refers specifically to the derivatives of structural decision variables and nonlinear physical properties, like the vapor-liquid equilibrium models (activity coefficients, fugacity, vapor pressures). The next subsection introduces a problem-specific GA using the smallest possible chromosome and a novel inheritance mechanism for generating globally optimal column sequences without the need of Hessian or gradient information at reasonable computational cost.

Problem Encoding for Optimal Multi-Component Column Sequences. Our novel chromosome contains all the necessary information for completely describing distillation networks composed of N operational tasks for separating a C -component mixture. The main innovation in the proposed chromosome is a specialized column sequencing problem representation that satisfies mass balances between intermediate and final product streams and ensures meaningful logical connectivity of the resulting candidate flowsheets. This type of encoding, leading to a drastic reduction in the number of equality constraints in the column separation problem, has not been discussed in the literature before.

We propose a real-valued string with three segments to avoid binary encoding/decoding steps. Figure 3 illustrates the novel problem-specific chromosome, in which m_i^j stands for the molar flow rates of species j in product i ; the variables r_i are reflux ratios of each column; and e encodes the structural representation of the separation network. Its digits determine whether

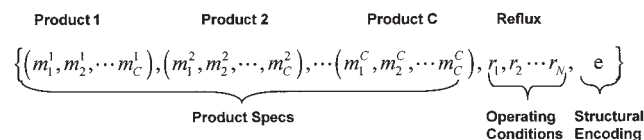


Figure 3. The novel chromosome encoding for the column sequencing design problem.

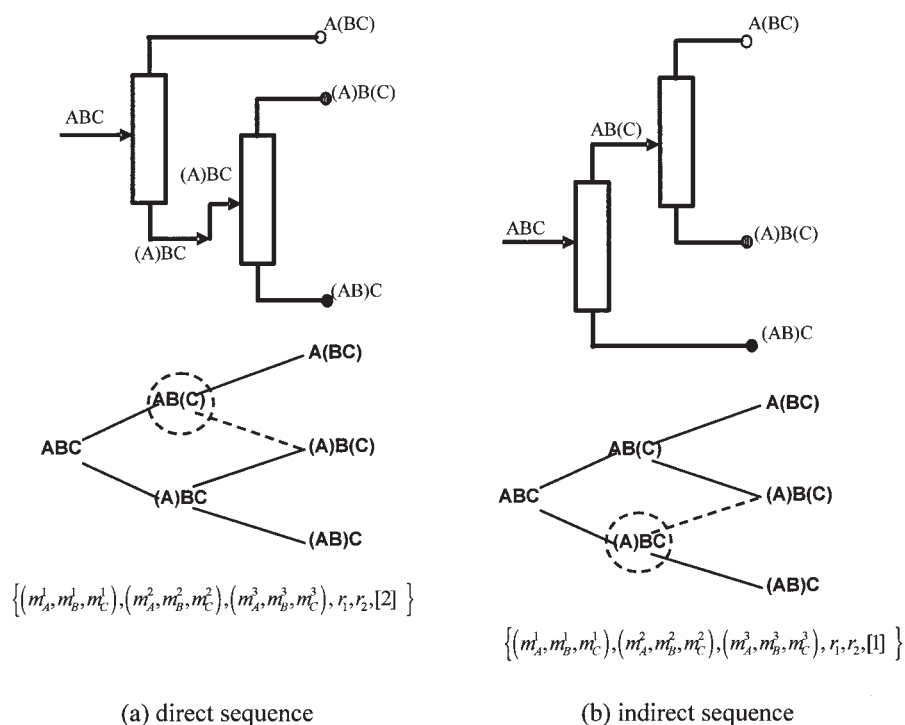


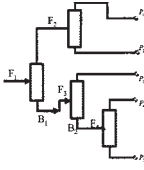
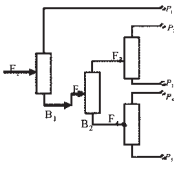
Figure 4. Basic configurations for the ternary separation problem and their chromosome representations.

distillate or bottoms are directed to the subsequent separation step.

Figure 4 exemplifies the representation of the ternary separation network whose structural parameters (e.g., the number one and two) refer to indirect and direct splits. The proposed problem encoding documented in Table 1 requires the smallest

possible design variable set to fully characterize all possible structural alternatives, their operating conditions, and associated distillate and bottom product targets. It also contains the data for inferring all intermediate feeds and other column parameters, such as the reboil ratio or the column height (e.g., Agrawal, 2003). The problem representation generalizes to any

Table 1. The Representation of Structural Parameters, e , for the Separation of a Five Component Mixture

	Options of Second Column	Options of Third Column	Options of Fourth Column
A	1	1	1
B	2	2	2
C		3	3
D			4
Examples			
	1	3	4
$e = 134$			
	2	2	4
$e = 224$			

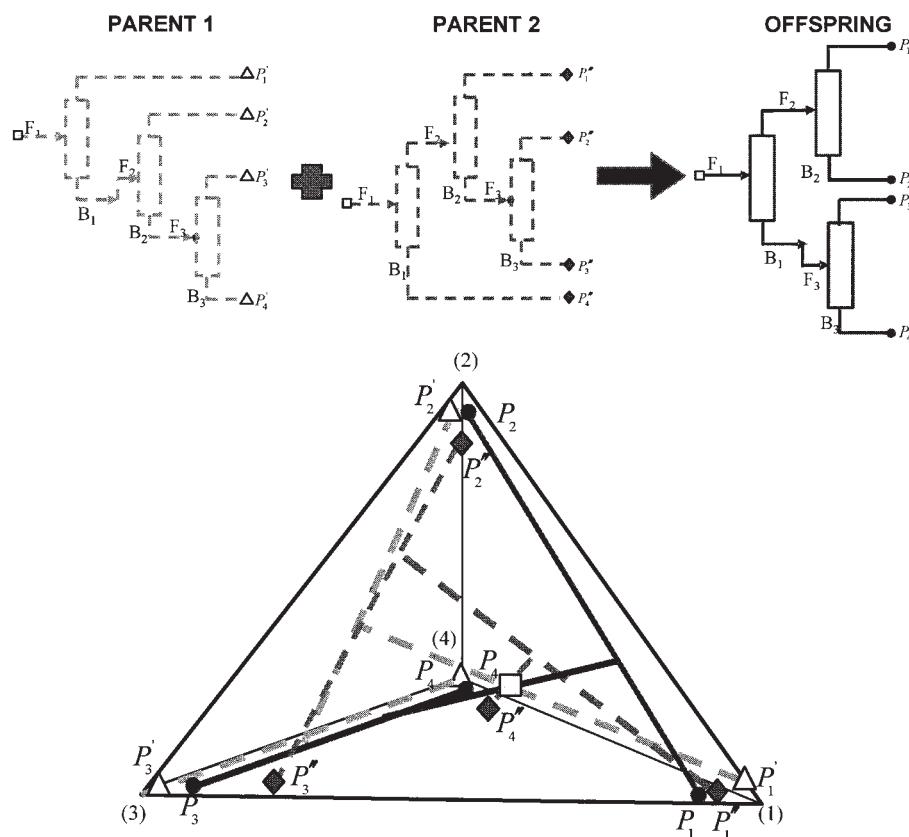


Figure 5. Geometric illustration of the crossover between two quaternary column sequences.

Parent product nodes (P_i', P_i'') lead to an offspring (P_i) combining their structural and operating design decisions.

number of species; it is not limited to sharp splits or a specific preference in the order of product cuts. The next subsection will demonstrate that this problem representation is also advantageous for implementing offspring inheritance by crossover and mutation.

Advancing Candidate Solutions in the Search Space

Crossover. The novel *crossover* operator implements an inheritance mechanism based on geometric considerations. Crossover applies to a pair of parents with probability P_c ; otherwise, the parents reproduce identically. Crossover of structural and parametric traits, such as ϕ , of the sequences' chromosomes can be expressed mathematically by Eqs. 8 and 9:

$$\phi|_{Child_1} = \alpha \phi|_{Parent_1} + (1 - \alpha) \phi|_{Parent_2} \quad (8)$$

$$\phi|_{Child_2} = (1 - \alpha) \phi|_{Parent_1} + \alpha \phi|_{Parent_2} \quad (9)$$

These formulae require a single random variable α drawn from the interval $[0, 1]$ using a standard random number generator (e.g., Press et al., 1995). It is easy to verify that offspring produced with these formulae possess product purities consistent with mass balances inside the composition space. When rounding the outcomes of the structural crossover to the next meaningful structural integer value, this simple inheritance mechanism covers every combination of column

sequences, including direct, indirect, and transition splits for any number of species.

Figure 5 provides a geometric illustration of the one-parameter crossover for a quaternary case. The child's composition can be interpreted as rotational variants of the parents' terminal products spanning a tetrahedron. The mass flow rates of the offspring's product composition resulting from crossover with $\alpha = 0.3$ are given in Table 2. Clearly, the mass balances are satisfied, automatically freeing the ensuing optimization from the need to enforce the component balances between intermediate feeds and their products. The crossover of structural parameter determines the connectivity of the candidate column sequence. Due to the mass balance consistency, the offspring sequence is always structurally possible (that is, no dangling feeds or products); feed and product compositions always lie within the composition space.

Mutation. An additional random element is needed in order to discover remote regions of the design space. *Mutation* is a process of generating candidate solutions by a random alteration of an offspring's traits. A random event with probability, P_m , determines whether the chromosome of a newly created offspring is submitted to mutation. Eq. 10 implements a one-parameter formula for a parametric design variable mutation, with ϕ^U being a proper upper bound of the trait. Structural mutation obeys Eq. 11. The parameters t , N_g , and b , represent-

Table 2. Illustration of Crossover Results of Two Parents with Random Variable ($\alpha = 0.3$): the Last Column Shows the Child's Final Product Mole Flowrate and Its Configurational Parameter, e

$\alpha = 0.3$	Parent ₁	Parent ₂	Child
F (Feed)	{10., 10., 10., 10.}	{10., 10., 10., 10.}	{10., 10., 10., 10.}
	Molar Product Flowrate, m [mol/h]		
P1	{7.0, 1.0, 1.0, 1.0}	{6.0, 2.0, 1.0, 1.0}	{6.3, 1.7, 1.0, 1.0}
P2	{1.0, 7.0, 1.0, 1.0}	{2.0, 6.0, 1.0, 1.0}	{1.7, 6.3, 1.0, 1.0}
P3	{1.0, 1.0, 7.0, 1.0}	{1.0, 1.0, 6.0, 2.0}	{1.0, 1.0, 6.3, 1.7}
P4	{1.0, 1.0, 1.0, 7.0}	{1.0, 1.0, 2.0, 6.0}	{1.0, 1.0, 1.7, 6.3}
	Structural Design Parameter, e		
	{2, 3}	{1, 2}	{1, 2}

ing the current generation index, the maximum number of generations, and an exponential parameter, restrict the perturbation step size of mutations with increasing generation count. The suggested mutation operator always produces viable candidate solutions consistent with mass balances within the composition space. Figure 6 depicts the outcome of a mutation for a quaternary mixture whose product set constitutes a random perturbation of the original tetrahedral product corners. Its column connectivity is also scrambled randomly, but is always structurally consistent by design, as shown in Table 2 and Figure 6.

Parametric mutation operator:

$$\phi' = \phi + (\phi^U - \phi) * \alpha * \left(1 - \frac{t}{N_g}\right)^b \quad (10)$$

Structural mutation operator:

$$e' = e * \alpha * \left(1 - \frac{t}{N_g}\right)^b \quad (11)$$

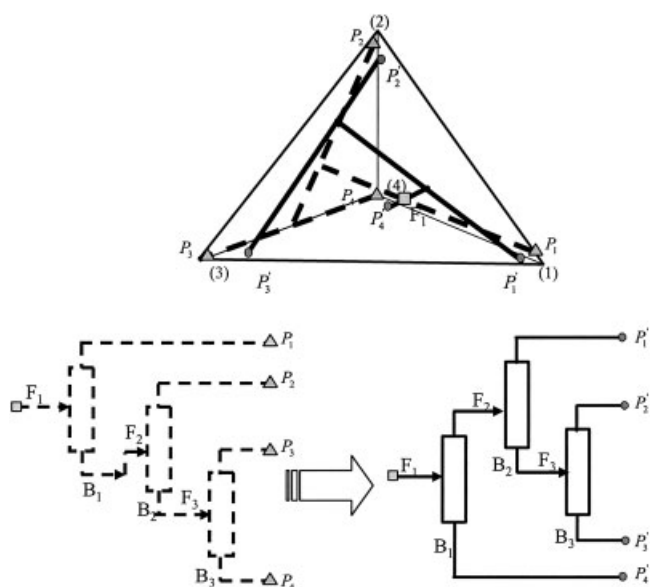


Figure 6. Geometric illustration of the mutation of a candidate solution to a quaternary separation problem in the composition space.

Fitness function

The optimal sequencing problem in Eq. 1 is constrained by a large number of equality constraints to satisfy the tray-by-tray equilibrium and component balances, as well as inequality constraints for enforcing the desired final product purity targets. Despite recent developments (Costa and Oliveira, 2001; Katare et al., 2004), GAs are impractical for optimization problems with a large number of constraints. We have already demonstrated above that the novel inheritance mechanisms (crossover and mutation) automatically handle mass balance and logical connectivity constraints. Furthermore, the BPD function, measuring the degree of realizability of the sequence, eliminates the need for enforcing tray-by-tray balances in the optimization problem. The use of penalty terms instead of tray equations and product inequalities convert the constrained sequencing problem of Eq. 1 into the unconstrained problem (Eq. 12) that is ideal for GA application. The *cost objective* (i.e., inverse of *fitness*) encompasses the operating cost as a function of the reflux ratio (which is indirectly related to reboil ratio s_i via energy balances), $\lambda_i r_i$ the capital cost associated with the required column height, $\mu_i h_i$, and penalty terms for feasibility of each column, P_i^F , and desired product purities, P_j^D .

$$\underset{d}{\text{Min Cost}} = \sum_{i=1}^N \lambda_i r_i + \sum_{i=1}^N \mu_i h_i + \sum_{i=1}^N P_i^F + \sum_{j=1}^C P_j^D \quad (12)$$

Cost. The operating cost accounts for the heating and cooling requirements of each column in the separation train. Simplified estimation models relate them to the reboiler and condenser heat loads (Modi and Westerberg, 1992). Capital cost can be estimated by simplified cost correlations accounting for the column shell, trays, condensers, and reboilers, such as Guthrie Correlations (Douglas, 1988). Column height relates to the dimensionless bubble point temperature range via the back transformation given in Eqs. A-2 and A-3 of the Appendix. The column diameter is usually chosen to avoid flooding (e.g., Douglas, 1988).

Rigorous Assessment of Realizable Column Profiles—Feasibility Penalty. The BPD rendering a single metric for the degree of realizability of the column encapsulates the complete set of column profile equations for the entire separation sequence. Note the massive reduction of the search space by compressing all internal column variables (tray, compositions, tray temperatures, etc) into a single scalar penalty for each column in the sequence (that is, the BPD). In the case studies,

the exponential penalty functions of Eq. 13 were used. f_i constitutes a tuning parameter typically set to a large number when compared to the total annualized cost of the sequence (e.g., total cost of a reference design $\times 1000$). An exponential operator punishes infeasible configurations, thus creating a strong driving force towards realizable column profiles.

$$P_i^F = (\exp^{BPD_i} - 1.0) \cdot f_i \quad (13)$$

Product Purity Penalty. Desired product targets can be enforced by adding penalty terms, P_j^D , as shown in Eq. 14 in which X_j^D is desired product specification and x_j^D is the composition achieved by each candidate solution. s_j is a tunable weight of the penalty chosen larger than the cost in Eq. 12. Two particular choices are worth discussing: (i) *Guaranteed Product Purity.* A weight significantly larger than the cost term ensures hard product purity standards as needed in high purity product recovery with tight products specifications. For this case, set s_j to $100 \times$ total cost of a reference design. (ii) *Purity Quality Trade-offs.* The product purity of optimal solutions is permitted to lie within a flexible range according to a trade-off between soft product quality and cost. Set s_j approximately equal to the total cost of a reference design.

The penalty term takes the role of the relative value/benefit in the sense of multi-objective optimization. This problem type has relevance in many environmental problems for assessing optimal levels of solvent recovery (e.g., Chakraborty and Linninger, 2003).

$$P_j^D = |X_j^D - x_j^D| \cdot s_j \quad (14)$$

Implementation

Figure 7 illustrates the information flow of the hybrid GA with temperature collocation. It receives an initial population of candidate solutions as input. The minimum bubble point distance function and the cost are computed for each starting individual. In each cycle of the master GA, wheel-of-fortune events choose pairs of individuals for reproduction. The likelihood for an individual's selection is proportional to the fitness, which is the inverse of cost as expressed in Eq. 12. Hence, fitter candidate solutions have a greater chance of reproducing. Two offsprings are created using the crossover methodology described in Eqs. 8 and 9. Another random event decides on mutation of the newly produced offspring according to Eqs. 10 and 11. For each new candidate solution, the fitness, including cost and the *BPD*, are evaluated with the help of the temperature collocation algorithm and stored according to Eq. 12. This cycle is repeated until the specified number of generations is reached. The final population typically contains clusters of multiple structures, each one corresponding to meaningful locally optimal solutions. The best performer in the final population or the best performer ever encountered are "declared" the optimal GA solution. This final result, however, is not globally optimal in a rigorous mathematical sense. Global optimization can only be ascertained by analytical proof or exhaustive enumeration, which is intractable for column sequencing problems.

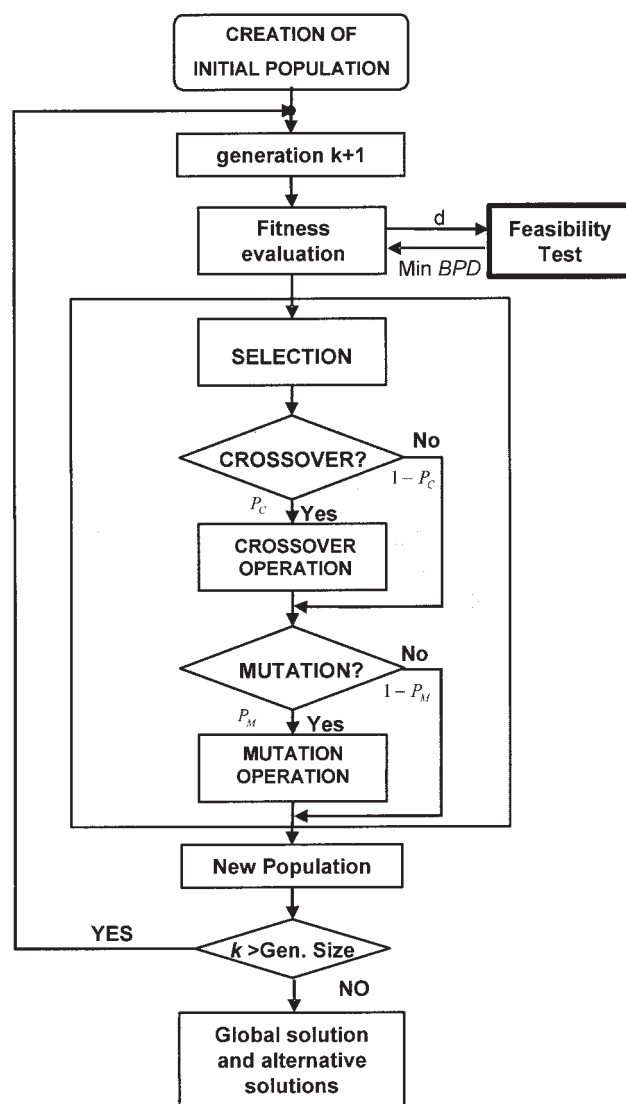


Figure 7. Information flow diagram of column design algorithm.

The large scale column profile computations for each design d are delegated to the feasibility test, returning the minimum bubble point distance (min *BPD*).

Applications

This section demonstrates the application of the new algorithm to challenging separation synthesis problems. Three problem types of practical significance will be presented: (i) optimal column sequences for up to five species involving a four-column network, (ii) complete solution maps to classes of separation synthesis problems, and (iii) breaking of azeotropes.

Optimal sequences of multi-component separations

Optimal Ternary Column Sequences. The separation of an ideal ternary mixture elucidates the convergence dynamics of the proposed design procedure. The task consists of finding the minimum total annualized cost (TAC) distillation sequence for separating an ideal mixture of three alkanes: pentane, hexane, and

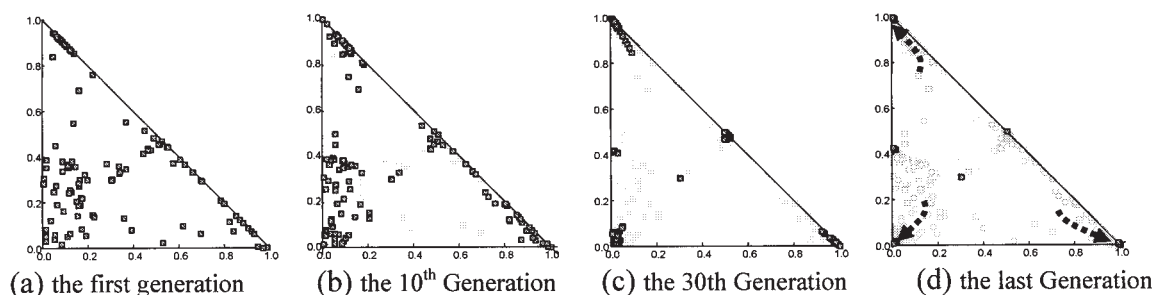


Figure 8. Evolution of candidate solutions to a ternary column synthesis problem in the composition space.

(a) The first generation, (b) the 10th generation, (c) the 30th generation, and (d) the last generation.

heptane. The product purity of the recovered species needs to be above 99%. The evolution of the GA optimization is depicted in four snapshots of consecutive populations in Figure 8. Candidate solutions in successive generations migrate from initially random product specifications towards the product corners. The members of the final generation are located within the desired purity target. At the same time, the GA search makes the column specifications feasible by appropriate variable adjustments. An additional feature of the GA is its ability to find several clusters of locally optimal solutions. The final population depicted in Figure 8d contains two optimal solution clusters.

Table 3 documents the ten best direct sequences, with the minimum cost of around 2.5×10^3 ; a cluster of more expensive indirect sequences is also listed. Figure 9 plots the actual tray composition profiles of the best direct sequence with the optimal trade-off between capital and operating cost. The intersection of the column profiles confirms that the columns are realizable. For additional validation of the computational results, we solved the same problem again but considered only operating cost in the objective. Accordingly, the optimal solution should encompass two columns operating at minimum reflux. Figure 9 shows the GA solution, with an intersection at the stripping pinch confirming the

minimum reflux operation, as expected. It is a remarkable property for a numerical approach to locate optimal solutions at the feasibility boundary robustly.

Quaternary and Five Component Mixtures. The proposed method extends seamlessly to multi-component separation problems. The algorithm was tested for the optimal separation of an ideal quaternary mixture into almost pure components. Figure 10a identifies the best configuration as a direct split sequence ([A][B][C][D]), with a total cost (TAC) of 5.98×10^3 ; Figure 10b displays the stripping and rectifying profiles of the first column in the optimal three-step separation network. This quaternary column implements a sloppy split because its liquid equilibrium tray composition profile does not approach the saddle pinch in the stripping section. Sloppy split solutions like this cannot be computed with existing geometric approaches. Table 4 lists the performance of four solution clusters corresponding to different structural solutions to the quaternary separation task, each one leading to distinct locally optimal separation sequences.

A five-component purification problem was solved to further demonstrate the algorithm's performance and robustness. Table 5 lists the species composition for a five-component mix-

Table 3. The Best Direct and Indirect Sequences for Separating a Ternary Mixture of Pentane, Hexane, and Heptane

No.	Chromosome								Structure	Cost ($\times 10^3$ \$)
	First Column				Second Column					
	X_A^D	X_B^D	X_A^B	R	X_A^D	X_B^D	X_A^B	R		
1	0.996800	0.003170	0.000726	1.088155	0.001404	0.986607	0.000213	1.322701	[2]	2.561
2	0.996816	0.003168	0.000729	1.323342	0.001425	0.986676	0.000211	1.300841	[2]	2.565
3	0.996830	0.003161	0.000730	1.310166	0.001346	0.986861	0.000265	1.310006	[2]	2.567
4	0.996815	0.003165	0.000727	1.321115	0.001362	0.987767	0.000248	1.310006	[2]	2.569
5	0.996834	0.003160	0.000730	1.312222	0.001364	0.986950	0.000251	1.165524	[2]	2.570
6	0.996811	0.003168	0.000728	1.323342	0.001425	0.986676	0.000200	1.300841	[2]	2.571
7	0.996814	0.003164	0.000727	1.319203	0.001393	0.986721	0.000222	1.332804	[2]	2.575
8	0.996724	0.003199	0.000723	1.305573	0.001382	0.987255	0.000225	1.307892	[2]	2.582
9	0.996836	0.003151	0.000724	1.327461	0.001413	0.987018	0.000203	1.337295	[2]	2.601
10	0.996750	0.003189	0.000725	1.307163	0.001409	0.986471	0.000206	1.316884	[2]	2.613
32	0.511036	0.488236	0.000622	1.01512	0.995393	0.003174	0.009239	2.273271	[1]	3.288
33	0.511082	0.488267	0.000618	1.01532	0.995423	0.003174	0.009901	2.285610	[1]	3.301
34	0.511081	0.488226	0.000636	1.01865	0.995434	0.003174	0.009106	2.492631	[1]	3.511
35	0.511055	0.488286	0.000610	1.01422	0.995384	0.003174	0.009754	2.527567	[1]	3.542
36	0.511083	0.488223	0.000619	1.01864	0.995429	0.003174	0.009082	2.531756	[1]	3.550
37	0.511052	0.488278	0.000672	1.01452	0.995423	0.003174	0.009384	2.537592	[1]	3.552
38	0.511090	0.488287	0.000632	1.01785	0.995401	0.003174	0.009783	2.536575	[1]	3.554
39	0.511071	0.488235	0.000611	1.01952	0.995450	0.003174	0.009942	2.579347	[1]	3.599
40	0.511064	0.488292	0.000610	1.01382	0.995710	0.003174	0.009643	2.592762	[1]	3.607
41	0.511048	0.488251	0.000613	1.09361	0.995476	0.003174	0.009817	2.534857	[1]	3.649

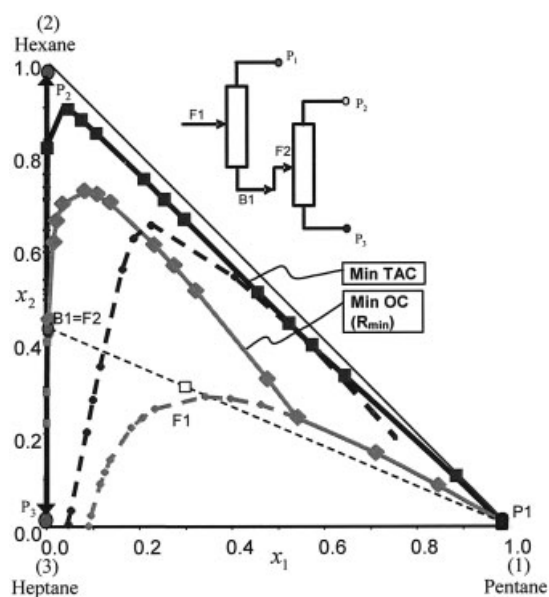


Figure 9. Comparison of two optimal strategies for solving a ternary column sequencing problem.

The composition profiles of the sequence leading to the minimum total annualized cost (that is, min TAC— \blacksquare) optimally balance the tray numbers with the refluxes, that is, capital versus operating cost. The minimum operating cost solution (min OC— \blacklozenge) operates at minimum reflux (and infinite number of trays), as expected.

ture of alcohols with almost constant relative volatilities (Westberg and Wahnschaft, 1996). According to Figure 11, there are 14 structurally different configurations to separate the feed into pure products using simple column configurations only (i.e., columns with one feed, one distillate, and one bottom

Table 4. The Minimum Cost Solutions for Different Distillation Column Configurations to Separate a Quaternary Mixture

Configuration	Column Product Sequence	Minimum Cost (\$)
	[A][B][C][D]	5.98×10^3
	[A][D][B][C]	6.52×10^3
	([A][B])([C][D]) [A][B][C][D]	6.61×10^3
	[D][A][B][C]	7.72×10^3
	[D][C][A][B]	8.40×10^3

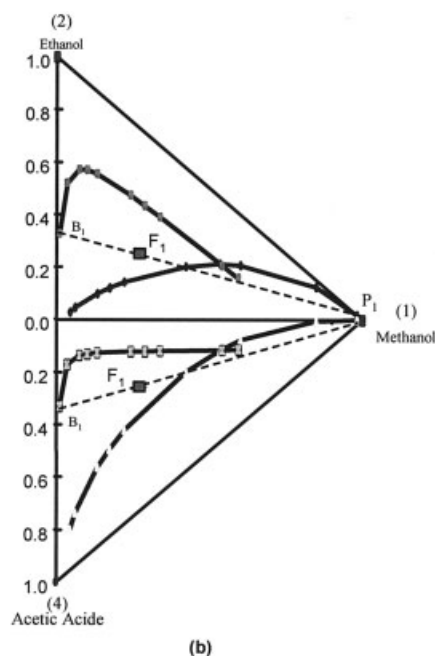
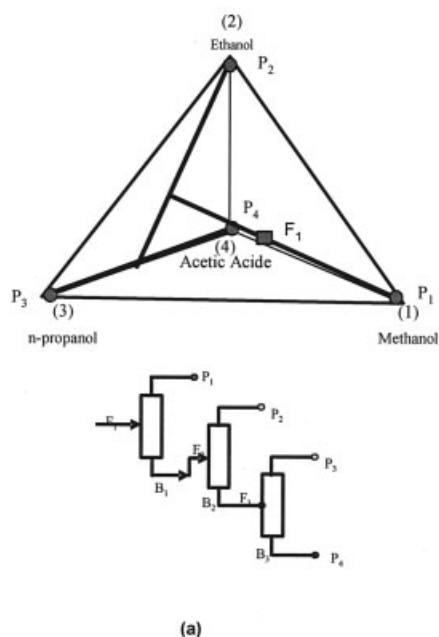


Figure 10. Configuration and column profiles of the first column in a quaternary separation train.

Table 5. Composition and Relative Volatility of a Five Species Alcohol Mixture

	Species	Composition (mol %)	Relative Volatility
A	Isobutanol	5	10.807
B	1-pentanol	10	5.773
C	1-hexanol	20	3.255
D	1-heptanol	50	1.809
E	1-octanol	15	1.000

Figure 12 depicts the liquid composition profiles of the four columns in the optimal column sequence in the bubble point temperature space. The first column, C_1 , recovers highly pure iso-butanol in the distillate, P_1 , from the five-component feed, F_1 . The bottom, B_1 , serves as the feed of the second column, C_2 , which produces two intermediate binary products to be purified in columns C_3 and C_4 . It is important to notice the methodology's ability to solve this complex separation problem with the smallest possible design input. Apart from product purity requirements, the novel procedure takes all choices relating to the intermediate product nodes as well as the resolution of internal degrees of freedom automatically and optimally without any informed initialization or problem-specific knowledge about the solution.

For validation purposes, the problem was solved again with total vapor rate as the objective (that is, minimum operating cost sequence). The structure of the optimal sequence is the same as reported by Westerberg and Wahnschaft (1996). However, the required minimum total vapor rate in the rigorously computed sequence was much higher than the value reported by the marginal cost method, which depends on strong simplifying assumptions. Moreover, the marginal cost method does not extend to general objectives nor to complex vapor liquid models.

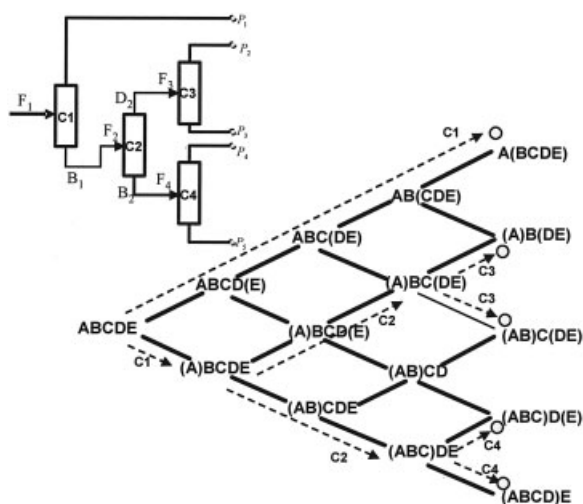


Figure 11. Network representation of a five species purification problem.

The product sequence of the structural solution (above) is depicted by the dashed path in the network (below).

Table 6. Column Height and Reflux Ratio of the Optimal Solution for the Separation of a Five Component Alcohol Mixture

	Height of Column (m)	Reflux Ratio, r
1	13.1	4.17
2	16.2	1.39
3	21.3	5.16
4	17.9	1.48

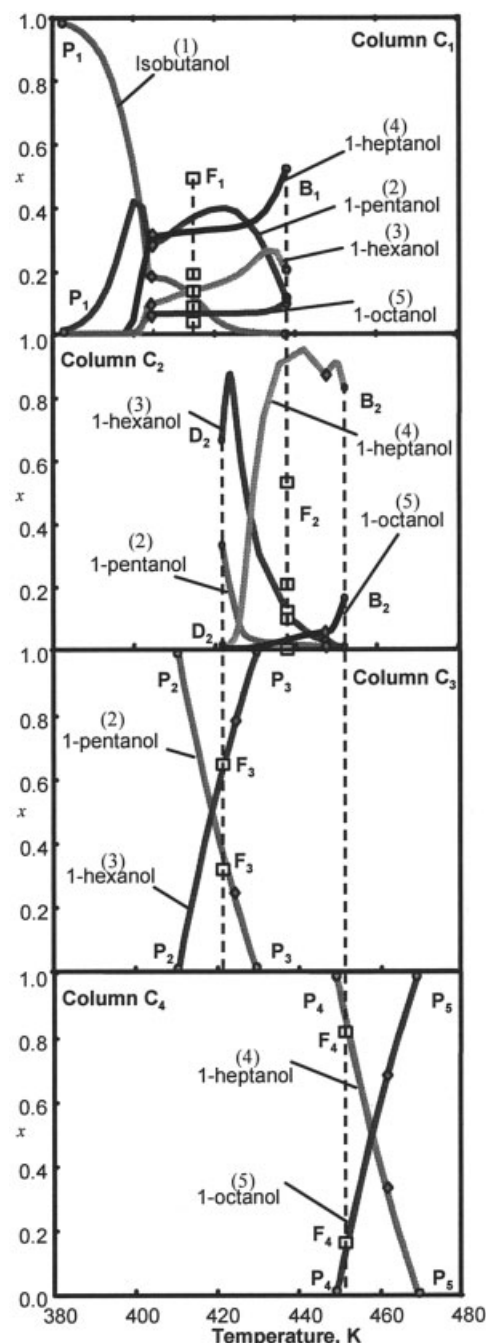


Figure 12. Display of column liquid composition profiles over bubble point temperatures of four-column sequence to optimally solve a five species separation synthesis problem.

Table 7. Operating Parameters and Configuration of the Optimal Solutions to a Ternary Separation Problem with Different Feed Compositions

No.	Feed	Operating Parameter (r_1 and r_2)	Configuration	CPU Time (s)
1	0.3, 0.3, 0.4	1.088, 1.323	[2]	296
2	0.3, 0.4, 0.3	1.294, 2.569	[2]	410
3	0.4, 0.3, 0.3	0.734, 2.661	[2]	410
4	0.4, 0.4, 0.2	1.798, 2.956	[2]	251
5	0.4, 0.2, 0.4	0.707, 2.642	[2]	355
6	0.2, 0.4, 0.4	3.274, 4.367	[2]	356
7	0.2, 0.6, 0.2	4.322, 5.817	[2]	298
8	0.6, 0.2, 0.2	1.999, 3.808	[2]	450
9	0.2, 0.2, 0.6	0.892, 2.713	[2]	364
10	0.3, 0.6, 0.1	4.817, 5.320	[2]	216
11	0.1, 0.8, 0.1	2.363, 2.815	[1]	333

Complete solution maps to classes of sequencing problems

The speed and robustness of the proposed method enables design engineers to address problems previously not amenable to numerical analysis. Consider the problem of finding the best separation network for varying feed compositions in a ternary mixture. For some region of the feeds in the composition triangle, a direct split may be optimal. In another region, the indirect split may be superior, yet close to the pure component side, a single column may be sufficient. It would be desirable to draw the exact boundaries for which a particular configuration is best. A two dimensional map of optimal structural solution spanning the entire composition space would ideally summarize the complete solution spectrum for a particular ternary mixture. We will demonstrate a first computational approach for answering this type of problem, never reported before in the literature.

Separate GA problems for different realization feed compositions were launched sequentially or in parallel. Although not

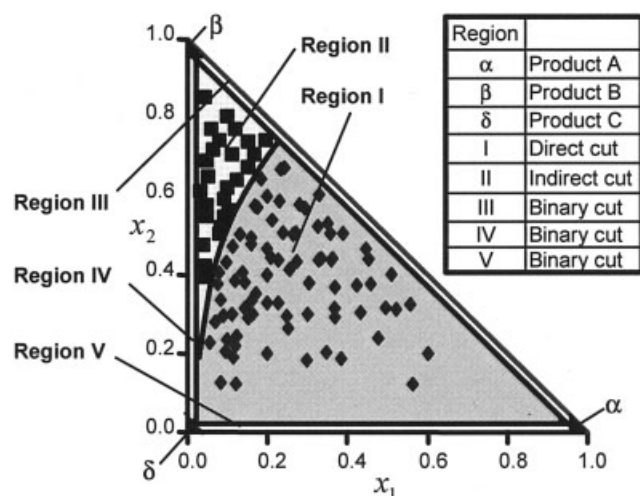


Figure 13. Complete solution map of optimal column sequences for separating a ternary mixture with different feed compositions.

All feeds located in Region I are best separated by a direct sequence.

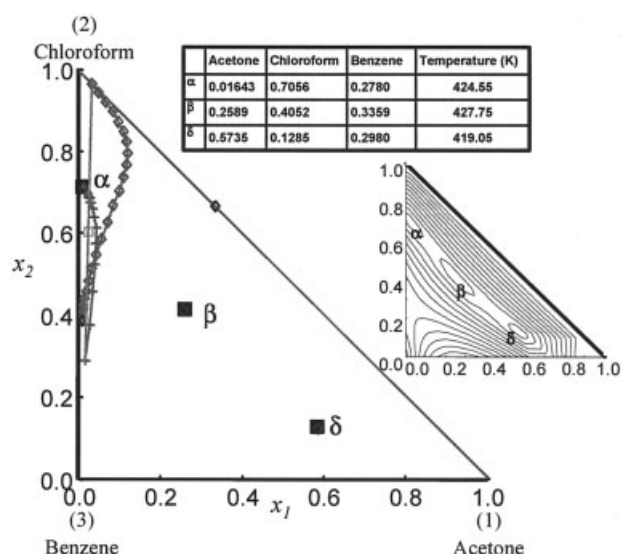


Figure 14. Pinch points of the stripping section for an azeotropic mixture determined by interval analysis.

essential to the principle, substantial time-savings are achievable by running different scenarios on parallel processors. For each sub-problem, the chromosome and the fitness of the optimal separation network were recorded. Table 7 documents the typical performance results for sub-processes associated with different feed compositions. Each sub-process consumed similar CPU time for 50 generations; the embedded temperature collocation algorithm ran robustly without convergence problems. When displaying the results of this massively parallel analysis, a complete solution map delineating the optimal achievable separation emerges as depicted in Figure 13.

The direct sequence is the optimal choice for feed compositions located in region I. Feeds in region II are separated more cheaply indirectly. A clear solution boundary delineates the spaces of optimal direct and indirect solutions. For completeness, feed compositions in regions III, IV, and V requiring but one binary column, and the product regions α , β , and δ without the need for separations are indicated. The accurate results of the complete solution map for varying feeds agree with simplified heuristics proposed by Malone et al. (1985). The complete solution maps computed with the novel approach can be used to decide on optimal column configurations when feed compositions are changing from day to day, as is the case in batch processes campaigns used in pharmaceutical manufacturing sites (e.g., Chakraborty and Linninger, 2002, 2003).

Breaking azeotropes

Pinch Point Location in Azeotropic and Non-Ideal Mixtures. Although the temperature collocation readily applies to vapor-liquid models of highly non-ideal and azeotropic mixtures, pinch point location requires special attention. Several methods, like homotopy-continuation methods and bifurcation analysis, are effective for the numerical solution of the pinch equations (e.g., Fidkowski et al., 1991). For azeotropic problems of this article, an Interval Newton with generalized bi-

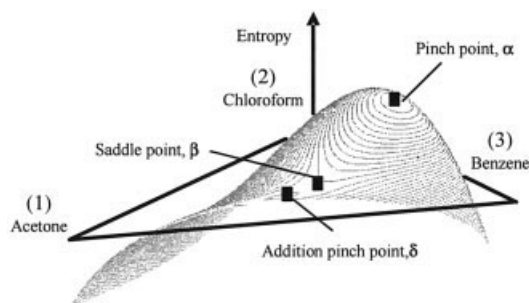


Figure 15. Entropy production surface and location of the pinch points of the stripping profile.

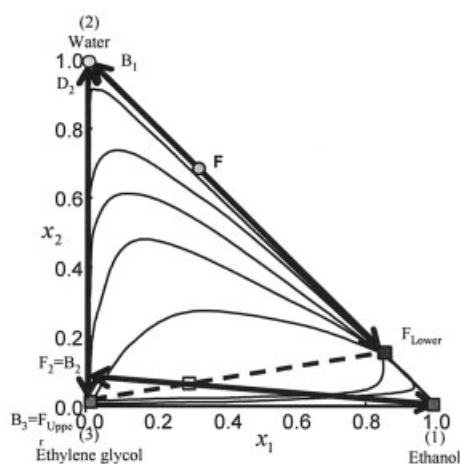
section using interval arithmetic (Kearfott, 1996) was implemented. This technique generates interval enclosures containing solutions of nonlinear equations. The sequence of solution boxes is searched exhaustively until all solutions are identified. Existence and uniqueness criteria offered by the Interval Newton method guarantee that all the roots are located. Although not encountered in the case studies, interval methods may require substantial computational effort when two bifurcation points are close to each other.

As an example for pinch location, consider a mixture of acetone-chloroform-benzene with a binary high boiling azeotrope. For a given feed, reflux ratio, and product composition, this highly non-ideal system exhibits three stationary nodes: a stable pinch point, α , a saddle point, β , and an additional pinch point, δ , as depicted in Figure 14. The residual error contour map of the stripping pinch equations exhibits three minima, corresponding to three solutions of the pinch equations. The flat contours around the solution explain why it is hard for conventional numerical techniques to identify all roots. The Interval Newton approach proved effective in locating all the pinch solutions for the azeotropic mixtures. Once all solutions of the pinch equations are known, the problem of categorizing the stationary nodes arises. It has been shown that entropy and

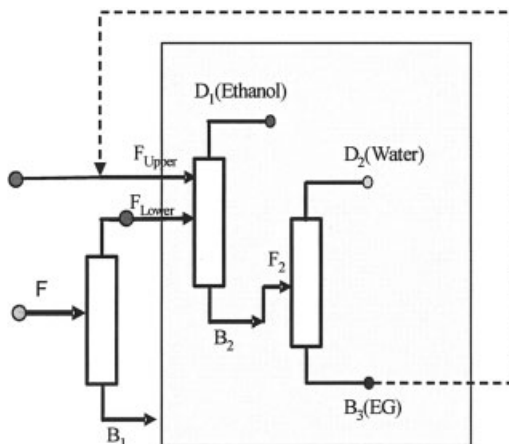
eigenvalues of Jacobian corresponding to the pinch equations allow a proper classification of the stationary points into unstable, stable, and saddle nodes (Bausa, et al., 1998; Julka and Doherty, 1993; Levy and Doherty, 1986a). The maximum of the entropy surface allowed us to detect node, α , as the stable node of the stripping profile, as shown in Figure 15. With the exception of the extra attention to locate the pinch points, temperature collocation proceeds identical to the ideal mixture case.

Classical Azeotropic Separation. Azeotropic mixtures are broken by extractive or homogeneous azeotropic distillation. In either process, a liquid separating agent, also known as an entrainer, is added to facilitate the separation (Doherty and Malone, 2001; Widago and Seider, 1996). For homogeneous azeotropic distillation, the separation agent is the heaviest species in the system. It does not form any azeotropes with the original components and is completely miscible with both in all proportions. Extractive distillation requires a double-feed column (Knapp and Doherty, 1994; Levy and Doherty, 1986b). The temperature collocation approach was extended to multiple-feed columns into which the entrainer-based separation introduces an additional middle column section. The new middle profile traverses the stripping profile and the pinch point of the middle section. The *BPD*, guaranteeing the feasibility of the two-feed column, is defined between the middle section (instead of the stripping profile) and the rectifying profile.

The extension of the multiple-feed column was applied for the optimal sequencing of a classical azeotropic separation of an ethanol-water mixture, with ethylene glycol as the entrainer. Figure 16 shows the residue map of a ternary mixture with one binary azeotrope. The flowsheet of a classical azeotropic mixture depicts the entrainer feed in the rectifying section. The algorithm discovers the classical flowsheet of the multi-section column profiles, with the azeotrope entering the second column and ethylene glycol added in the rectifying section, as depicted in Figure 17. The reflux ratio and the entrainer ratio as additional degrees of freedom were solved for simultaneously.



(a) Residue curve



(b) Schematic flowsheet of homogeneous azeotropic separation

Figure 16. Separation of a classical ternary azeotropic mixture (1. Ethanol; 2. Water; 3. Ethylene glycol).

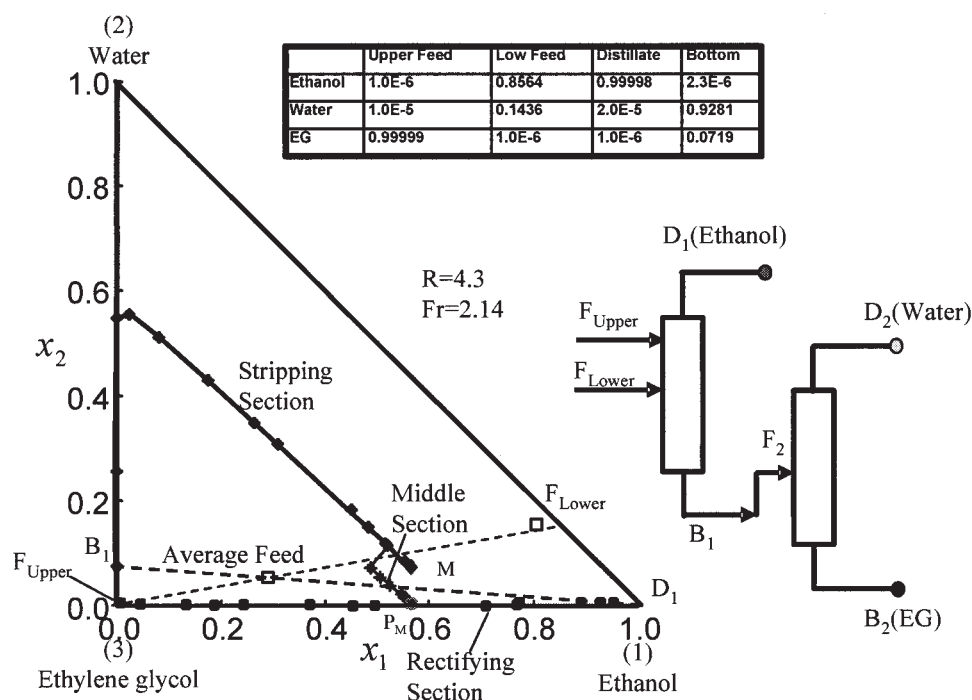


Figure 17. Optimal sequence composition profile of multi-section azeotropic column for the classical entrainer-based separation of a ternary azeotropic mixture (1. Ethanol; 2. Water; 3. Ethylene glycol).

Azeotropic Mixture with Curved Distillation Boundary. The novel approach can also identify column configurations that penetrate distillation boundaries observed by previous authors (e.g., Doherty and Malone, 2001; Fidkowski et al, 1993). The reason for the penetration lies in overlapping distillation regions. The bottom F_2 of Figure 18 is located in an intersecting region that is reachable even from a feed F_1 (Doherty and Malone, 2001). This phenomenon was found to occur in the separation of a chloroform-acetone-benzene mixture

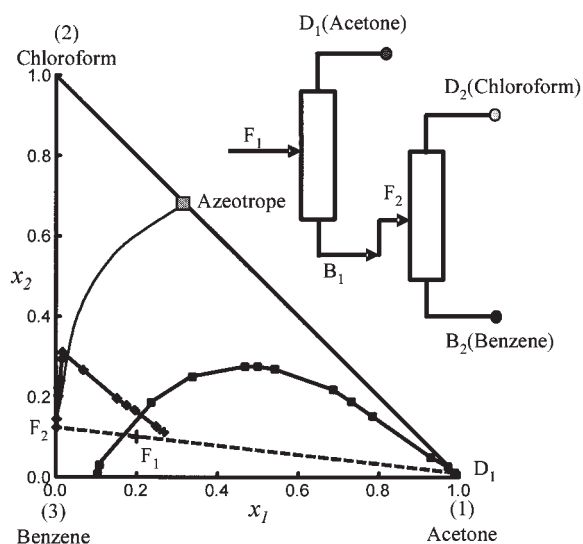


Figure 18. Liquid composition profiles of the first column in the optimal sequence traversing a distillation boundary.

whose residue curve map exhibits two distillation regions with a curved distillation boundary. The problem consisted in identifying the optimal realizable flowsheet for separating a feed mixture into products belonging to different distillation regions. The optimal sequence for the separation of 100 Kmol/h feed (feed composition: chloroform, 0.2; acetone, 0.1; benzene, 0.7) produces 99.1% acetone as distillate and a binary chloroform-benzene mixture as bottoms in the first column. The optimal profile of the first column, whose stripping profile penetrates the curved distillation boundary, is displayed in Figure 18. The vapor rate of the optimal solution is 115.21 Kmol/h. These results show that the well-known penetration of curved distillation boundaries can be correctly computed with the proposed method.

Discussion

Problem initialization

Synthesis problems leading to mixed integer non-linear mathematical programs (MINLP) formulations are very sensitive to good problem initialization. However, consistent column profile initialization for a multi-component separation problem is not a trivial task. We studied the sensitivity of the proposed approach with respect to three kinds of initial populations: patterned initial guesses, random initial guesses, and initialization with infeasible designs. In the ternary problems, the patterned initial guesses happened to produce a few feasible designs; while random initialization contained several feasible columns but none satisfied the product specifications. The random initialization spreads the initial generation throughout the search space. This is desirable, since the location of the optimum solution in the search space is unknown.

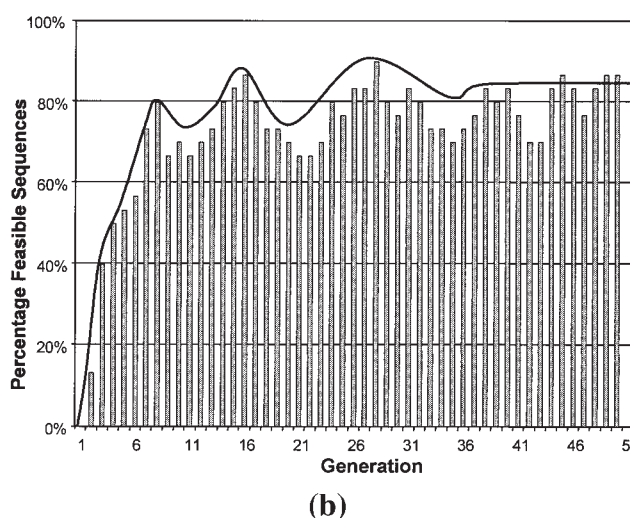
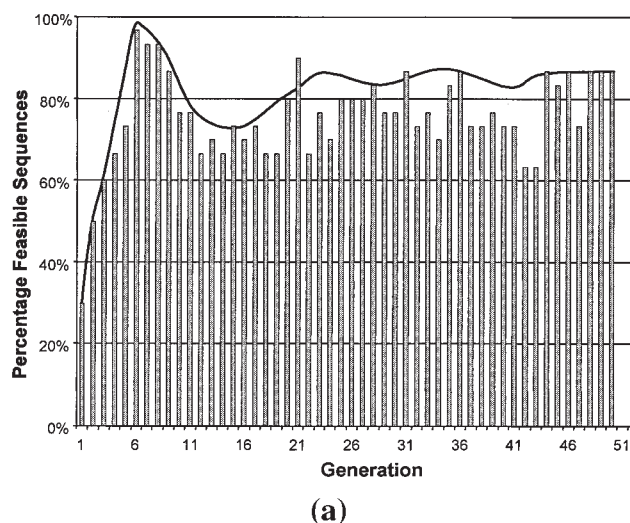


Figure 19. The percentage of feasible column configurations in the overall population as the GA search progresses. Both experiments converged to the same solution after 50 generations.

(a) Evolution of the percentage feasible sequences in the genetic pool starting from an initial population including some feasible designs. In the first iteration, the algorithm constructs feasible designs. In the later stages, the average percentage of feasible designs remains at around 80%.

(b) Evolution of feasible sequences starting from an initial population made up exclusively by infeasible guesses. Even without a single feasible solution in the initial population, the novel method generates feasible candidate solutions in a few iterations and converges to the same result as with better (feasible) initialization.

We also investigated the extreme condition of an initial population assembled entirely of infeasible column sequences. Figure 19a shows the GA evolution in terms of the feasible candidate solutions. The percentage of feasible sequences gradually increases from initially only nine to a stable level of about 80%. The genetic algorithm exploits partially feasible as well as close to feasible traits to construct improved candidate solutions. Figure 19b demonstrates that the genetic algorithm synthesizes feasible sequences converging to the same global solution, even from an initial population without a single feasible column. The demonstrated robustness in the absence of problem-specific initialization is highly relevant in practice, freeing the process engineer from the burden of supplying feasible initial

specifications in large multi-component synthesis problems. We conjecture that the robust global converge of the novel method could be of great value for the initialization of large-scale commercial flowsheet simulators.

Tuning and validation

The convergence of the algorithm is known to depend on the choice of tuning parameters: such as, crossover probability (P_c), mutation probability (P_m), the size of the population (N_p), and the maximum number of generations (N_g). Higher crossover probability P_c explores more solutions, reducing the chance of missing the global optimum. Low mutation probability P_m may miss globally optimal solu-

tions, while too high a value of P_m may destroy competitive traits. Similarly, large population and/or generation count intensifies the search. However, large population or generation counters augment the computation effort, often without improvement in the final result. In this article, typical values for P_c , P_m , N_p , and N_g were 0.75, 0.08, 30, and 51, respectively. Although there is no formal guarantee even when tuning parameters are chosen properly, the solutions presented here can be considered “globally” optimal, at least in a statistical sense.

The final “globally” optimal GA solutions were submitted to a local gradient-based optimizer (i.e., rSQP-NLP (Biegler, 1984; Cervantes and Biegler, 1998; Lucia, 2001; Tanartkit and Biegler, 1995)) that solved the sequencing problem, including full implementation of the feasibility criterion, with the help of finite element collocation of the column profile equations (Zhang and Linninger, 2004). For the reported case studies, the “optimization” of the GA solution did not produce significant changes, thus affirming at least local optimality of the simulation result.

Conclusions and Significance

The direct solution of separation synthesis problems with mathematical programming techniques is numerically challenging due to its size, discontinuities in the search space, and numerical instability outside the thermodynamic definition space. The problem size stems from the combinatorial permutations of possible column configurations and the stage numbers. In addition, each distillation train possesses a huge amount of internal state variables, such as vapor and liquid compositions corresponding to tray-by-tray equilibrium and operating line equations. The combination of structural decisions and continuous operating parameters leads to a discontinuous design space with a multitude of local minima. Finally, numerical instabilities occur frequently when column profiles are initialized outside thermodynamically consistent ranges. However, consistent initialization of feasible column profiles for multi-component separations is a challenging task in its own right.

The proposed computational design framework combines the advantages of stochastic search with a massively reduced problem representation. Size reductions of two orders in magnitude are due to two characteristics of the temperature collocation transformation: (i) mapping of discrete tray numbers into a dimensionless bounded bubble point temperature, and (ii) continuous representation of profile equations by polynomial collocation on finite elements.

The novel transformed problem converges robustly for feasible or infeasible design specifications; divergence due to improper initialization of infeasible specifications does not occur. It was also demonstrated that the algorithm’s ability to capitalize on partially infeasible or entirely unrealizable candidate solutions is necessary for discovering optimal solutions located at the boundaries of feasibility, such as minimum reflux column designs.

Temperature collocation also made it possible to cast the information pertaining to the internal composition profiles and operating conditions into a single scalar metric, i.e., bubble point distance (BPD). In effect, the optimal sequenc-

ing problem could be transformed into an unconstrained single-objective mathematical program readily solvable with a specialized genetic algorithm. The problem-specific GA uses the smallest possible design variable set, generating optimal solutions to complex separation problems without the need for proper initialization, a-priori problem-specific knowledge, or gradient information.

In the Applications section, computer-aided design solutions to challenging optimal design problems were presented, using only final product targets as input. In all cases, the method solved robustly at reasonable computational cost. These results might be useful for the robust initialization of large-scale flowsheet simulators. The performance of the novel computer-aided design approach allows the generation of complete solution maps for classes of optimal sequencing problems with varying feed compositions, exploiting the massively parallel problem implementation. This type of procedure is apt to generate high-level knowledge previously limited to the careful experts’ analyses.

In the future, it is planned to extend the new methodology to complex column configurations with variable numbers of columns, perhaps allowing for internal recycles. Another important subject is the inclusion of uncertainty in optimal separation network synthesis, hardly studied yet in the open literature.

Acknowledgments

Partial support for this research by NSF Grant DMI-0328134 is gratefully acknowledged. The authors thank Professor Lorenz T. Biegler from Carnegie Mellon University for providing the source code of his rSQP algorithm.

Notation

b	= exponent mutation coefficient
C	= number of species
d	= design variables, such as product purity, reflux ratio, etc.
e	= structure parameter
e'	= structure parameter
F_r	= entrainer ratio, $F_{\text{Upper}}/F_{\text{Lower}}$
K_i	= equilibrium constant of component i
h	= column height, m
m_i^j	= molar flow rate of component i of product j , Kmol/hr
N	= number of columns
N_g	= maximum number of generations
N_p	= size of population
P_c	= crossover probability
P_m	= mutation probability
P_i^f	= feasibility penalty of column i
P_j^b	= purity penalty for product j
r	= reflux ratio
s	= reboil ratio
t	= index of current generation
T	= bubble point temperature for an equilibrium tray, K
T_D, T_B	= bubble point temperature corresponding to distillate and bottoms compositions, K
x_i	= liquid mole fraction of component i
y_i	= vapor mole fraction of component i

Greek letters

α	= random number $\alpha \in [0, 1]$
ϕ	= chromosome trait of parent
ϕ'	= chromosome trait of child
ϕ^U	= upper bound for chromosome trait
λ, μ	= operating and capital cost coefficients

Subscripts

B = bottoms product
 c = component index
 D = distillation product
 i, j = species and product indices

Superscripts

R = rectifying section
 S = stripping section

Literature Cited

- Aggarwal, A., and C. A. Floudas, "Synthesis of General Distillation Sequence—Nonsharp Separation," *Comp. Chem. Eng.*, **14**, 631 (1990).
- Aggarwal, A., and C. A. Floudas, "Synthesis of Heat Integrated Non-Sharp Distillation Sequences," *Comp. Chem. Eng.*, **16**, 89 (1992).
- Agrawal, R., "Synthesis of Multi-Component Distillation Column Configurations," *AIChE J.*, **49**, 379 (2003).
- Andreacovich, M. J., and A. W. Westerberg, "An MILP Formulation for Heat-Integrated Distillation Sequence Synthesis," *AIChE J.*, **31**, 1461 (1985).
- Barnicki, S. D., and J. J. Siirola, "Process Synthesis Prospective," *Comp. Chem. Eng.*, **28**, 441 (2004).
- Bauer, M. H., and J. Stichlmair, "Synthesis and Optimization of Distillation Sequences for the Separation of Azeotropic Mixtures," *Comp. Chem. Eng.*, **19**, S15 (1995).
- Bauer, M. H., and J. Stichlmair, "Superstructure for the Mixed Integer Optimization of Nonideal and Azeotropic Distillation Process," *Comp. Chem. Eng.*, **20**, S25 (1996).
- Bauer, M. H., and J. Stichlmair, "Design and Economic Optimization of Azeotropic Distillation Processes Using Mixed-Integer Nonlinear Programming," *Comp. Chem. Eng.*, **22**, 1271 (1998).
- Bausa, J., R. V. Watzdorf, and W. Marquardt, "Shortcut Methods for Nonideal Multicomponent Distillation: 1. Simple Columns," *AIChE J.*, **44**, 2181 (1998).
- Biegler, L. T., "Solution of Dynamic Optimization Problems by Successive Quadratic Programming and Orthogonal Collocation," *Comp. Chem. Eng.*, **8**, 243 (1984).
- Cervantes, A., and L. T. Biegler, "Large Scale DAE Optimization Using Simultaneous Nonlinear Programming Formulations," *AIChE J.*, **44**, 1038 (1998).
- Chakraborty, A., and A. A. Linninger, "Plant-Wide Waste Management. 1. Synthesis and Multi-Objective Design," *Ind. Eng. Chem. Res.*, **41**, 4591 (2002).
- Chakraborty, A., and A. A. Linninger, "Plant-Wide Waste Management. 2. Decision Making Under Uncertainty," *Ind. Eng. Chem. Res.*, **42**, 357 (2003).
- Costa, L., and P. Oliveira, "Evolutionary Algorithms Approach to the Solution of Mixed Integer Non-Linear Programming Problems," *Comp. Chem. Eng.*, **25**, 257 (2001).
- Doherty, M. F., and M. F. Malone, *Conceptual Design of Distillation Systems*, McGraw-Hill, New York (2001).
- Douglas, J. M., *Conceptual Design of Chemical Processes*, McGraw-Hill, Boston (1988).
- Dunnebie, G., and C. C. Pantelides, "Optimal Design of Thermally Coupled Distillation Columns," *Ind. Chem. Eng. Res.*, **38**, 162 (1999).
- Fidkowski, Z. T., M. F. Malone, and M. F. Doherty, "Nonideal Multicomponent Distillation: Use of Bifurcation Theory for Design," *AIChE J.*, **37**, 1761 (1991).
- Fidkowski, Z. T., M. F. Malone, and M. F. Doherty, "Feasibility of Separations for Distillation of Nonideal Ternary Mixtures," *AIChE J.*, **39**, 1303 (1993).
- Floudas, C. A., and G. E. Paules IV, "A Mixed-Integer Nonlinear Programming Formulation for the Synthesis of Heat-Integrated Distillation Sequences," *Comp. Chem. Eng.*, **12**, 531 (1988).
- Frey, Th., M. H. Bauer, and J. Stichlmair, "MINLP-Optimization of Complex Column Configurations for Azeotropic Mixtures," *Comp. Chem. Eng.*, **21**, S217 (1997).
- Gen, M., and R. Cheng, *Genetic Algorithms and Engineering Optimization*, Wiley, New York (2000).
- Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, 2nd Edition, Addison-Wesley, New York (1989).
- Gopal, V., and L. T. Biegler, "Smoothing Methods for Complementarity Problems in Process Engineering," *AIChE J.*, **45**, 1535 (1999).
- Hauan, S., A. W. Westerberg, and M. L. Kristian, "Phenomena-Based Analysis of Fixed Points in Reactive Separation Systems," *Chem. Eng. Sci.*, **56**, 1053 (1999).
- Ismail, S. R., E. N. Pistikopoulos, and K. P. Papalexandri, "Modular Representation Synthesis Framework for Homogeneous Azeotropic Separation," *AIChE J.*, **45**, 1701 (1999).
- Ismail, S. R., P. Proios, and E. N. Pistikopoulos, "Modular Synthesis Framework for Combined Separation/Reaction Systems," *AIChE J.*, **47**, 629 (2001).
- Julka, V., and M. F. Doherty, "Geometric Behavior and Minimum Flows for Nonideal Multicomponent Distillation," *Chem. Eng. Sci.*, **45**, 1801 (1990).
- Julka, V., and M. F. Doherty, "Geometric Nonlinear Analysis of Multicomponent Nonideal Distillation: Simple Computer-Aided Design Procedure," *Chem. Eng. Sci.*, **48**, 1367 (1993).
- Katere, S., J. M. Caruthers, W. N. Delgass, and V. Venkatasubramanian, "An Intelligent System for Reaction Kinetic Modeling and Catalyst Design," *Ind. Eng. Chem. Res.*, **43**, 3484 (2004).
- Kearfott, R. B., *Rigorous Global Search: Continuous Problems*, Kluwer Academic Publishers, Dordrecht (1996).
- Knapp, J. P., and M. F. Doherty, "Minimum Entrainer Flows for Extractive Distillation: A Bifurcation Theoretic Approach," *AIChE J.*, **40**, 243 (1994).
- Lang, Y., and L. T. Biegler, "Distributed Stream Method for Tray Optimization," *AIChE J.*, **48**, 582 (2002).
- Leboreiro, J., and J. Acevedo, "Processes Synthesis and Design of Distillation Sequences Using Modular Simulators: A Genetic Algorithm Framework," *Comp. Chem. Eng.*, **28**, 1223 (2004).
- Levy, S. G., and M. F. Doherty, "A Simple Exact Method for Calculating Tangent Pinch Points in Multicomponent Nonideal Mixtures by Bifurcation Theory," *Chem. Eng. Sci.*, **41**, 3155 (1986a).
- Levy, S. G., and M. F. Doherty, "Design and Synthesis of Homogenous Azeotropic Distillations. 4. Minimum Reflux Calculations for Multiple-Feed Columns," *Ind. Eng. Chem. Fundam.*, **25**, 269 (1986b).
- Lewin, D. R., "A Generalized Method for HEN Synthesis Using Stochastic Optimization: II. The Synthesis of Cost-Optimal Networks," *Comp. Chem. Eng.*, **22**, 1387 (1998).
- Lewin, D. R., H. Wang, and O. Shalev, "A Generalized Method for HEN Synthesis Using Stochastic Optimization: I. General Framework and MER Optimal Synthesis," *Comp. Chem. Eng.*, **22**, 1503 (1998).
- Lucia, A., "Successive Quadratic Programming: Application in Distillation Systems," *The Encyclopedia of Optimization*, Vol. 5, C. A. Floudas and P. M. Pardalos, eds., Kluwer Acad. Publishers, Dordrecht, p. 393 (2001).
- Malone, M. F., and M. F. Doherty, "Reactive Distillation," *Ind. Eng. Chem. Eng.*, **39**, 3953 (2000).
- Malone, M. F., K. Glinos, F. E. Marquez, and J. M. Douglas, "Simple Analytical Criteria for the Sequencing of Distillation Columns," *AIChE J.*, **31**, 683 (1985).
- Michalewicz, Z., *Genetic Algorithms + Data Structures = Evolutionary Programs*, Springer-Verlag, Berlin (1992).
- Modi, A. K., and A. Westerberg, "Distillation Column Sequencing Using Marginal Price," *Ind. Eng. Chem. Eng.*, **31**, 839 (1992).
- Press, W. H., S. A. Teukolsky, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, New York (1995).
- Sargent, R. W. H., and K. Gaminbandara, "Introduction: Approaches to Chemical Process Synthesis," *Optimization in Action*, L. C. W. Dixon, ed., Academic Press, London (1976).
- Shah, P. B., and A. C. Kokossis, "New Synthesis Framework for the Optimization of Complex Distillation Systems," *AIChE J.*, **48**, 527 (2002).
- Siirola, J., "Industrial Application of Chemical Process Synthesis," *Advances in Chemical Engineering*, J. Anderson, ed., Academic Press, New York, p. 2 (1996).
- Smith, E. M., and C. C. Pantelides, "Design of Reaction/Separation Networks Using Detailed Models," *Comp. Chem. Eng.*, **19**, S83 (1995).

- Tanartkit, P., and L. T. Biegler, "Stable Decomposition for Dynamic Optimization," *Ind. Chem. Eng. Res.*, **34**, 1253 (1995).
- Villadsen, J. V., and M. L. Michelsen, *Solution of Differential Equations by Polynomial Approximation*, Prentice-Hall, New York (1978).
- Viswanathan, J., and I. E. Grossmann, "Optimal Feed Locations and Number of Trays for Distillation Columns with Multiple Feeds," *Ind. Chem. Eng. Res.*, **32**, 2942 (1993).
- Wahnschafft, O., J. Koehler, E. Blass, and A. Westerberg, "The Product Composition Regions of Single-Feed Azeotropic Distillation Columns," *Ind. Chem. Eng. Res.*, **31**, 2345 (1992).
- Wang, K., Y. Qian, Y. Yuan, and P. Yao, "Synthesis and Optimization of Heat Integrated Distillation Systems Using an Improved Genetic Algorithm," *Comp. Chem. Eng.*, **23**, 125 (1998).
- Watzdorf, R. V., J. Bausa, and W. Marquardt, "Shortcut Methods for Nonideal Multicomponent Distillation: 2. Complex Columns," *AIChE J.*, **45**, 1615 (1999).
- Westerberg, A., "Retrospective on Design and Process Synthesis," *Comp. Chem. Eng.*, **28**, 447 (2004).
- Westerberg, A. A., J. Lee, and S. Hauan, "Synthesis of Distillation-Based Processes for Non-Ideal Mixtures," *Comp. Chem. Eng.*, **24**, 2043 (2000).
- Westerberg, A., and O. Wahnschafft, "Synthesis of Distillation Based Separation Systems," *Advances in Chemical Engineering*, J. L. Anderson, ed., Academic Press, New York (1996).
- Widago, S., and W. Seider, "Azeotropic Distillation," *AIChE J.*, **42**, 96 (1996).
- Zhang, L., and A. A. Linninger, "Temperature Collocation Algorithm for Fast and Robust Distillation Design," *Ind. Chem. Eng. Res.*, **43**, 3163 (2004).

Appendix: Conversion of Bubble Point Temperature Range to Column Height

An expression for relating the bubble point temperature to column height can be derived as follows.

Equation A-1 expands the composition x_i as a function of height and bubble point temperature:

$$\frac{dx_i}{dh} = \frac{\partial x_i}{\partial T} \frac{\partial T}{\partial h} \quad (\text{A1})$$

Using Eqs. 2 and 3 to replace the term $\partial x/\partial T$ and eliminating terms $\partial x/\partial h$ on both sides of Eq. A-1, we obtain Eqs. A-2 and A-3, which relate the column height to the bubble point temperature. This approximation to the column height was used to estimate capital cost.

$$\frac{\partial h^R}{\partial T} = \frac{\sum_{j=1}^c \left(\frac{\partial K_j}{\partial T} x_j^R \right)}{\sum_{j=1}^c \left[\left(x_j^R - \frac{r+1}{r} y_j^R + \frac{1}{r} x_{D,j} \right) K_j \right]} \quad (\text{A2})$$

$$\frac{\partial h^S}{\partial T} = \frac{\sum_{j=1}^c \left(\frac{\partial K_j}{\partial T} x_j^S \right)}{\sum_{j=1}^c \left[\left(\frac{s}{s+1} y_j^S - x_j^S + \frac{1}{s+1} x_{B,j} \right) K_j \right]} \quad (\text{A3})$$

Manuscript received Dec. 31, 2004; revision received May 20, 2005; and final revision received Aug. 19, 2005.