AIChE

Design and Optimization of a Hybrid Distillation/Melt Crystallization Process

Meik Bernhard Franke
Dept. of Process Design, Bayer Technology Services GmbH, D-51368 Leverkusen, Germany

Norman Nowotny
Chair of Environmental Technology, Technical University of Dortmund, D-44227 Dortmund, Germany

Eugene Ndocko Ndocko
Dept. of Processing Technology, Bayer Technology Services GmbH, D-51368 Leverkusen, Germany

Andrzej Górak
Chair of Fluid Separation Processes, Technical University of Dortmund, D-44227 Dortmund, Germany

Jochen Strube
Institute for Separation and Process Technology, Technical University of Clausthal, D-38678 Clausthal-Zellerfeld, Germany

DOI 10.1002/aic.11605
Published online October 6, 2008 in Wiley InterScience (www.interscience.wiley.com).

Hybrid distillation/melt crystallization processes are widely used for the separation of isomer mixtures. The design of a hybrid separation processes is not a trivial task because of several structural and operational degrees of freedom. In this article, a new three-step design approach is proposed to address the above mentioned challenges. In the first step, process alternatives are generated by heuristic rules. In the second step, the generated process alternatives are optimized with respect to their energy consumption by using shortcut methods. In the third step, the most promising alternatives are rigorously optimized with respect to their total annualized costs, whereas the resulting mixed-integer nonlinear programming (MINLP) problem is solved using a modified generalized benders decomposition (GBD) algorithm to take the nonconvexities into account. The feasibility of this design approach is demonstrated by design of a hybrid distillation/melt crystallization process for separation of a ternary isomer mixture. © 2008 American Institute of Chemical Engineers *AIChE J,* 54: 2925–2942, 2008
Keywords: crystallization, distillation, optimization, process synthesis, simulation, process

Introduction

The combination of distillation and melt crystallization for separation of close-boiling isomer mixtures is a well-known example for a hybrid separation process. The hybrid distilla-

Correspondence concerning this article should be addressed to M. B. Franke at meikbernhard.franke@bayertechnology.com.

tion/melt crystallization process combines advantages of the distillation, which is the working horse of the chemical process industry and the melt crystallization in which very high separation factors per stage can be reached. Simultaneously, the combination of distillation and crystallization overcomes the shortcomings of the individual unit operations, i.e. high energy-consumption at small separation factors and limitation of yield by eutectics, respectively. Several hybrid processes for the separation of terphenyl,[1] xylene,[2,3] dichlorobenzene,[4,5] and diphenylmethane diisocyanate isomers[6] are reported in literature.

The detailed design of a hybrid separation process is a challenging task because of many degrees of freedom (DOF) involved (see also Figure 1), which can be divided into three classes:

(1) Process structure: type of unit operations included in the flowsheet, number and sequence of unit operations, and the location of recycle streams.

(2) Operating point: compositions of the process streams, so-called *hand-over purities*, which are discussed later in this article. As a general rule, the composition of a component which is separated in the specific unit will be chosen. For example, in Figure 1, the composition of the component P at the top of the distillation column D-1 is chosen to be the hand-over purity.

(3) Unit operation: structural DOF (e.g. number of stages for distillation) and operational DOF (e.g. yield on each stage for melt crystallization).

Whereas the DOF of Class 2 considers only continuous variables, in class one and three, integer as well as continuous variables have to be considered, which make the design problem more difficult. The free variables have to be chosen in such a way that the objective function, usually the total annualized costs, becomes minimal. The more components and unit operations are considered in the design, the number of DOF grows rapidly, which can be expressed by the formula of Thompson and King.[7] The design approaches coping with this problem can roughly be divided into heuristic and mathematical methods.

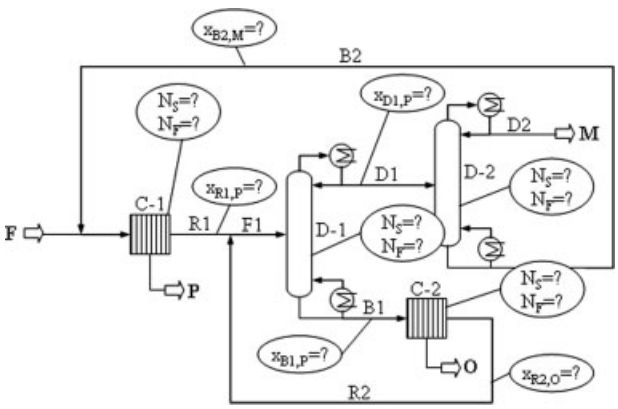

Figure 1. Degrees of freedom of a hybrid separation process.

$x_{R1,P}$, hand-over purity of residue of crystallization C-1 of component P; $x_{D1,P}$, hand-over purity of top product of distillation D-1 of component P; $x_{B1,P}$, hand-over purity of bottom product of distillation D-1 of component P; $x_{B2,M}$, hand-over purity of bottom product of distillation D-2 of component M; $x_{R2,O}$, hand-over purity of residue of crystallization C-2 of component O; N_S, number of stages of distillation or crystallization, respectively; N_F, feed stage of distillation or crystallization, respectively.

Heuristic design methods provide useful information in the early design stage with only few input data. However, well-known heuristics, which exist for distillation,[8] are missing for hybrid separation processes. Barnicki and Fair[9] developed a general knowledge-based design procedure for the separation of liquid mixtures and applied it to an isomer separation example. Because the method does not rely on detailed information, it can only be regarded as a rough guideline for the selection of separation methods. Simmrock et al.[10] describe the construction of the knowledge-based expert system TEAGPERT to generate process alternatives for azeotropic and close-boiling mixtures. Wahnschafft et al.[11] developed the software environment SPLIT for the systematic generation of flowsheets including different unit operations. Sequences developed by heuristic methods can often be improved by evolutionary methods.[12] Heuristic and evolutionary methods may lead to the selection of the nonoptimal sequence, what can be avoided using algorithmic methods.

The branch-and-bound method is a frequently applied algorithmic method.[13] Starting from the feed mixture, different process alternatives are developed by branching and backtracking, whereas every partial process with its costs is represented by a node. The algorithms stops if all nodes are more expensive than the current node and no further node can be visited, because no better solution can be obtained. Unlike branch-and bound, dynamic programming[14,15] proceeds backward from the products and requires that the optimal alternative is chosen whatever the preliminary decision may have been. The advantage of both methods is that not all possible flowsheet alternatives have to be developed and evaluated, so the computational effort is reduced. The disadvantage is that these methods can only be used efficiently when the interactions between the unit operations are small, e.g. if no recycle streams exist. Therefore, the branch-and-bound method and dynamic programming are mainly used for distillation train synthesis with sharp splits.

Bausa and Marquardt[16] used shortcut methods for the fast evaluation of different hybrid membrane/distillation processes on the basis of the minimum reboiler duty and minimum membrane area. Shortcut methods do not require the full specification of all process parameters and provide insight into the thermodynamic nature of the process. Recently, Wallert et al.[17] applied shortcut methods for the synthesis and evaluation of a hybrid distillation/melt crystallization processes.

Berry and Ng[18] also used simplified models, based on constant separation factors, for the synthesis of hybrid distillation/melt crystallization processes and proposed guidelines for flowsheet selection. Bek-Pedersen and Gani[19] developed a driving-force-based approach to find the optimal combinations of unit operations and applied this approach to a hybrid distillation/membrane process for separation of ethanol and water.

A complete automated method for rapid generation of distillations sequences with recycles for azeotropic homogeneous and heterogeneous mixtures has been developed by Wasylkiewicz and Castillo.[20] The method generates feasible

distillation sequences including nonisobaric distillation columns and decanters by using residue curves and information about all azeotropes in the mixture. The method is implemented in the computer program DISTIL.

All above mentioned methods consider mainly the DOF of class one, so that further effort (normally time-consuming simulation studies) is necessary to fix the DOF of Class 2 and 3.

In contrast, mathematical programming, especially mixed-integer nonlinear programming (MINLP), offers the possibility to solve the design problem simultaneously. Four main algorithms are published in literature and are listed here in chronological order: generalized benders decomposition (GBD),[21,22] branch-and-bound method (BB),[23] outer approximation (OA),[24] and extended cutting planes (ECP).[25] All MINLP methods base on an appropriate superstructure representing the set of feasible process alternatives and model the existence and nonexistence of process units by integer variables. Many applications of MINLP methods for design and optimization of distillation sequences are reported in literature. For example, Barttfeld et al.[26] applied MINLP methods to the synthesis and optimization of complex zeotropic and azeotropic distillation sequences with rigorous models. Recently, Kossack et al.[27] presented an optimization approach for distillation columns for homogenous azeotropic mixtures. The robustness of the method is increased by using a successive relaxed mixed-integer nonlinear programming algorithm (SR-MINLP) where the binary variables are forced to take discrete values by successively tightening the integer constraints.

However, for the synthesis and optimization of hybrid separation processes by rigorous models only a few applications are reported. Brusis[28] investigated the hybrid distillation/membrane process for the separation of a nonideal ternary mixtures with MINLP methods and rigorous models. Szitkai et al.[29] determined the optimal operating and structural parameters of the hybrid distillation/membrane process for the separation of ethanol and water by MINLP and Glanz[30] presented the MINLP optimization of hybrid distillation/extraction processes.

It should be noted that the creation of a superstructure is not a trivial task, especially for hybrid separation processes. However, the main drawback of all mathematical programming methods mentioned above is that they cannot guarantee to find the global optimum, if nonconvex equations are present. It should be noted that nonconvex equations occur in almost all chemical engineering applications because component balances contain the product of variable molar or mass flow rates and molar or mass fractions.

Deterministic global MINLP methods based on branch-and-bound strategies have been developed such as the branch-and-reduce algorithm by Ryoo and Sahinidis,[31] SMIN-αBB and GMIN-αBB by Adjiman et al.,[32] and the symbolic reformulation and spatial branch-and-bound algorithm by Smith and Pantelidis,[33] but due to their high computational effort they are not yet capable of solving large-scale process engineering problems. Stochastic methods like simulated annealing[34] and evolutionary algorithms[35] have been successfully applied to process engineering optimization problems. Because of their stochastic nature the likelihood of getting trapped in local solutions is reduced, but stochastic methods also require a high computational effort. Furthermore, they contain "tuning" parameters like the number of population members as in the case of evolutionary algorithms, which have to be determined in time-consuming trial-and-error procedures.

In summary, so far no algorithm exists which can solve the complete design problem to global optimality in a reasonable time. Therefore, a new approach, which will be presented in the next section of this article, has been developed toward a general design of hybrid separation processes.

Design Approach

The general three-step design approach is shown in Figure 2. In the first step, the unit operations which are included in the flowsheet are selected by heuristics or engineering judgement, and different process alternatives are developed based on defined heuristic rules. All alternatives are evaluated by shortcut methods on the basis of their energy consumption in the second step. In the third and last step, a reduced number of alternatives with the lowest energy consumption is selected and rigorously optimized by a modified GBD algorithm to find the best alternative with respect to the total annualized costs (TAC).

This approach overcomes two limitations of common MINLP approaches: First, the effort in the MINLP optimization is greatly reduced due to fact that only a reduced number of process alternatives has to be considered in the MINLP optimization. Second, the modified GBD algorithm takes the nonconvexities into account and therefore the optimization problem can be solved to global optimality, i.e. minimum total annualized costs.

It is important to note that this approach relies on the assumption that the energy consumption is a significant criterion to judge whether an alternative is promising or not. If

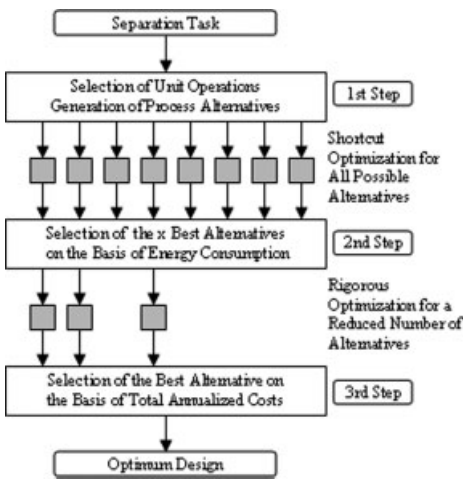

Figure 2. Three-step design approach for the synthesis of hybrid separation processes.

the assumption does not hold true, it may lead to wrong solutions, because no information feedback from the rigorous to the shortcut optimization step exists.

Process Example

The design approach for hybrid separation processes is explained with an industrial process example. The task is to separate a mixture of ortho, meta, and para isomers containing less than 1% of the low boiling meta component (M), about 66% of the intermediate boiling para component (P), and about 33% of the high boiling ortho component (O). For each component, a purity of 99% is demanded. For the considered example, distillation alone is feasible, but economically not attractive because of the low separation factors, as shown in Table 1, whereas the separation of the component P and O is especially difficult.

The combination with melt crystallization offers the advantage of obtaining almost pure product in a few stages with very high separation factors, but the yield is limited by the eutectic troughs as shown in Figure 4a. The eutectic troughs connect the binary eutectics E_1, E_2, and E_3 with the ternary eutectic E_4 and subdivide the triangular diagram into three saturation regions in which only one pure product can be obtained. For the feed F_2 shown in Figure 4a, the intermediate boiler P can be crystallized, whereas for crystallization of the other components further distillation steps are necessary. The solid product S_1, S_2, and S_3, the feed F_2, and the melt residue R_1, R_2, and R_3 lie on the balance line of the crystallization unit. By lowering the crystallization temperature, the intermediate boiler P precipitates and the residue composition R moves along the crystallization path F_2–R_3 until it reaches the eutectic trough. From this point, a second component crystallizes which is undesired and will not be considered in the design of the hybrid distillation/melt crystallization process. In this article, the composition at the point R_3 is called eutectic trough composition and the temperature at this point will be referred as eutectic trough temperature.

Generation of Process Alternatives

To systematically generate the flowsheet alternatives, the following rules are set up:

(1) A maximum of four unit operations are allowed in the flowsheet.

(2) No stream splitting is allowed.

(3) Only products on specification are allowed to leave the process. If a stream is on specification, this stream will not be recycled.

Table 1. Separation Factors $\alpha_{i,j}$ for the Isomer Mixture (Separation Factors are Approximated by the Ratio of the Vapor Pressures of the Low and High Boiling Component)

i\j	*M*	*P*	*O*
M	1	1.14	1.25
P		1	1.09
O			1

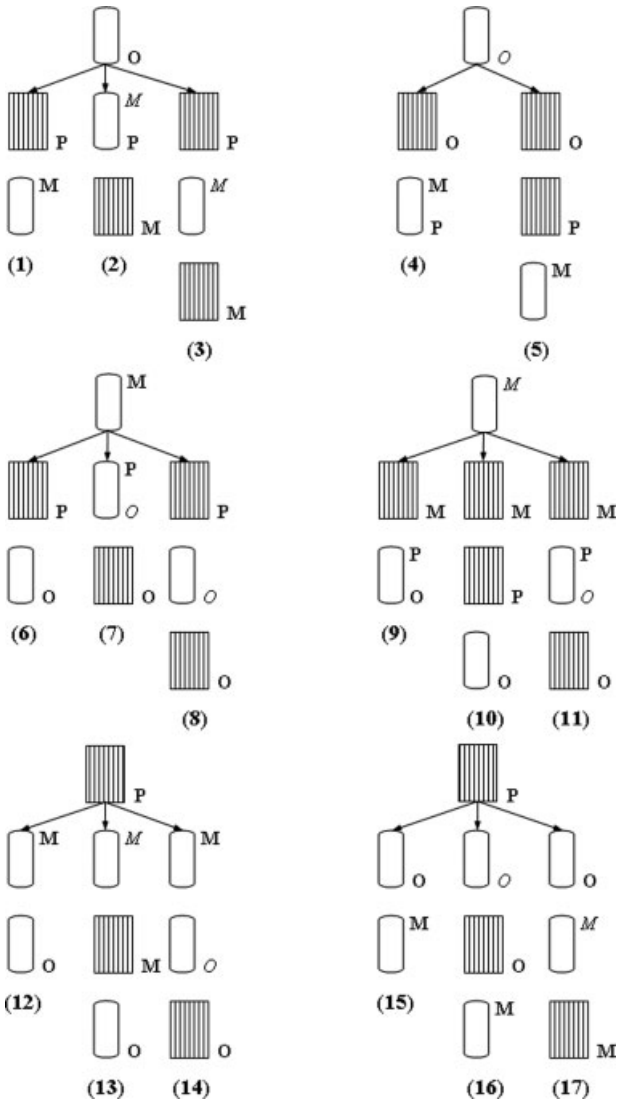

Figure 3. Developed hybrid distillation/melt crystallization sequences.

Bold letters, product on specification; italic letters, change of saturation region.

(4) Only simple, one-feed-two-products, distillation units are allowed.

(5) Distillation is used either to obtain product(s) or to change the saturation region.

When applying these rules, 19 feasible sequences can be developed. Figure 3 shows 17 sequences, whereas the two distillation sequences, i.e. the indirect (18) and direct (19) sequence, have to be added. As first separation step, the feed mixture is either separated by distillation of the components O or M (sequences 1–11) or by crystallization of component P (sequences 12–17). Process streams, which are not on specification, are processed in further separation steps until streams are obtained which are either on specification or contain only components which have been already obtained as products in previous separation steps. These streams are recycled, but the recycle streams are not shown in Figure 3.

The representation of the process alternatives in Figure 3 is explained with sequence 16 which is also shown in detail in Figure 8. The feed F and the recycle stream B2 are crystallized in the crystallizer C-1 to obtain the intermediate boiling component P. The residue R1 and recycled residue R2 of the second crystallizer C-2 are processed further in distillation D-1 to change to the ortho saturation region. The bottom product B1 is separated by crystallization C-2 to obtain the high boiling component O. The top product D1 is fed to a second distillation D-2 to get the low boiler M at top, whereas the bottom product B2 is recycled and mixed with the feed.

Modeling of Distillation

Shortcut model

The shortcut model is based on the well-known equation of Underwood,[36] which gives the minimum reflux ratio under the assumption of constant volatility of the components and constant molar overflow. The minimum heating energy for distillation is calculated by:

$$Q_{\mathrm{H,min}} = D(1 + v_{\min})\Delta h_{\mathrm{v}}, \tag{1}$$

The additional energy through preheating of the feed has to be considered:

$$Q_{\mathrm{HTX}} = F(h'_{\mathrm{F}} - h_{\mathrm{F}}), \tag{2}$$

Because the temperature levels of the distillation and crystallization differ significantly.

Besides the Underwood equation for calculation of the minimum reflux ratio, the Fenske,[37] Gilliland,[38] and Kirkbride[39] method are applied to calculate the minimum number of stages, the theoretical number of stages and the optimal feed stage, respectively, which are used as initial values for the rigorous optimization. It should be noted that the actual reflux ratio is assumed to be 1.1 times of minimum reflux ratio.

Rigorous model

The model is based on the well known equilibrium stage concept, e.g. described by Stichlmair and Fair[40] (see Appendix A). To model different feed positions and different numbers of stages an approach, similar to the one of Viswanathan and Grossmann,[41] has been used. The feed stage location is modeled by the binary variables $y_{\mathrm{F},n}$:

$$F_n \leq F_{\max} \cdot y_{F,n}, \tag{3}$$

whereas the constant $F_{\max}$ is set equal to the feed rate F_n. If $y_{\mathrm{F},n}$ is set equal to 1, the corresponding stage is supplied by the feed flow. By introducing the mass balance

$$F = \sum_{n=\mathrm{NF}z-7}^{\mathrm{NF}z+7} F_n \tag{4}$$

it is ensured that the total feed is distributed to all possible feed stages. For the considered example, only one out of 15 possible feed stages are allowed by using the following integer constraint:

$$\sum_{n=\mathrm{NF}z-7}^{\mathrm{NF}z+7} y_{\mathrm{F},n} = 1. \tag{5}$$

The number of stages is modeled by relocating the reboiler vapor stage. Binary variables denote whether the reboiler vapor is supplied to a stage or not:

$$R_n \leq R_{\max} \cdot y_{\mathrm{S},n}, \tag{6}$$

whereas $R_{\max}$ results from Eq. 49. The mass balance is satisfied by:

$$R = \sum_{n=N\max-9}^{N\max} R_n. \tag{7}$$

The following integer constraint ensures that only one out of the last 10 bottom stages is chosen for the reboiler vapor:

$$\sum_{n=N\max-9}^{N\max} y_{\mathrm{S},n} = 1. \tag{8}$$

A maximum of 150 stages is allowed to avoid columns over 80 m or column splitting. It should be noted that Eqs. 5 and 8 define special ordered sets of Type 1 (SOS-1) (see also Eq. 50).

The number of stages which is calculated by

$$N_{\mathrm{S}} = \sum_{n=N\max-9}^{N\max} \mathrm{ord}(n) \cdot \alpha_{\mathrm{S},n} \tag{9}$$

with

$$\alpha_{\mathrm{S},n} - y_{\mathrm{S},n} = 0 \tag{10}$$

is used to determine the purchased costs of the distillation column (see Appendix B). The continuous variables $\alpha_{\mathrm{s},n}$ are introduced in Eq. 9 to obtain the integer variables only in linear equations as it is the case for the common formulation of the MINLP optimization problem (see Eq. 42).

Modeling of Melt Crystallization

In layer melt crystallization,[42] which is considered in this article, a crystal layer is formed at a cooled wall. Because of the high growth rates, liquid melt is included in the crystal layer so that the desired purity is usually not obtained in one stage as shown in Figure 4. The purity is increased by adding purification stages, whereas the yield is maximized by additional stripping stages.

Melt crystallization is a batch process, i.e. in each stage cooling, crystallizing and melting is done consecutively. Additional purification steps like sweating and washing are possible, but not considered in this article. The process is usually carried out in one crystallizer, and the residue and crystal product quantities are stored in buffer tanks.[42] For N

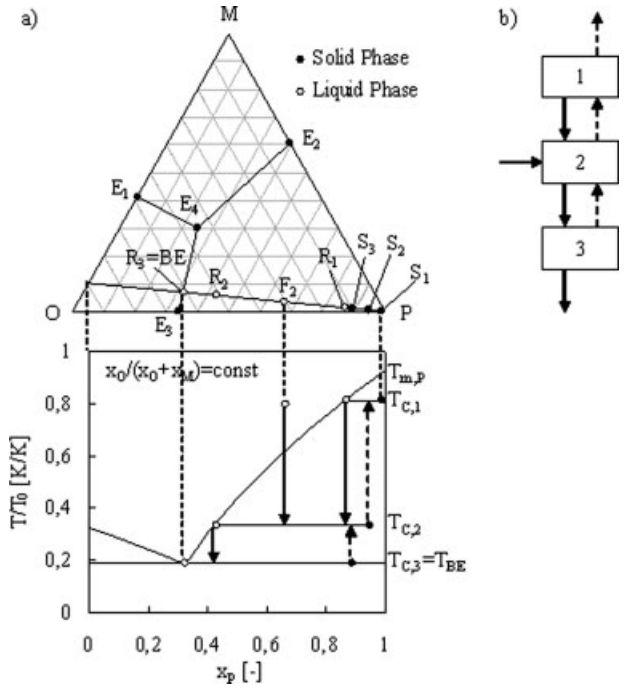

Figure 4. (a) Multi-stage crystallization in polythermal ternary diagram and phase diagram with corresponding temperatures and (b) corresponding cascade with one feed stage (Stage 2), one purification stage (Stage 1) and one stripping stage (Stage 3).

crystallization stages a maximum of N buffer tanks is required. Although the different crystallization stages are run through consecutively, the mass balance of an N stage crystallization process is identical to that of a continuous counter-current mass-transfer cascade (see Figure 4b). The batch nature of the process has to be considered only in the sizing routines (see Appendix D).

Melt crystallization shortcut model

The melt crystallization shortcut model relies on the following assumptions:

- The desired crystal product is pure and obtained in one stage.
- The crystallization temperature is equal to the eutectic trough temperature.
- The cooling energy is assumed to be equal to the energy for crystallization.

To solve the model Eqs. 14–22, the crystallizing component denoted by the index cryst has to be known. It holds: The component with the highest saturation temperature at feed composition starts to precipitate first, which is described by the following if-statement:

$$x_{\text{cryst}} = \begin{cases} x_{\text{M}} & \text{if} \quad T_{\text{sat,M}} > T_{\text{sat,P}} > T_{\text{sat,O}} \\ x_{\text{P}} & \text{if} \quad T_{\text{sat,P}} > T_{\text{sat,M}} > T_{\text{sat,O}} \\ x_{\text{O}} & \text{if} \quad T_{\text{sat,O}} > T_{\text{sat,M}} > T_{\text{sat,P}} \end{cases} \tag{11}$$

with

$$T_{\text{sat},i} = \frac{1}{\frac{1}{T_{\text{m},i}} - \frac{R}{\Delta h_{\text{m},i}} \ln(\gamma_i x_{\text{F},i})} \text{ for } i = M, P, O, \tag{12}$$

whereas the activity coefficients are calculated by the Wilson model.

Depending on the ratio of the noncrystallizing components in the feed, i.e. impurity components denoted by the index *imp1* and *imp2*, a second component precipitates when the crystallization path hits the eutectic trough. For example, in Figure 4a, the crystallization path intersects the eutectic trough E_3–E_4 and therefore component O crystallizes. The second crystallizing component is denoted by the index cryst2 and determined by

$$x_{\text{cryst2}} = \begin{cases} x_{\text{imp1}} & \text{if} \quad \frac{x_{\text{F,imp1}}}{x_{\text{F,imp2}}} > \frac{x_{\text{E4,imp1}}}{x_{\text{E4,imp2}}} \\ x_{\text{imp2}} & \text{if} \quad \frac{x_{\text{F,imp1}}}{x_{\text{F,imp2}}} < \frac{x_{\text{E4,imp1}}}{x_{\text{E4,imp2}}} \end{cases}. \tag{13}$$

The equations for the shortcut model are given by the total mass balance of the crystallization unit

$$F = S + R, \tag{14}$$

the mass balance of the crystallizing component

$$F \cdot x_{\text{F,cryst}} = S \cdot x_{\text{S,cryst}} + R \cdot x_{\text{R,cryst}}, \tag{15}$$

the summation of the mole fractions to one for the residue and the solid product

$$\sum x_{\text{S},i} = 1 \quad \text{and} \quad \sum x_{\text{R},i} = 1, \tag{16}$$

respectively, the assumption that pure crystals are obtained

$$x_{\text{S,cryst}} = 1 \tag{17}$$

and that therefore no other components are trapped in the crystal

$$x_{\text{S,imp1}} = 0 \tag{18}$$

and the assumption that the relation of the impurities in the feed and the residue are the same:

$$\frac{x_{\text{F,imp1}}}{x_{\text{F,imp2}}} = \frac{x_{\text{R,imp1}}}{x_{\text{R,imp2}}}. \tag{19}$$

The mole fraction of the crystallizing component in the residue is given by

$$x_{\text{R,cryst}} = \frac{1}{\gamma_{\text{cryst}}} \exp\left(\frac{\Delta h_{\text{m,cryst}}}{R}\left(\frac{1}{T_{\text{m,cryst}}} - \frac{1}{T_{\text{C}}}\right)\right) \tag{20}$$

with the crystallization temperature

$$T_{\text{C}} = \frac{1}{\frac{1}{T_{\text{m,cryst2}}} - \frac{R}{\Delta h_{\text{m,cryst2}}} \ln(\gamma_{\text{cryst2}} x_{\text{R,cryst2}})}. \tag{21}$$

By knowing the flow rate of solid product the cooling energy is now calculated:

$$Q_{C,\min} = S \cdot \Delta h_{m,cryst}. \tag{22}$$

Rigorous melt crystallization model

The rigorous melt crystallization model bases on the shortcut model (see Eqs. 11–22). The total mass balance and the mass balance of crystallizing component are modified to take the additional inlet streams of the neighboring stages into account (see also Figure 4b):

$$F_n + S_{n+1} + R_{n-1} = S_n + R_n \tag{23}$$

and

$$F_n x_{F,n,cryst} + S_{n+1} x_{S,n+1,cryst} + R_{n-1} x_{R,n-1,cryst} = S_n x_{S,n,cryst} + R_n x_{R,n,cryst} \tag{24}$$

respectively. Possible feed streams, residue, and crystal product from the neighboring stages are mixed before crystallized. The flow rate and the composition of the mixing stream are calculated by

$$M_n = F_n + S_{n+1} + R_{n-1} \tag{25}$$

and

$$M_n x_{M,n,i} = F_n x_{F,n,i} + S_{n+1} x_{S,n+1,i} + R_{n-1} x_{R,n-1,i} \tag{26}$$

respectively. This mixing stream composition substitutes the feed composition of Eqs. 12, 13, and 19 of the shortcut model.

The residue composition on each stage is not restricted to the eutectic trough composition; instead the residue composition is determined by the yield:

$$q_n = \frac{S_n}{M_n}. \tag{27}$$

It is worth noting that the energy and mass balance are not coupled like in adiabatic distillation, because crystallization and melting are done consecutively. Therefore, the yield is an additional optimization variable. An auxiliary variable θ_n which is allowed to take values between 0 and 1 is introduced on each stage

$$\theta_n = \frac{x_{F,n} - x_{R,n}}{x_{F,n} - x_{BE,n}} \quad \text{with} \quad 0 \le \theta_n \le 1 \tag{28}$$

to ensure that the residue composition lies between feed and eutectic trough composition. The eutectic trough composition is calculated by solving Eq. 12 for the two saturated components simultaneously.

In the rigorous melt crystallization model, the trapping of impurities in the crystal phase is considered. This complicated physical behavior is expressed in a very simplified way by a distribution coefficient which is the ratio between the impurity concentration in the solid and liquid phase,[43] p. 93:

$$k = \frac{1 - x_{S,n,cryst}}{1 - x_{F,n,cryst}}. \tag{29}$$

It should be noted that the coefficient k is influenced by the feed concentration, the crystal growth rate, and flow conditions and has been determined experimentally for each system.

A superstructure of the crystallization unit has been developed to account for different feed stages and number of stages. The feed stage is modeled analogous to distillation by the following equations:

$$F_n \le F_{max} \cdot y_{F,n}, \tag{30}$$

$$\sum_{n=1}^{5} y_{F,n} = 1, \tag{31}$$

$$F = \sum_{n=1}^{5} F_n, \tag{32}$$

whereas F_{max} is always set equal to the feed stream F.

If the yield $q_n - y_{S,n} \le 0$ is zero on a stage, no solid product is obtained, and therefore, this stage does not contribute to the separation. Consequently, the existence or nonexistence of stages is modeled by the following constraint

$$q_n - y_{S,n} \le 0, \tag{33}$$

whereas for $y_{S,n} = 1$, the yield may take any value between zero and one and for $y_{S,n} = 0$, the yield is zero. The constraint

$$y_{S,n+1} - y_{S,n} \le 0 \tag{34}$$

ensures that the stages are ordered sequentially. The continuous variable $\alpha_{S,n}$ denotes the existence or nonexistence of the buffer tanks as shown in Appendix D:

$$\alpha_{S,n} - y_{S,n} = 0. \tag{35}$$

As for the distillation, by using the variable $\alpha_{S,n}$, the integer variables appear only in linear equations.

The maximum number of stages is fixed at five for all optimization studies. The minimum number of stages for crystallization results from Eq. C6 derived in Appendix C.

No explicit energy balance is used for the rigorous crystallization model, instead the cooling energy is estimated by the equation of Wellinghoff and Wintermantel[44]:

$$Q_C = S_1 \Delta h_{m,cryst} K \left(1 + \frac{Q_{pump}}{Q_m} + \frac{Q_{app}}{Q_m}\right) \tag{36}$$

with

$$0.1 \le \frac{Q_{pump}}{Q_m} \le 1.2 \quad \text{and} \quad 1.0 \le \frac{Q_{app}}{Q_m} \le 3.0 \tag{37}$$

and the so called crystallization effort:

$$K = \frac{\sum S_n}{S_1}. \tag{38}$$

The temperatures of the crystal product and residue outlet streams correspond to their saturation temperatures, which are calculated by Eq. 12. The calculation of dimensions and costs of the melt crystallizer are shown in Appendix D.

Optimization

Optimization with shortcut models

The total energy demand of the process is minimized using shortcut models:

$$Z = \min_{\mathbf{x}} f(\mathbf{z}) = w_1 \sum Q_{\mathrm{H,min},i} + w_2 \sum Q_{\mathrm{C,min},i} \tag{39}$$

with respect to

$$\begin{aligned} &\mathbf{h}(\mathbf{x},\mathbf{z}) = \mathbf{0} \\ &\mathbf{x} \in X = \{\mathbf{x} | \mathbf{x}^{\mathrm{l}} \le \mathbf{x} \le \mathbf{x}^{\mathrm{u}}, \mathbf{x} \in R^{\mathrm{nx}}\}. \\ &\mathbf{z} \in Z = \{\mathbf{z} | \mathbf{z}^{\mathrm{l}} \le \mathbf{z} \le \mathbf{z}^{\mathrm{u}}, \mathbf{z} \in R^{\mathrm{nz}}\} \end{aligned} \tag{40}$$

The total energy demand consists of the reboiler duties of the distillation columns and the cooling energy of the crystallization units, whereas the energies are weighted with cost factors for cooling and heating energy. The constraints are given by the model equations like Eqs. 1, 2, and 11–22 and the fixed product specifications. Optimization variables are the hand-over purities $\mathbf{x}$ in all interconnecting streams.

For example, the optimization variables of sequence 16 are the hand-over purities of the high boiler P in D1 and B1 as well as the hand-over purity of the low boiler M in B2, so the DOF are three for the shortcut optimization in this case. Binary variables have not to be considered in the shortcut optimization because the number of stages of the distillation and the crystallization model are assumed to be infinite and one, respectively.

Optimization with rigorous models

In the rigorous optimization, the total annualized costs (TAC) of the process are minimized

$$\begin{aligned} Z = \min_{\mathbf{u},\mathbf{x},\mathbf{y}} f(\mathbf{z}) = C_{\mathrm{lot,a}} &= C_{\mathrm{op,a}} + C_{\mathrm{inv,a}} \\ &= C_{\mathrm{op}} \cdot t_{\mathrm{a}} + 0,25 \cdot 4 \cdot C_{\mathrm{eqp}} \end{aligned} \tag{41}$$

with respect to

$$\begin{aligned} &\mathbf{h}(\mathbf{u},\mathbf{x},\mathbf{z}) = \mathbf{0} \\ &\mathbf{z} - \mathbf{y} = \mathbf{0} \\ &\mathbf{z} - \mathbf{M}\mathbf{y} \le \mathbf{0} \\ &\mathbf{u} \in U = \{\mathbf{u} | \mathbf{u}^{\mathrm{l}} \le \mathbf{u} \le \mathbf{u}^{\mathrm{u}}, \mathbf{u} \in R^{\mathrm{nu}}\} \\ &\mathbf{x} \in U = \{\mathbf{x} | \mathbf{x}^{\mathrm{l}} \le \mathbf{x} \le \mathbf{x}^{\mathrm{u}}, \mathbf{x} \in R^{\mathrm{nx}}\} \\ &\mathbf{y} \in Y = \{\mathbf{y} | \mathbf{D}\mathbf{y} \le \mathbf{d}, \mathbf{E}\mathbf{y} = \mathbf{e}, \mathbf{y} \in \{0,1\}^{\mathrm{ny}}\} \\ &\mathbf{z} \in \mathbf{Z} = \{\mathbf{z} | \mathbf{z}^{\mathrm{l}} \le \mathbf{z} \le \mathbf{z}^{\mathrm{u}}, \mathbf{z} \in R^{\mathrm{nz}}\} \end{aligned} \tag{42}$$

The TAC are the sum of the annualized capital costs and the annual operating costs assuming 8000 h operation per year. The capital costs are estimated to be 25% of the installed equipment costs. The costs of the installed equipment are calculated by multiplying the purchased equipment costs of Appendix B and D with the Factor 4. The constraints are given by the model equations involving only continuous variables (e.g. Eqs. A1–A10 and 23–29), the inequalities involving continuous as well as integer variables (Eqs. 3, 6, 10, 30, 33, and 35), purely integer constraints (Eqs. 5, 8, 31, and 34), and the product specifications. The optimization variables are the hand-over purities $\mathbf{x}$, the unit operation variables $\mathbf{u}$ and the binary variables $\mathbf{y}$.

For example, for sequence 16 the hand-over purities of the high boiler O in D1 and B1, the hand-over purity of the low boiler M in B2 as well as the hand-over purities of the two residue streams R1 and R2, the internal variables of the crystallizers, i.e. the yield, and the binary variables denoting the feed stages and number of stages may be set independently. It should be noted that the reflux ratio and the reboiler duty are dependent variables because they are already fixed by the hand-over purities and/or product specifications, also the yield on two stages of each crystallizer is set by the hand-over purities and product specifications. For sequence 16, the number of DOF is therefore 13, i.e. five DOF during operation and eight DOF of the process topology. If a crystallizer has more than two stages, the number of DOF will be increased by one per each additional stage.

To solve the optimization problem, the GBD algorithm is applied. The solution procedure involves iterations between the NLP subproblem and MILP master problem. The NLP subproblem is obtained by fixing the integer variables of problem (42):

$$\begin{aligned} &Z_{\mathrm{u}} = v(\mathbf{y}^k) = \min_{\mathbf{u},\mathbf{x}} f(\mathbf{z}) \\ &w.r.t. \quad \mathbf{h}(\mathbf{u},\mathbf{x},\mathbf{z}) = \mathbf{0} \\ &\qquad \mathbf{z} - \mathbf{y}^k = \mathbf{0} \\ &\qquad \mathbf{z} - \mathbf{M}\mathbf{y}^k \le \mathbf{0} \\ &\qquad \mathbf{u} \in X = \{\mathbf{u} | \mathbf{u}^{\mathrm{l}} \le \mathbf{u} \le \mathbf{u}^{\mathrm{u}}, \mathbf{u} \in R^{\mathrm{nu}}\} \\ &\qquad \mathbf{x} \in X = \{\mathbf{x} | \mathbf{x}^{\mathrm{l}} \le \mathbf{x} \le \mathbf{x}^{\mathrm{u}}, \mathbf{x} \in R^{\mathrm{nx}}\} \\ &\qquad \mathbf{z} \in \mathbf{Z} = \{\mathbf{z} | \mathbf{z}^{\mathrm{l}} \le \mathbf{z} \le \mathbf{z}^{\mathrm{u}}, \mathbf{z} \in R^{\mathrm{nz}}\} \end{aligned} \tag{43}$$

The solution of the NLP subproblem gives the optimal solution for a particular combination of integer variables including the variable values, the Lagrange multipliers of each constraint and an upper bound on the final optimum. Problem (43) is also called perturbation function which will be referred to later.

The variables and Lagrange multipliers are used to set up the Lagrangian function which represents a local linearization around the optimum solution of the last NLP subproblem:

$$L^k(\mathbf{y},\mathbf{z}^k,\lambda^k,\mu^k) = f(\mathbf{z}^k) + (\boldsymbol{\lambda}^k)^{\mathrm{T}}(\mathbf{z}^k - \mathbf{y}) + (\boldsymbol{\mu}^k)^{\mathrm{T}}(\mathbf{z}^k - \mathbf{M}\mathbf{y}). \tag{44}$$

The master problem is constructed from the Lagrangian function and the purely integer constraints:

$$\begin{aligned} &Z \ge Z_{\mathrm{l}} = \min_{\mu_{\mathrm{B}},y} \mu_B \\ &w.r.t. \quad \mu_{\mathrm{B}} \ge L^k(\mathbf{y},\mathbf{z}^k,\boldsymbol{\lambda}^k,\boldsymbol{\mu}^k), \quad k = 1,\ldots,K \\ &\qquad \mathbf{y} \in Y = \{\mathbf{y} | \mathbf{D}\mathbf{Y} \le \mathbf{d}, \mathbf{E}\mathbf{y} = \mathbf{e}, \mathbf{y} \in \{0,1\}^{\mathrm{ny}}\} \end{aligned} \tag{45}$$

The solution of the MILP master problem yields a new combination of integer variables and a lower bound upon the solution. In each iteration step a new Lagrangian function is added, so that the solution of the master problem provides a

sequence of nondecreasing lower bounds. The algorithm stops when the lower bound of the master problem equals the upper bound of the NLP subproblem.

The GBD algorithm does not guarantee a global optimum, if nonconvex equations are present, because the solution of the NLP subproblem may not be unique or the master problem may not provide valid lower bounds so that the optimal solution is cut-off. The uniqueness of the solution of each NLP subproblem has been verified by perturbation of the initial values of the continuous optimization variables (see Table A1 for hand-over purities) by 0.1 or the maximum allowable concentration range. However, the second reason turned out to be crucial because the master problem did not provide valid lower bounds in all cases.

To show how the finding of the optimal solution of the optimization problem (42) is prevented when using the standard GBD algorithm, the objective function values, i.e. TAC, of sequence 16 are plotted for different numbers of stages of the distillation column D-1. For the analysis, only the number of stages of the distillation column D-1 is varied, but all other continuous and integer variables are fixed at their optimum values. The values of the objective function are obtained by the solution of the following perturbation function:

$$\begin{aligned} & v(y_{\mathrm{S},N\max-9},\ldots,y_{\mathrm{S},N\max}) = \min_{\mathbf{u}} f(\mathbf{z}) \\ & w.r.t. \quad \mathbf{h}(\mathbf{u},\mathbf{x}^k,\mathbf{z}) = \mathbf{0} \\ & \qquad \alpha_{\mathrm{S},n} - y_{\mathrm{S},n} = 0 \\ & \qquad R_n - R_{\max} y_{\mathrm{S},n} \le 0 \\ & \qquad \mathbf{u} \in U = \{\mathbf{u} | \mathbf{u}^{\mathrm{l}} \le \mathbf{u} \le \mathbf{u}^{\mathrm{u}}, \mathbf{u} \in R^{\mathrm{nu}}\} \\ & \qquad \mathbf{z} \in Z = \{\mathbf{z} | \mathbf{z}^{\mathrm{l}} \le \mathbf{z} \le \mathbf{z}^{\mathrm{u}}, \mathbf{z} \in R^{\mathrm{nz}}\} \end{aligned} \tag{46}$$

with

$$n = N_{\max} - 9, \ldots, N_{\max}. \tag{47}$$

Because the column specifications are fixed, problem (46) is not a real optimization problem, but only a simple simulation problem.

The corresponding Lagrange underestimator is calculated by

$$L(y_{\mathrm{S},N\max-9},\ldots,y^*_{\mathrm{S},n}\ldots,y_{\mathrm{S},N\max}) = f(\mathbf{z}^k) + \sum_{n=N\max-9}^{N\max} \lambda_n^k(\alpha_{\mathrm{S},n}^k - y_{\mathrm{S},n}) + \sum_{n=N\max-9}^{N\max} \mu_n^k(R_n^k - R_{\max} y_{\mathrm{S},n}) \tag{48}$$

with

$$R_{\max} = M \cdot R_n^*, \tag{49}$$

whereas the asterix indicates at which stage the linearization is done. The constant M in Eq. 49 is called big-M constant and will discussed in the subsequent paragraphs. It assumed that the summation condition is satisfied:

$$\sum_{n=N\max-9}^{N\max} y_{\mathrm{S},n} = 1, \tag{50}$$

i.e. only one stage can be chosen as the stage for the reboiler vapor (SOS-1).

The objective function values are plotted in Figure 5. In this so-called pseudo-continuous representation the discrete values are connected by straight lines, whereas the real run of the curve may be different.

The piecewise linear Lagrange function - constructed around stage number 134—is also shown in Figure 5 for different numbers of the big-M constant. For $N_{\mathrm{S}} = 134$, the perturbation function and the Lagrange underestimator are identical by definition. For stages below 134 the Lagrange function underestimates, the perturbation function correctly, whereas for stages above 134 the validity of the underestimator depends strongly on the big-M constant. For $M = 1$, the Lagrange function of the neighboring Stage 135 is approximated correctly, but for stage numbers higher than 135 the Lagrange function lies above the perturbation function and therefore the optimal solution ($N_{\mathrm{S}} = 136$) cannot be reached within the optimization. For $M = 2$ and $M = 3$, the Lagrange function provides a correct underestimation for more numbers of stages, but the quality of the underestimation, i.e. the distance between the perturbation function and the Lagrange underestimator, also becomes worse.

To explain why nonvalid underestimators are obtained, the perturbation function (46) is split into two partial perturbation functions:

$$\begin{aligned} & v(y_{\mathrm{S},N\max-9,\ldots,}y_{\mathrm{S},N\max}) = \min_{\mathbf{u}} f(\mathbf{z}) \\ & u.B.v. \quad \mathbf{h}(\mathbf{u},\mathbf{x}^k,\mathbf{z}) = \mathbf{0} \\ & \qquad \alpha_{\mathrm{S},n} - y_{\mathrm{S},n} = 0 \end{aligned} \tag{51}$$

and

$$\begin{aligned} & v(y_{\mathrm{S},N\max-9,\ldots,}y_{\mathrm{S},N\max}) = \min_{\mathbf{u}} f(\mathbf{z}) \\ & u.B.v. \quad \mathbf{h}(\mathbf{u},\mathbf{x}^k,\mathbf{z}) = \mathbf{0} \\ & \qquad R_n - R_{\max} y_{\mathrm{S},n} \le 0 \end{aligned} \tag{52}$$

with

$$n = N_{\max} - 9, \ldots, N_{\max}, \tag{53}$$

whereas problem (51) and (52) are written without definition of the variable sets.

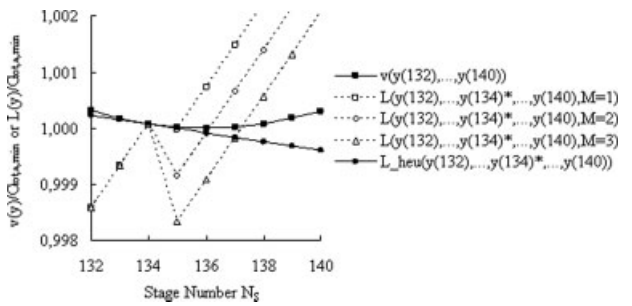

Figure 5. Pseudo-continuous perturbation function and corresponding Lagrange underestimators as well as heuristic Lagrange underestimator for column D-1 of sequence 16.

See Appendix E for numerical representation of the functions.

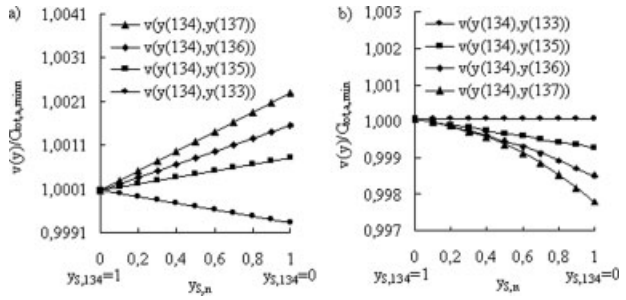

Figure 6. Perturbation function in dependence of ($y_{S,134}$, $y_{S,133}$), ($y_{S,134}$, $y_{S,135}$), ($y_{S,134}$, $y_{S,136}$), and ($y_{S,134}$, $y_{S,137}$) of column D-1 of sequence 16.

(a) problem (51) and (b) problem (52).

The perturbation function (51) describes the sensitivity of the objective function with respect to the number of stages for a fixed location of the reboiler vapor stage, whereas the perturbation function (52) shows the influence on the objective function at a fixed number of stages, but varying reboiler vapor locations.

Figure 6a shows the perturbation function (51) when going to different numbers of stages, i.e. 133, 135, 136, and 137, starting always from Stage 134. It becomes clear that the functions are almost linear and therefore can be approximated correctly by the corresponding Lagrange functions.

In Figure 6b, which shows the perturbation function (52), it can be seen that the function depending on (y_{134}, y_{133}) is constant. The reason is that the inequality constraint of problem (46) is not active and the Lagrange multiplier is zero. For variables with neighboring index (y_{134}, y_{135}), the function is almost linear and therefore can be approximated correctly, whereas for (y_{134}, y_{136}) and (y_{134}, y_{137}) the perturbation function shows a nonconvex behavior. The degree of nonconvexity is increasing when going from Stage 134 to more distant stages, but the slope, represented by the Lagrange multiplier, is always the same. To get a correct underestimation for stages above Stage 135, valid big-M constants have to be used, but the determination of a valid big-M constant is difficult because second-order information, i.e. the curvature of the perturbation function, is needed. But even if valid big-M constants are found, the algorithm needs many iterations, because the approximation is not tight (see also Figure 5). From a practical point of view, the use of the standard GBD algorithm with the big-M formulation is not a good choice for the solution of optimization problem (46).

It should be noted that the same qualitative progression of the perturbation function as shown in Figure 6 is observed for ideal distillation columns, i.e. constant molar overflow and constant separation factor, with a linear cost model. So it can be stated clearly that the nonconvexities in the presented example are not caused by thermodynamic nonidealities like the Wilson model, which is used here to calculate the activity coefficients in the liquid phase or the nonlinear cost functions rather than by bilinear terms of the equilibrium stage model itself, i.e. the multiplication of molar flow rates and mole fractions (see Eq. A2), which are the only nonlinearity in the model.

To overcome this problem of nonvalid underestimators the general GBD algorithm is modified. Based on the observations that the pseudo-continuous representation of the perturbation function is convex and the perturbation function of a neighboring stage variable is almost linear a heuristic underestimator is proposed (see Appendix F for the derivation):

$$L_{\mathrm{heu}}(y_S) = f(\mathbf{z}^k) + \begin{pmatrix} \lambda^k_{N\max-9} - \lambda^k_{N\max-8} & \text{if} \quad N^k_S = N_{\max} - 9 \\ \frac{1}{2}(\lambda^k_{N\max-9} - \lambda^k_{N\max-7}) & \text{if} \quad N^k_S = N_{\max} - 8 \\ \vdots & \\ \frac{1}{2}(\lambda^k_{N\max-2} - \lambda^k_{N\max}) & \text{if} \quad N^k_S = N_{\max} - 1 \\ (\lambda^k_{N\max-1} - \lambda^k_{N\max}) & \text{if} \quad N^k_S = N_{\max} \end{pmatrix} \left(\sum_{n=N\max-9}^{N\max} \alpha^k_{S,n}\mathrm{ord}(n) - \sum_{n=N\max-9}^{N\max} y_{S,n}\mathrm{ord}(n) \right). \quad (54)$$

It should be noted that N^k_S and N_S can be substituted by the corresponding expressions with $\alpha^k_{S,n}$ and $y_{S,n}$, respectively and vice versa (see Eq. 9).

As shown in Figure 5, this new representation leads to valid lower bounds for the optimization problem. It uses Lagrange multipliers of the neighboring elements to extrapolate the values of the Lagrange function for the remaining elements. This approximation is valid as long as the pseudo-continuous representation of the perturbation

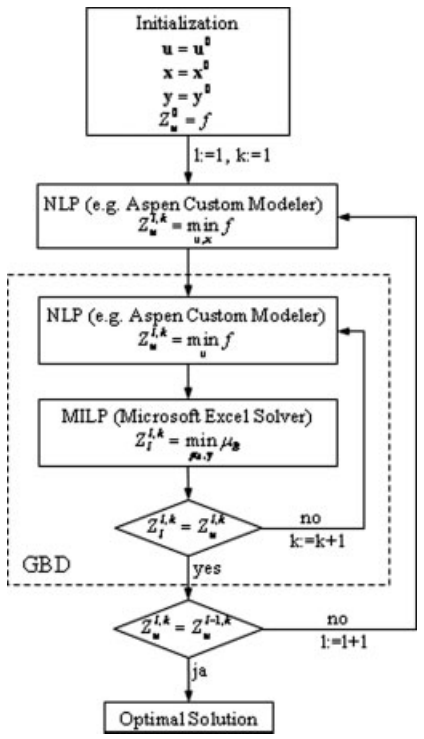

Figure 7. Modification of the GBD[21] optimization algorithm.

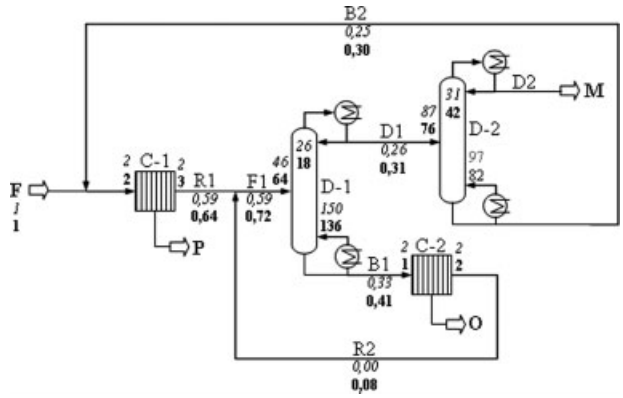

Figure 8. Flowsheet of sequence 16.

Italic and bold numbers refer to the flow rates, feed stages, number of stages, and reflux ratios before and after optimization, respectively.

function is convex and the perturbation function of two neighboring y-variables can be approximated correctly by the Lagrange function. The heuristic underestimator does not require the calculation of other data than required by the standard GBD algorithm and does not involve big-M constants.

The Lagrange function provides generally a much better approximation of the perturbation function when varying the feed stages and holding the number of stages constant. However, when recycle streams are present, the feed flow rate is not constant any more and estimation of the big-M constant becomes as difficult as for the case when the number of stages is varied. Therefore, the GBD algorithm is decomposed into an inner optimization problem where the hand-over compositions are fixed and the internal variables are optimized in each unit operation separately and an outer optimization problem where only the hand-over compositions are varied.

Figure 7 shows the approach explained above. The inner and outer optimization problems are solved iteratively until no improvement of the objective function is observed. The MILP problems are solved by the Microsoft Excel Solver,[45] whereas the NLP solution is done by the NLP algorithm of the commercial process simulator, e.g. Aspen Custom Modeler.

It should be noted that no special attention has to be paid to the optimization of the number of stages and feed stage of the crystallizer, because always valid underestimators have been obtained.

Results

Results of optimization with shortcut models

The results of the shortcut optimization procedure are shown in Table 2. It is obvious that sequence 15 has the lowest energy demand, but the difference between 15 and the sequences 17, 12, 13, 16, and 14 is smaller than 5%. The next three best sequences are sequence 6, 8, and 10 which have an energy demand about 20% higher than the optimal shortcut optimization sequence, i.e. sequence 15.

All sequences mentioned above have in common that they involve the crystallization of the intermediate boiler P. In sequences 12 to 17 the feed mixture is crystallized first. This variant seems promising because a large amount of the feed can be discharged from the process with a relatively low energy demand. In sequences 6, 8, and 10 the high boiler M is separated first by distillation or a combination of distillation and crystallization. These sequences need a considerable amount of energy to separate the minor component M with a feed mole fraction less than 1%.

All sequences that separate the heavy boiler O by distillation first or do not crystallize the intermediate boiler P show an energy demand at least 50% higher than that of the optimal sequence. It should be noted that the hand-over purities of all sequences involving crystallization of the high boiling component O are at their bounds, i.e. 0.99, because the minimum reflux ratio is not affected by the bottom composition. To overcome this problem the number of stage has to be fixed and the investment cost must be considered, to trade-off the number of stages and the reboiler duty as done in the rigorous optimization.

On the basis of the results in Table 2 sequences 12 to 17 are chosen to be rigorously optimized, whereas sequences 1, 6, 18, and 19 are optimized to check if the minimum energy demand is a reasonable criterion to choose promising alternatives. So the number of alternatives has been reduced at about one third, from 19 to 6, by the shortcut optimization procedure.

Results of optimization with rigorous models

The rigorous optimization procedure is explained exemplary for sequence 16. The initial hand-over purities are taken from the shortcut optimization, whereas the initial values for the feed stages and numbers of stages of the distillation columns are calculated by the shortcut methods. Because

Table 2. Ranking of Sequences after Shortcut Optimization (Numbering of the Distillation Columns and Crystallizers in the Order of Appearance in Figure 3)

Rank	Sequence	Description	$\frac{Q_{D-1}}{Q_{tot,min}}$ [a]	$\frac{Q_{D-2}}{Q_{tot,min}}$ [a]	$\frac{C_{C-1}}{Q_{tot,min}}$ [a]	$\frac{C_{C-2}}{Q_{tot,min}}$ [a]	$\frac{Q_{tot}}{Q_{tot,min}}$
1	**15**	C-P-D-O-D-M	0.913	0.034	0.053	0.000	**1.000**
2	**17**	C-P-D-O-D-C-M	0.913	0.034	0.053	0.001	**1.001**
3	**12**	C-P-D-M-D-O	0.135	0.831	0.053	0.000	**1.019**
4	**13**	C-P-D-C-M-D-O	0.135	0.831	0.053	0.001	**1.020**
5	**16**	C-P-D-C-O-D-M	0.913	0.034	0.053	0.034	**1.034**
6	**14**	C-P-D-M-D-C-O	0.135	0.831	0.053	0.034	**1.053**
7	**6**	D-M-C-P-D-O	0.205	0.947	0.053	0.000	**1.205**
8	**10**	D-C-M-C-P-D-O	0.228	0.929	0.001	0.053	**1.211**
9	**8**	D-M-C-P-D-C-O	0.208	0.944	0.053	0.034	**1.239**
10	**1**	D-O-C-P-D-M	1.197	0.250	0.053	0.000	**1.500**
11	**3**	D-O-C-P-D-C-M	1.197	0.250	0.053	0.001	**1.501**
12	**5**	D-C-O-C-P-D-M	1.197	0.251	0.033	0.053	**1.534**
13	**18**	D-O-D-M-P	1.978	0.137	0.000	0.000	**2.115**
14	**2**	D-O-D-P-C-M	1.978	0.142	0.000	0.000	**2.120**
15	**4**	D-C-O-D-M-P	1.978	0.139	0.033	0.000	**2.150**
16	**19**	D-M-D-P-O	0.331	1.837	0.000	0.000	**2.168**
17	**9**	D-C-M-D-P-O	0.331	1.837	0.000	0.000	**2.168**
18	**7**	D-M-D-P-C-O	0.342	1.836	0.033	0.000	**2.211**
19	**11**	D-C-M-D-P-C-O	0.344	1.836	0.000	0.033	**2.213**

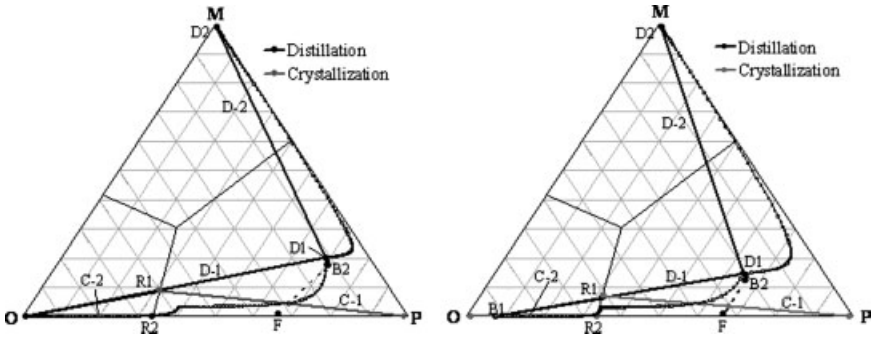

Figure 9. Representation of sequence 16 in triangular diagram before (left) and after (right) optimization.

there is no shortcut method for melt crystallization available, the second stage is defined to be the feed stage and the number of stages is set at the minimum number, i.e. two stages for this example.

The results of the initial and the final (seventh) optimization run are presented in Figure 8 and the triangular diagrams in Figure 9. The residue streams of the crystallizers C-1 and C-2 stay at their eutectic compositions even after the optimization which holds true for all examined sequences. The crystallizer C-1 involves one feed stage, one rectifying stage and one stripping stage. The crystallizer C-2 has one feed stage and one stripping stage. The main column D-1 involves 136 stages with the feed on the 64th stage. The distillation unit D-2 is smaller (82 stages and the feed on the 76th stage) and consists mainly of rectifying stages to obtain the high boiling component M. Both columns, D-1 and D-2, are operating very close to their minimum reflux, whereas the reflux ratios are 18.1 and 41.7, respectively. Although the reflux ratio of the distillation D-2 is extremely high, the energy demand is moderate due to the small amount of top product.

The hand-over purities have a significant influence on the objective function which is indicated by Table A1. Without changing the binary variables the objective function is improved by about 9 % after solving the first NLP problem. Although the flow rates of the internal streams, F1, D1, B1, B2, and R2, increase, the total energy demand declines. This is because of the less sharp separation in distillation D-1, i.e. the top and bottom concentration of high boiler O change from 0.113 to 0.209 and from 0.99 to 0.923, respectively, whereas the energy demand of the distillation D-2 increase because the concentration of the low boiler M in the bottom goes down from 0.176 to 0.124. Generally speaking, the outer NLP problem can be considered a simple, but very important trade-off between the amount of recycle and the sharpness of separation.

There is also strong influence of the feed stage position and the number of stages of the crystallizers on the total annualized costs, instead the influence of feed stages and the number of stages of the distillation columns is small. This is not a special problem of isomer separations, because even for other systems the number of stages is not sensitive when the reflux ratio is close to the minimum reflux ratio, as Bauer[46] showed. To avoid additional iterations with only slight improvement of the objective function, the algorithm could be stopped when the improvement is small, e.g. less than 0.1%.

Table 3 shows that the objective function values for sequence 12 by 17 are similar. However, the indirect variant, that uses the first distillation to obtain pure low boiler or move in its saturation region, is favored as shown for sequences 15, 16, and 17. This is in agreement with the well-known heuristic to separate minor components last and also a result of shortcut optimization. The addition of a second crystallization unit to produce the low boiler O leads to lower TAC in sequence 16, which also turns out to be the cost optimal sequence, and sequence 14, because the energy demand as well as the number of stages of the previous distillation is reduced compared to sequences without additional crystallizers, i.e. 15 and 12. This variant should become even more promising when the maximum number of stages in the distillation columns is further restricted. The reason why sequence 16 and 14 are favored here compared to the shortcut optimization is the invariance of the reboiler duty with respect to the hand-over purity due to the assumption of infinite stages in the shortcut distillation model.

The addition of a crystallization unit to obtain the low boiler M, as shown in sequences 13 and 17, is not advantageous for the examined sequences, but as indicated by shortcut optimization there may be sequences, e.g. 2 and 9, where this is favorable, especially when the low boiler feed concentration of the previous distillation is small, i.e. about 1%. Sequences 1 and 6 are not competitive as already indicated by the shortcut optimization. The costs of the pure distillation sequences are almost four times higher than the hybrid processes. The ranking obtained by the rigorous optimization

Table 3. Total Annualized Costs (TAC) of the Rigorous Optimized Sequences (TAC are Referred to the TAC of Sequence 16)

Sequence	Description	$\frac{C_{D-1}}{C_{tot,a,min}}$	$\frac{C_{D-2}}{C_{tot,a,min}}$	$\frac{C_{C-1}}{C_{tot,a,min}}$	$\frac{C_{C-2}}{C_{tot,a,min}}$	$C_{\frac{tot}{C_{tot,a,min}}}$
16	C-P-D-C-O-D-M	0.710	0.045	0.200	0.045	**1.000**
15	C-P-D-O-D-M	0.756	0.045	0.205	0.000	**1.006**
17	C-P-D-O-D-C-M	0.756	0.045	0.205	0.002	**1.008**
14	C-P-D-M-D-C-O	0.117	0.654	0.204	0.045	**1.020**
12	C-P-D-M-D-O	0.119	0.693	0.210	0.000	**1.022**
13	C-P-D-C-M-D-O	0.119	0.693	0.210	0.002	**1.024**
6	D-M-C-P-D-O	0.164	0.768	0.220	0.000	**1.152**
1	D-O-C-P-D-M	1.051	0.166	0.242	0.000	**1.459**
18	D-O-D-M-P	3.255	0.533	0.000	0.000	**3.788**
19	D-M-D-P-O	0.773	3.190	0.000	0.000	**3.963**

agrees with that of the shortcut optimization, except for the sequences 14 and 16 which have moved upward in the ranking.

Conclusions

In this article, a new three-step design method for hybrid distillation/melt hybrid processes is presented and illustrated by the separation of a ternary isomer mixture. In a first step different sequences are generated by heuristic rules. These sequences are evaluated in a second step by shortcut methods on the basis of energy consumption in order to identify the most promising alternatives. In the third and last step a reduced number of promising sequences is rigorously optimized by MINLP methods and the best sequence on the basis of total annualized costs is chosen.

It was found that the shortcut optimization provides a good basis for the ranking of the sequences, although the rigorous optimization is necessary to finally choose the best sequence on the basis of total annualized cost.

To overcome the problem of getting trapped in local solutions when non convex equations are present the GBD algorithm for the rigorous optimization was modified in two points: First, a new heuristic Lagrange underestimator for the number of distillation stages is proposed which leads to valid underestimators without using big-M constants. Second, the optimization problem is decomposed into an inner GBD and outer optimization where only the hand-over purities are varied in order to reduce nonconvexities. It should be noted that even with the modified GBD algorithm it cannot be guaranteed to find the global optimum due to assumptions underlying the heuristic Lagrange underestimator. However, parameter studies show that the algorithm is able to find the global optimum.

The approach can easily be extended to other unit operations, like extraction or absorption and gives the possibility for the cost optimal design of many different hybrid and non hybrid separation processes.

Acknowledgments

The authors thank the Bayer Technology Services GmbH, Leverkusen, Germany for financing this work. The helpful discussions with Edmund Dikow are gratefully acknowledged. Without the information supplied by Jesko Zank and Jan Thomas Leu this work would not have been possible. The authors also thank their project partners Claudia Wallert and Wolfgang Marquardt from Lehrstuhl für Prozesstechnik, RWTH Aachen, Germany for their co-operation as well as Manfred Stepanski from Sulzer Chemtech, Winterthur, Switzerland for his valuable advice in the field of melt crystallization.

Notation

a = price coefficient
A = area [m^2]
b = price coefficient
B = bottom flow rate [mol/s]
c = price coefficient
c_{cool} = price coefficient in Eq. D1 [[€] s/(J h)]
c_{steam} = price coefficient in Eq. B1 [[€] s/(mol h)]
C = costs [[€]/a] or [[€]/h] or [[€]]
D = distillate flow rate [mol/s]
D = diameter [m]
$\mathbf{d}$ = vector of integer inequality constraints
$\mathbf{D}$ = matrix of integer inequality constraints
$\mathbf{e}$ = vector of integer equality constraints
$\mathbf{E}$ = matrix of integer equality constraints
f = objective function value [[€]/a]
F = feed flow rate [mol/s]
h = specific enthalpy [J/mol]
$\mathbf{h}$ = vector of nonlinear equations
H = height [m]
HETP = height of theoretical stage [m]
k = distribution coefficient [mol/mol]
k = iteration counter
$\boldsymbol{K}$ = K-value [−]
K = crystallization effort defined by Eq. 38 [−]
l = iteration counter
L = Lagrangian function [[€]/a]
L = Length [m]
L = liquid flow rate [mol/s]
M = big-M constant [−]
M = mixing stream [mol/s]
M = molar mass [kg/mol]
$\mathbf{M}$ = matrix of big-M constants
MS = Marshall-Swift index [−]
n = mole flow [mol/s]
N = integer number
q = yield [mol/mol]
Q = energy [W]
Q_{app} = cooling energy for apparatus [W]
Q_{pump} = energy dissipated by pumping the cooling medium [W]
r = repetition factor [−]
R = residue flow rate [mol/s]
R = reboiler vapor flow rate [mol/s]
R = set of real numbers
S = solid product flow rate [mol/s]
t = time [s]
T = temperature [K]
U = set of internal optimization variables
$\mathbf{u}$ = vector of internal optimization variables
v = reflux ratio [−]
v = crystal growth rate [m/s]
V = vapor flow rate [mol/s]
V = volume [m^3]
w = weight factor [−]
x = liquid or solid mole fraction [mol/mol]
X = set of hand-over purities
$\mathbf{x}$ = vector of hand-over purities
y = vapor mole fraction [−]
y = binary variable
Y = set of binary variables
$\mathbf{y}$ = vector of binary variables
$\mathbf{z}$ = vector of dependent variables
Z = set of dependent variables
Z = objective function

Subscripts

0 = reference
a = annual
B = bottom product
BE = eutectic trough
Buffer = buffer tank
C = cooling
cryst = crystallizing component
Cryst = crystallizer
Dew = dew point
eqp = equipment
F = feed
H = heating
heu = heuristic
HTX = heat exchanger
i = index for component
imp = impurity
int = internals
jac = jacket
l = lower bound

L = liquid
m = melting
M = meta isomer
max = maximum
min = minimum
n = index for stage
NFz = number of central feed stage
N_{max} = maximum stage number
op = operating
O = ortho component
P = para component
R = reboiler vapor
R = residue
S = solid product
S = stage
sat = saturation
t = total
T = tube of crystallizer
u = upper bound
V = vapor

Superscripts

k = iteration counter
l = iteration counter
l = lower bound
nu = dimension of vector of internal optimization variables
nx = dimension of vector of hand-over purities
ny = dimension of vector of binary variables
nz = dimension of vector of dependent variables
u = upper bound

Greek letters

α = auxiliary variable
α = split factor [−]
γ = liquid activity coefficient [−]
Δh_m = melt enthalpy [J/mol]
Δh_v = evaporation enthalpy [J/mol]
λ = Lagrange multiplier of equality constraints
$\boldsymbol{\lambda}$ = vector of Lagrange multipliers of equality constraints
μ = Lagrange multiplier of inequality constraints
$\boldsymbol{\mu}$ = Lagrange multiplier
Θ = auxiliary variable
ρ = density [kg/m3]

Literature Cited

1. Jenkins RL. *Separation of Diphenylbenzene Isomers by Distillation and Crystallization*. U.S. Patent 2,489,215, 1949.
2. Weedman JA. *Separation of Xylene Isomers by Crystallization and Distillation*. U.S. Patent 3,067,270, 1962.
3. Stepanski M, Haller U. Meta-Xylol wirtschaftlich gewinnen. *Sulzer Tech Rev*. 2000;3:8–9.
4. Rittner S, Steiner R. Die Schmelzkristallisation von organischen Stoffen und ihre groβtechnischen Anwendung. *Chem Ing Tech*. 1985;57:91–102.
5. Ruegg PJ. Destillation plus kristallisation. *Chem Ind*. 1989;11:83–85.
6. Stepanski M, Fässler P. Destillation und kristallisation kombiniert. *Sulzer Tech Rev*. 2002;4:14–16.
7. Thompson RW, King CJ. Systematic synthesis of separation sequences. *AIChE J*. 1972;18:941–948.
8. Seider WD, Seader JD, Lewin DR. *Process Design Analysis*. New York: John Wiley, 1999.
9. Barnicki SD, Fair JR. Separation system synthesis: a knowledge-based approach. 1. Liquid mixture separations. *Ind Eng Chem Res*. 1990;29:421–432.
10. Simmrock KH, Fried A, Welker R. Beratungssystem für die Trennung engsiedender und azeotroper Gemische. *Chem Ing Tech*. 1991;63:593–604.
11. Wahnschafft OM, Jurain TP, Westerberg AW. SPLIT: a separation process designer. *Comput Chem Eng*. 1991;15:565–581.
12. Seader JD, Westerberg AW. A combined heuristic and evolutionary strategy for synthesis of simple separation sequences. *AIChE J*. 1977;23:951–954.
13. Biegler LT, Grossmann IE, Westerberg AW. *Systematic Methods of Chemical Process Design*. Upper Saddle River, New Jersey, USA: Prentice-Hall, 1997.
14. Hendry JE, Hughes RR. Generation of process flowsheets. *Chem Eng Prog*. 1972;68:71–76.
15. Mitten LG. Multistage optimization. *Chem Eng Prog*. 1963;59:52–60.
16. Bausa J, Marquardt W. Shortcut design methods for hybrid membrane/distillation processes for the separation of nonideal multicomponent mixtures. *Ind Eng Chem Res*. 2000;39:1658–1672.
17. Wallert C, Marquardt W, Leu JT, Strube J. Design and optimization of layer crystallization processes. In: Puigjaner L, editor. *European Symposium on Computer Aided Process Engineering—15*. Barcelona, Spain: Elsevier, 2005:871–876.
18. Berry DA, Ng KM. Synthesis of crystallisation-distillation hybrid separation processes. *AIChE J*. 1997;43:1751–1762.
19. Bek-Pedersen E, Gani R. Design and synthesis of distillation systems using a driving-force-based approach. *Chem Eng Proc*. 2004;43:251–262.
20. Wasylkiewicz SK, Castillo FJL. Automatic synthesis of complex separation sequences with recycles. *Comput Aided Chem Eng*. 2001;9:591–596.
21. Geoffrion AM. Generalized benders decomposition. *J Optim Theory Appl*. 1972;10:237–262.
22. Floudas CA. *Nonlinear and Mixed-Integer Optimization*. New York: Oxford University Press, 1995.
23. Gupta OK, Ravindran V. Branch and bound experiments in convex nonlinear integer programming. *Manage Sci*. 1985;31:1533–1546.
24. Duran MA, Grossmann IE. An outer approximation algorithm for a class of mixed integer nonlinear programs. *Math Program*. 1986;36:307–339.
25. Westerlund T, Skrifvars H, Harjunkoski I, Pörn R. An extended cutting plane method for a class of non-convex MINLP problems. *Comput Chem Eng*. 1998;22:357–365.
26. Barttfeld M, Aguirre PA, Grossmann IE. A decomposition method for synthesizing complex column configurations using tray-by-tray GDP models. *Comput Chem Eng*. 2004;28:3441–3457.
27. Kossack S, Krämer K, Marquardt W. Efficient optimization-based design of distillation columns for homogenous azeotropic mixtures. *Ind Eng Chem Res*. 2006;45:8492.
28. Brusis D. Synthesis and Optimisation of Distillation Processes with MINLP Techniques (PhD Thesis). München, Germany: Technische Universität München, 2003.
29. Szitkai Z, Lelkes Z, Rev E, Fonyo Z. Optimization of hybrid ethanol dehydration systems. *Chem Eng Proc*. 2002;41:631–646.
30. Glanz S. Synthese und Strukturoptimierung von Prozessen zur Trennung heterogener Flüssigkeitsgemische (PhD Thesis). München, Germany: Technische Universität München, 1998.
31. Ryoo HS, Sahinidis NV. Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Comput Chem Eng*. 1995;19:551–566.
32. Adjiman CS, Androulakis IP, Floudas CA. Global optimization of MINLP problems in process synthesis and design. *Comput Chem Eng*. 1997;21:S445–S450.
33. Smith EMB, Pantelides CC. Global optimization of general process models. In: Grossmann IE, editor. *Global Optimization in Engineering Design*. Dordrecht: Kluwer, 1996:355–386.
34. Floquet P, Pibouleau L, Domenech S. Separation sequence synthesis: how to use simulated annealing procedure? *Comput Chem Eng*. 1994;18:1141–1148.
35. Gross B, Roosen P. Total process optimization in chemical engineering with evolutionary algorithms. *Comput Chem Eng*. 1998;22:S229–S236.
36. Underwood AJV. Fractional distillation of multicomponent mixtures. *Chem Eng Prog*. 1948;44:603–614.
37. Fenske MR. Fractionation of straight-run Pennsylvania gasoline. *Ind Eng Chem*. 1932;24:482–485.
38. Gilliland ER. Multicomponent rectification. Estimation of the number of theoretical plates as a function of the reflux ratio. *Ind Eng Chem*. 1940;32:1220–1223.
39. Kirkbride CG. Process design procedure for multicomponent fractionators. *Pet Refiner*. 1944;23:321–332.

40. Stichlmair JG, Fair JR. *Distillation—Principles and Practices*. New York: Wiley-VCH, 1998.
41. Viswanathan J, Grossmann IE. Optimal feed locations and number of trays for distillation columns with multiple feeds. *Ind Eng Chem Res*. 1993;32:2942–2949.
42. Stepanski M, Schäfer E. Separate organics by melt crystallization: a guide when and how to use this technique. In: Ulrich J, Glade H, editors. *Melt Crystallization*. Aachen: Shaker, 2003:167–189.
43. Arkenbout GF. *Melt Crystallization Technology*. Lancaster, Pennsylvania, USA: Technomic Publishing Company, 1995.
44. Wellinghoff G, Wintermantel K. Schmelzkristallisation—theoretische Voraussetzungen und technische Grenzen. *Chem Ing Tech*. 1991;63:881–891.
45. Fylstra D, Lasdon L, Watson J, Waren A. Design and use of the microsoft excel solver. *Interfaces*. 1998;28:29–55.
46. Bauer MH. Synthese und Optimierung nichtidealer Rektifizierprozesse (PhD Thesis). München, Germany: Technische Universität München, 1997.
47. Peters MS, Timmerhaus KD. *Plant Design and Economics for Chemical Engineers*, 4th ed. New York: McGraw-Hill, 1990.

Appendix A: Rigorous Distillation Model

The model is based on the equilibrium stage concept, e.g. described by Stichlmair and Fair.[40] Any stage modeled by the total mass balance

$$F_n + R_n + L_{n-1} + V_{n+1} = L_n + V_n, \tag{A1}$$

the mass balance for each component i

$$F_n x_{F,n,i} + R_n y_{R,n,i} + L_{n-1} x_{n-1,i} + V_{n+1,i} y_{n+1,i} = L_n x_{n,i} + V_n y_{n,i}, \tag{A2}$$

the phase equilibrium

$$y_{n,i} = K_{n,i} x_{n,i}, \tag{A3}$$

the relationship for the summation of the mole fractions

$$\sum_i x_{n,i} - \sum_i y_{n,i} = 0, \tag{A4}$$

and the heat balance

$$F_n h_{F,n} + R_n h_{R,n} + V_{n+1} h_{V,n+1} + L_{n-1} h_{L,n-1} = L_n h_{L,n} + V_n h_{V,n}. \tag{A5}$$

The reboiler is modeled by the total mass balance

$$L_{N\max} = B + R \tag{A6}$$

and the mass balance of each component i

$$L_{N\max} x_{L,N\max,i} = B x_{B,i} + R y_{R,i}. \tag{A7}$$

The mole fractions and temperature of the liquid inlet and bottom product stream are assumed to be equal:

$$x_{L,N\max,i} = x_{B,i} \quad \text{and} \quad T_{L,N\max} = T_B. \tag{A8}$$

The temperature of the vapor outlet stream is equal to the dew point temperature:

$$T_R = T_{Dew}(y_{R,i}). \tag{A9}$$

The energy required in the reboiler is then calculated by:

$$Q_H = B h_B + R h_R - L_{N\max} h_{L,N\max}. \tag{A10}$$

The analysis of the DOF for the reboiler shows that there is one degree of freedom left. Here, the composition of one component in the bottom product stream is either fixed as product specification or to be determined in an optimization problem as hand-over purity. The model equations for the total condenser are similar and therefore not mentioned here.

Appendix B: Cost Model for Distillation

The operating costs of a distillation unit are dominated by the steam utility costs which are taken into account by using actual steam prices:

$$C_{op} = \frac{Q_H + Q_{HTX}}{\Delta h_{v,steam}} \cdot c_{steam}, \tag{B1}$$

The costs for cooling water are neglected.

The equipment costs of a distillation unit consist of the costs for heat exchangers, i.e. the heat exchanger for feed preheating, reboiler and condenser, and the costs for the column jacket and internals:

$$C_{eqp} = \sum C_{HTX,i} + C_{jac} + C_{int}. \tag{B2}$$

The costs of the heat exchangers are calculated by

$$C_{HTX} = \left(\frac{MS}{MS_0}\right) \cdot a \cdot A, \tag{B3}$$

whereas for determination of the heat transfer area mean logarithmic temperature differences and constant heat transfer coefficients are used. The costs of the column jacket and internals are calculated by

$$C_{jac} = \left(\frac{MS}{MS_0}\right) \cdot b \cdot D^{1.07} \cdot H^{0.8} \tag{B4}$$

and

$$C_{int} = \left(\frac{MS}{MS_0}\right) \cdot c \cdot D^{1.55} \cdot H, \tag{B5}$$

respectively. For the calculation column diameter, it is assumed that the F-factor has a value of 1.6 $\sqrt{Pa}$. The column height is determined by assuming an HETP of 0.5:

$$H = N_S \cdot \mathrm{HETP} + H_0. \tag{B6}$$

Additional height H_0 of 4 m for the top and bottom section of the column is considered. The parameters a, b, and c have been determined by using relevant cost data from literature[47] or an in-house databank.

Appendix C: Minimum Number of Crystallization Stages

Analogously to the Fenske equation in distillation, an equation for the minimum number of stages for crystallization can be derived.

The impurities distribute on each stage n between the solid and liquid phase according to (see also Eq. 29):

$$k \cdot \left(1 - x_{\mathrm{R},n,\mathrm{cryst}}\right) = \left(1 - x_{\mathrm{S},n,\mathrm{cryst}}\right), \tag{C1}$$

whereas k is assumed to be constant. The mole fraction of the crystallizing component on any stage is calculated by:

$$x_{\mathrm{S},n,\mathrm{cryst}} = 1 - k \cdot \left(1 - x_{\mathrm{R},n,\mathrm{cryst}}\right). \tag{C2}$$

Under the assumption that the yield on each stage is zero, which corresponds to the total reflux mode in distillation, $x_{\mathrm{R},n,\mathrm{cryst}}$ and $x_{\mathrm{S},n+1,\mathrm{cryst}}$ are equal and Eq. C2 can be written as

$$x_{\mathrm{S},n,\mathrm{cryst}} = 1 - k \cdot \left(1 - x_{\mathrm{S},n+1,\mathrm{cryst}}\right). \tag{C3}$$

A general form for the mole fraction of the crystallizing component on the first stage is obtained by inserting the corresponding equations of Eq. C3 into each other:

$$x_{\mathrm{S},1,\mathrm{cryst}} = 1 - k^{n-1} \cdot \left(1 - x_{\mathrm{S},n,\mathrm{cryst}}\right). \tag{C4}$$

Substituting the bracket on the right hand side by Eq. C1 gives:

$$x_{\mathrm{S},1,\mathrm{cryst}} = 1 - k^{n-1} \cdot k \cdot \left(1 - x_{\mathrm{R},n,\mathrm{cryst}}\right). \tag{C5}$$

When n is set equal to $N_{\min}$, Eq. C5 is rearranged to finally derive the expression for the minimum number of stages:

$$N_{\min} = \frac{\log\left(\frac{1-x_{\mathrm{S},1,\mathrm{cryst}}}{1-x_{\mathrm{R,N,cryst}}}\right)}{\log k}. \tag{C6}$$

Appendix D: Cost Model for Melt Crystallization

The operating costs per hour for cooling are calculated by:

$$C_{\mathrm{op}} = Q_{\mathrm{C}} \cdot c_{\mathrm{cool}}. \tag{D1}$$

The equipment costs of a melt crystallization unit consist of the costs of the crystallizer itself whereas a falling film crystallizer[42] is considered and the costs of the buffer tanks:

$$C_{\mathrm{eqp}} = C_{\mathrm{cryst}} + \sum C_{\mathrm{Buffer}},n + C_{\mathrm{Buffer,R}}. \tag{D2}$$

The crystallizer costs are calculated by

$$C_{\mathrm{Cryst}} = \left(\frac{\mathrm{MS}}{\mathrm{MS}_0}\right) \cdot a \cdot A_{\mathrm{Cryst}}^{0.65}. \tag{D3}$$

The crystallizer surface consists of the surface of N_{T} single tubes:

$$A_{\mathrm{Cryst}} = N_{\mathrm{T}} \cdot A_{\mathrm{T}} = N_{\mathrm{T}} \cdot \pi \cdot D_{\mathrm{T}} \cdot L_{\mathrm{T}}. \tag{D4}$$

The time needed for one crystallization stage is

$$t_{\mathrm{Cryst},n} = \frac{1}{v_{\mathrm{Cryst}}}\left(\frac{D_{\mathrm{T}}}{2} - \sqrt{\left(\frac{D_{\mathrm{T}}}{2}\right)^2 - \frac{m_{\max,\mathrm{T}}}{\rho_{\mathrm{S}}\pi L_{\mathrm{T}}}}\right), \tag{D5}$$

assuming that the growth rate is constant and the maximum possible crystal mass is obtained in each tube. Because the crystal mass obtained on each stage is different corresponding to the yield, some stages are repeated to fully load the crystallizer. Therefore, repetition factors[42] r_n are introduced:

$$r_n = \frac{S_n}{S_1}. \tag{D6}$$

The total time needed for all crystallization stages is then determined by

$$t_{\mathrm{tot}} = \frac{\sum r_n t_{\mathrm{Cryst},n}}{r_1}, \tag{D7}$$

when neglecting the time for cooling and melting.

The number of tubes and the crystallizer area are determined by:

$$N_{\mathrm{T}} = \frac{t_{\mathrm{tot}} \cdot S_1 \cdot M_{\mathrm{S}}}{m_{\max,\mathrm{T}}} \tag{D8}$$

and

$$A_{\mathrm{Cryst}} = N_{\mathrm{T}} \cdot A_{\mathrm{T}} = N_{\mathrm{T}} \cdot \pi \cdot D_{\mathrm{T}} \cdot L_{\mathrm{T}}, \tag{D9}$$

respectively.

The costs of the buffer tanks are calculated by:

$$C_{\mathrm{Buffer},n} = \left(\frac{\mathrm{MS}}{\mathrm{MS}_0}\right) \cdot b \cdot V_{\mathrm{Buffer},n}^{0.56}. \tag{D10}$$

The volumes of the buffer tanks are fixed by the amounts of residue and feed which have to be stored for the time t_{tot}:

$$V_{\mathrm{Buffer},n} = \alpha_{\mathrm{S},n} \cdot \frac{(R_{n-1} + F_n) \cdot M_{\mathrm{R}} \cdot t_{\mathrm{tot}}}{\rho_{\mathrm{R}}},\ n = 1, \ldots, 5. \tag{D11}$$

Additionally, a buffer tank for the residue has to be considered:

$$V_{\mathrm{Buffer,R}} = \frac{R_5 \cdot M_{\mathrm{R}} \cdot t_{\mathrm{tot}}}{\rho_{\mathrm{R}}}. \tag{D12}$$

The parameters a and b in Eqs. D3 and D10 and the cost have to determined by using relevant cost data or an in-house databank.

Appendix E: Lagrange Function for Number of Stages

The Lagrange function in (48) for the number of stages is calculated by:

$$\begin{aligned} L\left(y_{\mathrm{S},132}, \ldots, y^*_{\mathrm{S},134}, \ldots, y_{\mathrm{S},140}\right)\Big/C_{\mathrm{tot,a,min}} = {} & 1.000072186 \\ & - 0.0981 \cdot \left(0 - y_{\mathrm{S},132}\right) \\ & - 0.0988 \cdot \left(0 - y_{\mathrm{S},133}\right) - 0.0996 \cdot \left(1 - y_{\mathrm{S},134}\right) \\ & - 0.1003 \cdot \left(0 - y_{\mathrm{S},135}\right) \\ & - 0.1011 \cdot \left(0 - y_{\mathrm{S},136}\right) - 0.1018 \cdot \left(0 - y_{\mathrm{S},137}\right) \\ & - 0.1025 \cdot \left(0 - y_{\mathrm{S},138}\right) \\ & - 0.1033 \cdot \left(0 - y_{\mathrm{S},139}\right) - 0.1040 \cdot \left(0 - y_{\mathrm{S},140}\right) \\ & + 0 \cdot \left(0 - y_{\mathrm{S},132} \cdot M\right) + 0 \cdot \left(0 - y_{\mathrm{S},133} \cdot M\right) \\ & + 0 \cdot \left(1 - y_{\mathrm{S},134} \cdot M\right) \quad . \\ & + 8.261 \cdot 10^{-4} \cdot \left(0 - y_{\mathrm{S},135} \cdot M\right) + 8.261 \cdot 10^{-4} \\ & \cdot \left(0 - M \cdot y_{\mathrm{S},136}\right) \\ & + 8.261 \cdot 10^{-4} \cdot \left(0 - y_{\mathrm{S},137} \cdot M\right) + 8.261 \cdot 10^{-4} \\ & \cdot \left(0 - M \cdot y_{\mathrm{S},138}\right) \\ & + 8.261 \cdot 10^{-4} \cdot \left(0 - y_{\mathrm{S},139} \cdot M\right) + 8.261 \cdot 10^{-4} \\ & \cdot \left(0 - M \cdot y_{\mathrm{S},140}\right) \end{aligned} \tag{E1}$$

The heuristic Lagrange underestimator in (54) is calculated by:

$$L(y_{S,132},\ldots,y^*_{S,134},\ldots,y_{S,140})/C_{lot,a,min} = 1.000072186 + 7.846\cdot 10^{-5}.(134 - 132\cdot y_{S,132} - 133\cdot y_{S,133} - 134\cdot y_{S,134} - 135\cdot y_{S,135} - 136\cdot y_{S,136}. - 137\cdot y_{S,137} - 138\cdot y_{S,138} - 139\cdot y_{S,139} - 140\cdot y_{S,140}) \quad \text{(E2)}$$

Appendix F: Heuristic Lagrangian Function

The heuristic Lagrange underestimator bases on the idea that the convex pseudo-continuous can be underestimated by a linear approximation at the stage number N^k_S:

$$L_{heu}(N_S) = f(\mathbf{z}^k) + \frac{\partial f}{\partial N^k_S}(N_S - N^k_S) \text{ with } N^k_S, N_S = \{N_{max}-9,\ldots,N_{max}\} \quad \text{(F1)}$$

Because the gradient cannot be calculated, the average of the gradient of the Lagrange functions with the discrete variables with the left and right neighboring indices are taken, $L_L(y_{S,n-1}, y^*_{S,n})$ and $L_R(y^*_{S,n}, y_{S,n+1})$. The general Lagrange functions with the discrete variables with the neighboring indices are as follows:

$$L_L\left(y_{S,n-1}, y^*_{S,n}\right) = f\left(\mathbf{z}^k\right) + \lambda^k_{n-1}\left(\alpha^k_{S,n-1} - y_{S,n-1}\right) + \lambda^k_n\left(\alpha^k_{S,n} - y^*_{S,n}\right) \quad \text{(F2)}$$

and

$$L_R\left(y^*_{S,n}, y_{S,n+1}\right) = f\left(\mathbf{z}^k\right) + \lambda^k_n\left(\alpha^k_{S,n} - y^*_{S,n}\right) + \lambda^k_{n+1}\left(\alpha^k_{S,n+1} - y_{S,n+1}\right), \quad \text{(F3)}$$

whereas Eq. 6 is changed into

$$R_n = R\cdot\alpha_{S,n} \quad \text{(F4)}$$

by setting

$$R_{max} = R \quad \text{and} \quad \alpha_{S,n} - y_{S,n} = 0. \quad \text{(F5)}$$

Because the equality constraints are satisfied from the last NLP subproblem, following equations hold true:

$$\alpha^k_{S,n} = 1 \quad \text{and} \quad \alpha^k_{S,n-1} = \alpha^k_{S,n+1} = 0 \quad \text{(F6)}$$

and

$$y_{S,n-1} + y^*_{S,n} = 1 \quad \text{and} \quad y^*_{S,n} + y_{S,n+1} = 1. \quad \text{(F7)}$$

By inserting Eqs. F6 and F7 in F2 and F3 one gets:

$$L_L\left(y^*_{S,n}\right) = f(\mathbf{z}^k) + \left(y^*_{S,n} - 1\right)\left(\lambda^k_{n-1} - \lambda^k_n\right) \quad \text{(F8)}$$

and

$$L_R\left(y^*_{S,n}\right) = f(\mathbf{z}^k) + \left(1 - y^*_{S,n}\right)\left(\lambda^k_n - \lambda^k_{n+1}\right), \quad \text{(F9)}$$

respectively, and the average gradient can be derived:

$$\frac{\partial f}{\partial N^k_S} = \frac{1}{2}\left(\left(\frac{\partial f}{\partial N^k_S}\right)_L + \left(\frac{\partial f}{\partial N^k_S}\right)_R\right) = \frac{1}{2}\left(\left(\lambda^k_{n-1} - \lambda^k_n\right) + \left(\lambda^k_n - \lambda^k_{n+1}\right)\right) = \frac{1}{2}\left(\lambda^k_{n-1} - \lambda^k_{n+1}\right) \quad \text{(F10)}$$

For $N^k_S = N_{max} - 9$ and $N^k_S = N_{max}$, the Lagrange functions with variables with the left and right neighboring indices are not existing. Therefore, the gradient of the existing Lagrange function is used:

$$\frac{\partial f}{\partial N^k_S} = \left(\frac{\partial f}{\partial N^k_S}\right)_R = \left(\lambda^k_{N\,max-9} - \lambda^k_{N\,max-8}\right) \quad \text{(F11)}$$

and

$$\frac{\partial f}{\partial N^k_S} = \left(\frac{\partial f}{\partial N^k_S}\right)_L = \left(\lambda^k_{N\,max-1} - \lambda^k_{N\,max}\right), \quad \text{(F12)}$$

respectively. By substituting of Eqs. F11, F12, and F13 in Eq. F1 the following expression for the heuristic Lagrange underestimator can be derived:

$$L_{heu}(\mathbf{y}_S) = f(\mathbf{z}^k) + \begin{pmatrix} \left(\lambda^k_{N^k_S} - \lambda^k_{N^k_S+1}\right) & \text{if } N^k_S = N_{max} - 9 \\ \frac{1}{2}\left(\lambda^k_{N^k_S-1} - \lambda^k_{N^k_S+1}\right) & \text{if } N_{max} - 8 \le N^k_S \le N_{max} - 1 \\ \left(\lambda^k_{N^k_S-1} - \lambda^k_{N^k_S}\right) & \text{if } N^k_S = N_{max} \end{pmatrix} (N_S - N^k_S), \quad \text{(F13)}$$

whereas N^k_S and N_S can be substituted by the corresponding expressions with $\alpha^k_{S,n}$ and $y_{S,n}$, respectively (see Eq. 9).

Appendix G: Iteration Progress of Modified GBD Algorithm

Table A1. Flow Rates, Referred to the Feed Flow Rate, and Mole Fractions of Internal Streams for Each Outer Iteration Step (Sequence 16)

Outer Iteration	Stream	F1	D1	B1	B2	R2
Initial	Mole flow n/n_0	0.59	0.26	0.33	0.26	0.00
	x_M	0.090	0.201	0.000	**0.176**	0.000
	x_P	0.312	**0.686**	**0.010**	0.707	0.337
	x_O	0.598	0.113	0.990	0.117	**0.663**
NLP 1 (9.01 %)	Mole flow n/n_0	0.73	0.35	0.38	0.34	0.05
	x_M	0.076	0.158	0.000	**0.138**	0.000
	x_P	0.316	**0.599**	**0.054**	0.613	0.337
	x_O	0.608	0.243	0.946	0.249	**0.663**
NLP 2 (6.12 %)	Mole flow n/n_0	0.72	0.34	0.39	0.33	0.06
	x_M	0.066	0.143	0.000	**0.122**	0.000
	x_P	0.319	**0.616**	**0.061**	0.631	0.337
	x_O	0.615	0.241	0.939	0.247	**0.663**
NLP 3 (0.23 %)	Mole flow n/n_0	0.72	0.32	0.40	0.32	0.07
	x_M	0.065	0.144	0.000	**0.122**	0.000
	x_P	0.319	**0.627**	**0.068**	0.643	0.337
	x_O	0.616	0.228	0.932	0.234	**0.663**
NLP 4 (0.08 %)	Mole flow n/n_0	0.72	0.32	0.40	0.31	0.08
	x_M	0.064	0.144	0.000	**0.122**	0.000
	x_P	0.319	**0.634**	**0.072**	0.650	0.337
	x_O	0.617	0.222	0.928	0.228	**0.663**
NLP 5 (0.03 %)	Mole flow n/n_0	0.72	0.31	0.41	0.31	0.08
	x_M	0.064	0.147	0.000	**0.124**	0.000
	x_P	0.319	**0.637**	**0.075**	0.654	0.337
	x_O	0.617	0.216	0.925	0.222	**0.663**
NLP 6 (0.01 %)	Mole flow n/n_0	0.72	0.31	0.41	0.30	0.08
	x_M	0.063	0.146	0.000	**0.124**	0.000
	x_P	0.320	**0.641**	**0.075**	0.658	0.337
	x_O	0.617	0.212	0.925	0.218	**0.663**
NLP 7 (0.005 %)	Mole flow n/n_0	0.72	0.31	0.41	0.30	0.08
	x_M	0.063	0.147	0.000	**0.124**	0.000
	x_P	0.320	**0.643**	**0.077**	0.661	0.337
	x_O	0.617	0.209	0.923	0.215	**0.663**

Percentage value in the bracket denotes the improvement compared to the previous iteration. The optimization variables (hand-over purities) are in bold.

Manuscript received Jan. 19, 2008, and revision received June 15, 2008.