Monotonic Reformulation and Bound Tightening for Global Optimization of Ideal Multi-Component Distillation Columns

**Nick Mertens · Christian Kunde ·
Achim Kienle · Dennis Michaels**

Abstract This paper addresses the problem of determining cost-minimal process designs for ideal multi-component distillation columns. The special case of binary distillation was considered in the former work [2]. Therein, a problem-specific bound-tightening strategy based on monotonic mole fraction profiles of single components was developed to solve the corresponding MINLPs (mixed-integer nonlinear problems) globally. In the multi-component setting, the mole fraction profiles of single components may not be monotonic. Therefore the bound-tightening strategy from the binary case cannot be applied directly. In this follow-up paper, a model reformulation for ideal multi-component distillation columns is presented. The reformulation is achieved by suitable aggregations of the involved components. Proofs are given showing that mole fraction profiles of aggregated components are monotonic. This property is then used to adapt the bound-tightening strategy from the binary case to the proposed model reformulation. Computational results are provided that indicate the usefulness of both the model reformulation and the adapted bound tighten-

N. Mertens
Department of Mathematics, TU Dortmund University, Germany
E-mail: nick.mertens@tu-dortmund.de

C. Kunde
Institute for Automation Engineering, Otto von Guericke Unversity Magdeburg, Magdeburg, Germany
E-mail: christian.kunde@ovgu.de

A. Kienle
Institute for Automation Engineering, Otto von Guericke Unversity Magdeburg, Magdeburg, Germany
Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany
E-mail: achim.kienle@ovgu.de

D. Michaels
Department of Mathematics, TU Dortmund University, Germany
E-mail: dennis.michaels@tu-dortmund.de

ing technique for deterministic global optimization of ideal multi-component distillation column designs.

1 Introduction

The focus of this paper is on deterministic global mixed-integer nonlinear optimization of distillation columns, often dominating the cost of chemical production processes. Despite recent progress in deterministic global optimization (e.g., see [5] and [7] for recent surveys), those problems are often very difficult to solve due to high computational effort that is caused by non-convexities and a large number of variables and equations. Deterministic global optimization of distillation processes is, hence, a very challenging task.

In the literature, rigorous deterministic global optimization of distillation processes is not well-covered. Instead, restrictive model assumptions are used to get so-called short-cut models that are easier to solve. In a recent publication [19], optimal sequencing of multi-component distillation columns at minimum reflux, i.e., smallest feasible internal vapor und liquid flows, using short-cut models is calculated with deterministic global optimization methods.

By dropping simplifying assumptions such as minimum reflux, tray-to-tray distillation models offer a wider range of validity compared to short-cut models. The increased computational effort is significantly reduced by applying problem-specific global optimization strategies as demonstrated for ideal binary mixtures in our previous work [2]. Therein, it has been shown that computational cost could be reduced by orders of magnitude using a specific bound tightening strategy based on monotonicity of molar fractions, i.e., fractions of mole numbers of individual components relative to the total mole number of all components, throughout the column. The application was demonstrated for a hybrid process combining distillation and crystallization units for the separation of two isomers that are difficult to separate by distillation alone. However, an extension to multi-component mixtures is non-trivial because molar fraction profiles associated with certain components typically show non-monotonic behavior.

In this follow-up paper, a reformulation of a distillation column model is presented to overcome this problem. The reformulation is obtained by aggregation of single components, resulting in a linear transformation of the variables used for molar fractions. For mixtures with ideal liquid and vapor behavior, we prove that the transformed variables show the desired monotonic behavior, which allows us to extend the bound tightening strategy from the binary case to the multi-component setting.

Bound tightening, also known as domain reduction or range reduction, is a common strategy in global optimization to reduce the initial domain of the problem variables without cutting off the optimal solutions. This results in two

benefits in view of deterministic global optimization. First of all, a reduction of the search space is achieved. But most importantly, available global optimization software often relies on the use of convex relaxations for nonlinear mixed-integer optimization problems. In general, such convex relaxations can be tightened when the underlying domains are reduced.

In the literature, it is mainly distinguished between two basic types of domain reduction. Feasibility based bound tightening (FBBT) cuts off infeasible points using the constraints of the underlying problem. Optimization based bound tightening (OBBT) applies optimization techniques in order to derive tighter variable bounds. Both types are integrated in many state-of-the-art global optimization software packages (e.g., see BARON [22], COUENNE [6], SCIP [26], and GloMIQO [18]).

For FBBT, standard methods are often based on interval arithmetic (e.g., see [21]) and the description of nonlinearities using expression trees (e.g., see [23,25]). Bounds on the variables can be propagated onto the nonlinear expressions via forward propagation. Also, the other way around, tighter bounds on the variables can be computed using the bounds on the nonlinearities (backward propagation). This procedure can be iterated until no further strengthening of the bounds is achieved. The iteration can be formally interpreted as an operator which is shown in [4] to have a limit point. For special problems, this point can be determined using polynomial time algorithms. In [3], the concept of FBBT is expanded on convex combinations of two linear constraints. In [9], additional results on constraint satisfaction problems with quadratic constraints are presented.

The key idea of OBBT is to consecutively minimize and maximize each variable appearing in the problem on the feasible set. This is in general as hard as finding the optimal solution of the problem itself, so a common approach is to only use (linear) relaxations of the feasible set. This procedure can be iterated multiple times in order to further tighten the bounds. OBBT can be more effective, but is often much more time-consuming. It is, hence, used very rarely or only at the root node of the Branch-and-Bound tree. For more details, we refer to work [11]. An early reference in which OBBT is applied in the context of heat exchanger networks is [20]. In [8], a so-called nonlinearities removal domain reduction is introduced by fixing the value of a variable whose domain is intended to be tightened. Under some assumptions, the equivalence of this approach to standard procedures is shown, and it is concluded that general domain reduction techniques are independent in the ordering of the considered variables. In [11], Lagrangian variable bound constraints are described, making it easier to propagate bounds throughout the branching tree. Additionally, enhancements on the computation of OBBT are given.

According to this terminology, the bound tightening strategy that is developed in this paper for ideal multi-component distillation column models belongs to FBBT. We especially make use of interval arithmetics that we apply to two types of well-structured model constraints. Preliminary results of this work have already been presented in a conference paper [17]. The additional

content in the present publication comprises three points. First of all, a more detailed description and interpretation of the model and the reformulation is given. Furthermore, we provide mathematical background and proofs regarding the monotonicity property and the domain reduction strategies. Finally we present a computational study including a diverse set of test instances and solver selections.

The remainder of the paper is structured as follows. In Section 2, we present the ideal multi-component distillation column model we are working with. In Section 3, we derive an alternative model formulation by introducing suitable *aggregated components*. In Section 4, we prove that the transformed variables associated with each newly introduced aggregated component fulfill the desired monotonicity property. This is exploited in Section 5 in order to extend the bound tightening strategy from the binary case [2] to the general ideal multi-component case. In Section 6, we solve several numerical test examples to global optimality and experimentally analyze the influence of the developed techniques on the running time.

2 Distillation Column Model

In this section, the model under consideration is presented. As in previous work [2,15], we focus on a tray-by-tray model of a distillation column in steady state, and assume ideal liquid and vapor phase, total condenser and total reboiler, single liquid feed flow at boiling temperature, and constant molar overflow. Notation and model description are basically taken from [2,15] and adapted, if necessary. For a general introduction to the topic of thermal separation processes, we refer to [16]. Table 1 displays the name and meaning of all parameters, variables and indices that we are using in this work.

We are given a mixture consisting of n single components labeled by $1, \ldots, n$. The order of the components is defined with respect to the boiling point. Here, component 1 is the component with the lowest boiling point and component n refers to the component with the highest boiling point. In the model, the composition of a mixture is given in terms of molar fractions. Thus, the sum of molar fractions over all components is equal to one at every position of the column, which is known as the *summation conditions*.

Table 1 displays the name and meaning of all used parameters, variables and indices. A sketch of a distillation column is shown in Figure 1. The mixture enters the column at the feed tray with molar feed flow F and initial composition x_i^{in}, $i = 1, \ldots, n$. At the top tray (condenser), the distillate molar flow D leaves the column with composition x_i^{dist}, $i = 1, \ldots, n$, and at the bottom tray (reboiler), the molar flow B leaves the column with composition x_i^{bot}, $i = 1, \ldots, n$. V denotes the vapor flow that streams upwards through the column. The *overall mass balance equations*

$$Fx_i^{\text{in}} = Dx_i^{\text{dist}} + Bx_i^{\text{bot}}, \quad i = 1, \ldots, n, \tag{1}$$

Parameters:	
u	Upper bound on the number of trays (length)
α_i	Constant relative volatility of component i
σ	Position of the split between less and more volatile components
π	Purity requirement on the components
F	Feed molar flow entering the column, given in $\mathrm{mol\,s}^{-1}$

Variables:	
V	Vapor flow streaming upwards through the column, given in $\mathrm{mol\,s}^{-1}$
D	Distillate molar flow withdrawn at the condenser, given in $\mathrm{mol\,s}^{-1}$
B	Bottom molar flow withdrawn at the reboiler, given in $\mathrm{mol\,s}^{-1}$
ν_{r}	Ratio of downward flow to upward flow in the rectifying section
ν_{s}	Ratio of upward flow to downward flow in the stripping section
l	Number of trays in the part of the column specified by superscript
β	Binary coupling variable determining the position of the feed tray
x_i/y_i	Liquid/Vapor molar fraction of component i
X_k/Y_k	Liquid/Vapor molar fraction of the aggregated component k

Indices:	
in	Feed flow
dist	Distillate flow
bot	Bottom flow
feed	Feed tray
feed-1	Tray above the feed tray
feed$+1$	Tray below the feed tray
col	Whole column
rect	Rectifying section
l_{r}	Tray number in rectifying section
strip	Stripping section
l_{s}	Tray number in stripping section

Table 1: List of parameters, variables and indices used in our model

ensure that the amount of component i entering the column coincides with the overall amount of component i leaving the column.

Rectifying section (above the feed tray) and stripping section (below the feed tray) can contain several trays. Trays of the rectifying section are numbered from the top to the bottom by $l_{\mathrm{r}} = 1, \ldots, l^{\mathrm{rect}}$, and trays in the stripping section are numbered from the bottom to the top by $l_{\mathrm{s}} = 1, \ldots, l^{\mathrm{strip}}$. Variables used for molar fractions of component i in liquid and in vapor phases are denoted by x_i and y_i, respectively. In order to specify the tray a variable is
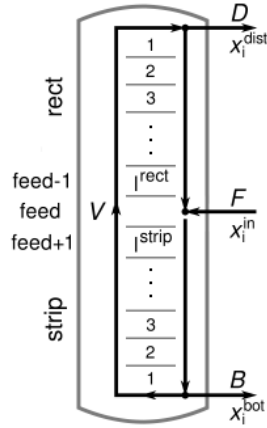
Fig. 1: A fixed design of a distillation column. Numbers of trays in the rectifying and stripping sections are l^{rect} and l^{strip}. The total number of trays in the column is given by $l^{\text{col}} = l^{\text{rect}} + l^{\text{strip}} + 1$.

associated with, we introduce superscripts "feed", "feed-1", "feed+1", "rect" and "strip". The script "feed-1" denotes the tray above the feed tray and the script "feed+1" denotes the tray below the feed tray. Trays from the rectifying and stripping sections are additionally equipped with their associated tray number as subscript. Mass balances comprising the first tray and a number of consecutive trays are established for the stripping and the rectifying section as well as a mass balance comprising the feed tray. The mass transfer in liquid and vapor phase through the column is then described by the *component mass balance equations*

$$
\begin{aligned}
y_{i,l_{\mathrm{r}}+1}^{\text{rect}} &= \nu_{\mathrm{r}}\, x_{i,l_{\mathrm{r}}}^{\text{rect}} + (1 - \nu_{\mathrm{r}})\, y_{i,1}^{\text{rect}}, \\
\nu_{\mathrm{s}}\, y_i^{\text{feed}} + x_i^{\text{feed}} &= \nu_{\mathrm{s}}\, y_i^{\text{feed+1}} + \nu_{\mathrm{r}}\, \nu_{\mathrm{s}}\, x_i^{\text{feed-1}} + (1 - \nu_{\mathrm{r}}\, \nu_{\mathrm{s}})\, x_i^{\text{in}}, \\
x_{i,l_{\mathrm{s}}+1}^{\text{strip}} &= \nu_{\mathrm{s}}\, y_{i,l_{\mathrm{s}}}^{\text{strip}} + (1 - \nu_{\mathrm{s}})\, x_{i,1}^{\text{strip}},
\end{aligned}
\tag{2}
$$

for $i = 1, \ldots, n$, $l_{\mathrm{r}} = 1, \ldots, u^{\text{rect}}$ and $l_{\mathrm{s}} = 1, \ldots, u^{\text{strip}}$, where u^{rect} and u^{strip} denote upper bounds imposed on l^{rect} and l^{strip}, respectively. We remark that in the equations (2) the subscripts indicating trays formally range to $u^{\text{rect}} + 1$ and $u^{\text{strip}} + 1$, respectively. This way, two artificial trays are introduced to the model. These two trays are later used to model the coupling of the feed tray with the rectifying and the stripping sections (see equations (6)). The auxiliary variables $\nu_{\mathrm{r}}, \nu_{\mathrm{s}} \in [0,1]$ defined as

$$
\nu_{\mathrm{r}} = \frac{V - D}{V} \quad \text{and} \quad \nu_{\mathrm{s}} = \frac{V}{V + B}
\tag{3}
$$

describe the ratio of upward and downward molar flows in the rectifying and stripping section.

The separation behavior of component i is given by its volatility. Components with higher volatility accumulate in the vapor phase, while components with lower volatility accumulate in the liquid phase. We assume constant relative volatilities of the components, expressed by parameters $\alpha_i > 0$, for $i = 1, \ldots, n$. Due to our assumption on the order of the components, we have that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$. At all trays, the interactions of the mole fractions in the vapor phase and in the liquid phase are given by the *phase equilibrium equations*

$$y_{i,l_r}^{\text{rect}} = \frac{\alpha_i x_{i,l_r}^{\text{rect}}}{\sum_{j=1}^n \alpha_j x_{j,l_r}^{\text{rect}}}, \; y_i^{\text{feed}} = \frac{\alpha_i x_i^{\text{feed}}}{\sum_{j=1}^n \alpha_j x_j^{\text{feed}}}, \; y_{i,l_s}^{\text{strip}} = \frac{\alpha_i x_{i,l_s}^{\text{strip}}}{\sum_{j=1}^n \alpha_j x_{j,l_s}^{\text{strip}}} \quad (4)$$

for $i = 1, \ldots, n$, $l_r = 1, \ldots, u^{\text{rect}} + 1$ and $l_s = 1, \ldots, u^{\text{strip}} + 1$. Note that the index j is used to indicate the respective component in the sums.

The vapor flow at the top of the column is completely condensed and the liquid flow fed back to the column at the bottom is completely vaporized. *Total condenser* and *total reboiler* are modeled by

$$x_i^{\text{dist}} = y_{i,1}^{\text{rect}} \quad \text{and} \quad x_i^{\text{bot}} = x_{i,1}^{\text{strip}}, \qquad i = 1, \ldots, n. \tag{5}$$

The total number l^{col} of trays used in a distillation column is given by the number of trays used in the rectifying section, the number of trays used in the stripping section, and the feed tray. To specify l^{col} in our model, the following *coupling conditions* are imposed.

$$x_i^{\text{feed-1}} = \sum_{l_r=1}^{u^{\text{rect}}} \beta_{l_r}^{\text{rect}} x_{i,l_r}^{\text{rect}}, \qquad x_i^{\text{feed}} = \sum_{l_r=1}^{u^{\text{rect}}} \beta_{l_r}^{\text{rect}} x_{i,l_r+1}^{\text{rect}}, \quad i = 1, \ldots, n,$$

$$y_i^{\text{feed+1}} = \sum_{l_s=1}^{u^{\text{strip}}} \beta_{l_s}^{\text{strip}} y_{i,l_s}^{\text{strip}}, \quad x_i^{\text{feed}} = \sum_{l_s=1}^{u^{\text{strip}}} \beta_{l_s}^{\text{strip}} x_{i,l_s+1}^{\text{strip}}, \quad i = 1, \ldots, n,$$

$$(6a)$$

$$l^{\text{col}} = \sum_{l_r=1}^{u^{\text{rect}}} \beta_{l_r}^{\text{rect}} l_r + \sum_{l_s=1}^{u^{\text{strip}}} \beta_{l_s}^{\text{strip}} l_s + 1,$$

$$\sum_{l_r=1}^{u^{\text{rect}}} \beta_{l_r}^{\text{rect}} = 1, \quad \beta_{l_r}^{\text{rect}} \in \{0,1\}, \; l_r = 1, \ldots, u^{\text{rect}}, \tag{6b}$$

$$\sum_{l_s=1}^{u^{\text{strip}}} \beta_{l_s}^{\text{strip}} = 1, \quad \beta_{l_s}^{\text{strip}} \in \{0,1\}, \; l_s = 1, \ldots, u^{\text{strip}}.$$

Note that the binary variables $\beta_{l_r}^{\text{rect}}$ and $\beta_{l_s}^{\text{strip}}$ attain value one if and only if tray l_r of the rectifying section and tray l_s of the stripping section are chosen to be the trays above and below the feed tray in the column.

The purpose of the distillation column is to separate the more volatile components from the less volatile components under given purity constraints.

Let the predefined *split* parameter $\sigma \in \{1, \ldots, n-1\}$ be the index such that the components 1 to σ belong to the more volatile part and components $\sigma + 1$ to n belong to the less volatile part of the mixture, and let $\pi^{\mathrm{dist}}, \pi^{\mathrm{bot}} \in [0, 1]$ denote the purity requirements imposed on the more volatile components at the condenser and on the less volatile components at the reboiler. Then, the *purity constraints* are given as follows:

$$\sum_{i=1}^{\sigma} x_i^{\mathrm{dist}} \geq \pi^{\mathrm{dist}} \quad \text{and} \quad \sum_{i=\sigma+1}^{n} x_i^{\mathrm{bot}} \geq \pi^{\mathrm{bot}}. \tag{7}$$

The objective function of our column model reflects the total annualized cost of the distillation process that needs to be minimized. In our work, we make use of the following cost function that is taken from previous work [15].

$$\begin{aligned}
\mathrm{cost} \; = \; &\lambda_1 V \; + \; \lambda_2 l^{\mathrm{col}} \; + \; \lambda_3 (\lambda_4 V \; + \; \lambda_5 B) l^{\mathrm{col}} \\
&+ \; \lambda_6 (\lambda_4 V \; + \; \lambda_5 B)^{\gamma_1} (\lambda_7 l^{\mathrm{col}} \; + \; \lambda_8)^{\gamma_2} \; + \; \lambda_9 (\lambda_4 V \; + \; \lambda_5 B)^{\gamma_3} l^{\mathrm{col}},
\end{aligned} \tag{8}$$

with coefficients as specified in Table 2. This objective function was originally developed for the distillation of dodecanal and 2-methylundecanal. However, its structure is typical for economical cost estimation and therefore suitable for the computational studies in our work.

Coefficent	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
Value	17544	173.6	2009.7	0.2378	0.0221	2364.5
Unit	€a^{-1}mol^{-1}s	€a^{-1}	€a^{-1}	mol^{-1}s	mol^{-1}s	€a^{-1}
Coefficient	λ_7	λ_8	λ_9	γ_1	γ_2	γ_3
Value	0.2	4	-171.4	0.533	0.5	0.82
Unit			€a^{-1}			

Table 2: Coefficients for the cost function

The resulting mixed-integer nonlinear optimization problem is given by

$$\min (8) \quad \text{s.t. } (1) - (7). \tag{9}$$

3 Model Reformulation by Aggregating Components

In the binary case, mole fractions associated with a single component behave monotonic through the distillation column. Based on that property, a problem-specific bound tightening strategy for binary distillation column design problems has been developed in our previous work [2].

To formalize the notion of monotonicity, we consider the distillation column tray by tray from the bottom to the top. We say that a component i shows

monotonic behaviour through the distillation column when the sequence of respective molar fractions in the liquid phase are either non-decreasing or non-inreasing, i. e. either

$$x_{i,1}^{\mathrm{strip}} \leq \cdots \leq x_{i,l^{\mathrm{strip}}}^{\mathrm{strip}} \leq x_i^{\mathrm{feed}} \leq x_{i,l^{\mathrm{rect}}}^{\mathrm{rect}} \leq \cdots \leq x_{i,1}^{\mathrm{rect}}$$

or

$$x_{i,1}^{\mathrm{strip}} \geq \cdots \geq x_{i,l^{\mathrm{strip}}}^{\mathrm{strip}} \geq x_i^{\mathrm{feed}} \geq x_{i,l^{\mathrm{rect}}}^{\mathrm{rect}} \geq \cdots \geq x_{i,1}^{\mathrm{rect}}$$

holds. A sequence of values associated with liquid phase (or vapor phase) mole fractions of a single component, considered from the bottom to the top, is also refered to as a *mole fraction profile*.

However, the molar fractions of a single component may not behave monotonic in the multi-component setting. Such a typical situation is illustrated in Figure 2 (a) for component 2 (blue-colored dashed curve) and component 3 (yellow-colored dotted curve). This fact makes it hard to generalize the bound
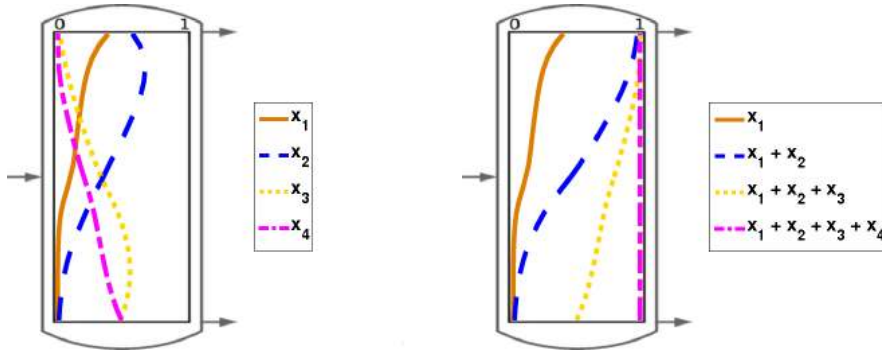
(a) Molar fraction profiles of single components (original model formulation)

(b) Profiles of transformed variables for aggregated components (aggregated model formulation)

Fig. 2: The figure shows molar fraction profiles and profiles of transformed variables in liquid phase for a four component mixture.

tightening strategy for the binary case [2] to the multi-component case, directly.

We overcome this problem by aggregating, for $k = 1, \ldots, n$, the first k components. This is achieved by summing up the corresponding variables used for molar fractions at every position of the distillation column.

The driving force for separating the components of a mixture using distillation are differences in the volatilities of the single components. For components sorted by decreasing volatility, any group of the first k components has an effective volatility larger than that of the complementary group of $n - k$ components. Therefore, the same direction of the driving force and thus monotonicity of the aggregated molar fractions is expected over the whole column.

To be more precise, let x_i and y_i (for $i = 1, \ldots, n$) be the variables used for molar fractions of component i in the liquid and in the vapor phase at an arbitrary position. The associated sub- and superscripts indicating the specific position are omitted for a clean presentation. We label the *aggregated components* by $k = 1, \ldots, n$ and introduce *aggregated concentration variables* X_k and Y_k for the liquid and vapor phases. The original variables used for molar fractions and the new aggregated concentration variables are linearly linked to each other by the following bijective relations.

$$X_k = \sum_{i=1}^{k} x_i \quad \text{and} \quad Y_k = \sum_{i=1}^{k} y_i, \quad k = 1, \ldots, n. \tag{10}$$

We also introduce an (aggregated) *dummy component* 0 for which the liquid and vapor phase concentrations X_0 and Y_0 are zero at all positions. Hence we can formulate the inverse to the relations (10) by

$$x_i = X_i - X_{i-1}, \quad \text{and} \quad y_i = Y_i - Y_{i-1}, \quad k = 1, \ldots, n. \tag{11}$$

By definition and due to the summation conditions (i.e. $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i = 1$), we have

$$0 = X_0 \leq X_1 \leq \cdots \leq X_n = 1 \quad \text{and} \quad 0 = Y_0 \leq Y_1 \leq \cdots \leq Y_n = 1. \tag{12}$$

Note that the equations (3), (6b) and the objective function (8) do not depend on the molar fractions. Hence, they remain unchanged in our aggregated model formulation. Observe further that the overall mass balance equations (1), the component mass balance equations (2), the coupling conditions (6a), total condenser and total reboiler conditions (5) and the purity constraints (7) are linear in x and y while using the same coefficients for every $i = 1, \ldots, n$ respectively. Therefore, we obtain the corresponding constraints for each aggregated component k by summing up the corresponding conditions associated with the first k single components. Only the phase equilibrium equations (4) are not linear in the original concentration variables and need to be adapted by applying the inverse relations (11). The aggregated model formulation reads as

− The *aggregated overall mass balance equations*:

$$F X_k^{\mathrm{in}} = D X_k^{\mathrm{dist}} + B X_k^{\mathrm{bot}}, \quad k = 1, \ldots, n. \tag{13}$$

− The *aggregated component mass balance equations*:

$$
\begin{aligned}
Y_{k,l_{\mathrm{r}}+1}^{\mathrm{rect}} &= \nu_{\mathrm{r}} X_{k,l_{\mathrm{r}}}^{\mathrm{rect}} + (1 - \nu_{\mathrm{r}}) Y_{k,1}^{\mathrm{rect}}, \\
\nu_{\mathrm{s}} Y_k^{\mathrm{feed}} + X_k^{\mathrm{feed}} &= \nu_{\mathrm{s}} Y_k^{\mathrm{feed}+1} + \nu_{\mathrm{r}}\, \nu_{\mathrm{s}}\, X_k^{\mathrm{feed}\text{-}1} + (1 - \nu_{\mathrm{r}}\, \nu_{\mathrm{s}})\, X_k^{\mathrm{in}}, \\
X_{k,l_{\mathrm{s}}+1}^{\mathrm{strip}} &= \nu_{\mathrm{s}} Y_{k,l_{\mathrm{s}}}^{\mathrm{strip}} + (1 - \nu_{\mathrm{s}})\, X_{k,1}^{\mathrm{strip}}
\end{aligned}
\tag{14}
$$

for $k = 1, \ldots, n$, $l_{\mathrm{r}} = 1, \ldots, u^{\mathrm{rect}}$ and $l_{\mathrm{s}} = 1, \ldots, u^{\mathrm{strip}}$.

– The *auxiliary variables constraints*:

$$\nu_{\mathrm{r}} = \frac{V - D}{V} \quad \text{and} \quad \nu_{\mathrm{s}} = \frac{V}{V + B}. \tag{15}$$

– The *aggregated phase equilibrium equations*:

$$\begin{aligned}
Y_{k,l_{\mathrm{r}}}^{\mathrm{rect}} &= \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{r}}}^{\mathrm{rect}} - X_{j-1,l_{\mathrm{r}}}^{\mathrm{rect}})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{r}}}^{\mathrm{rect}} - X_{j-1,l_{\mathrm{r}}}^{\mathrm{rect}})}, \\
Y_{k}^{\mathrm{feed}} &= \frac{\sum_{j=1}^{k} \alpha_j (X_{j}^{\mathrm{feed}} - X_{j-1}^{\mathrm{feed}})}{\sum_{j=1}^{n} \alpha_j (X_{j}^{\mathrm{feed}} - X_{j-1}^{\mathrm{feed}})}, \\
Y_{k,l_{\mathrm{s}}}^{\mathrm{strip}} &= \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}}^{\mathrm{strip}} - X_{j-1,l_{\mathrm{s}}}^{\mathrm{strip}})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}}^{\mathrm{strip}} - X_{j-1,l_{\mathrm{s}}}^{\mathrm{strip}})}
\end{aligned} \tag{16}$$

for $k = 1, \ldots, n$, $l_{\mathrm{r}} = 1, \ldots, u^{\mathrm{rect}} + 1$ and $l_{\mathrm{s}} = 1, \ldots, u^{\mathrm{strip}} + 1$. Note that in this case also the index j is used to indicate the respective aggregated components in the sums.

– The *aggregated* constraints for the *total condenser/reboiler*:

$$X_{k}^{\mathrm{dist}} = Y_{k,1}^{\mathrm{rect}} \quad \text{and} \quad X_{k}^{\mathrm{bot}} = X_{k,1}^{\mathrm{strip}}, \qquad k = 1, \ldots, n. \tag{17}$$

– The *aggregated coupling conditions*:

$$\begin{aligned}
X_{k}^{\mathrm{feed\text{-}1}} &= \sum_{l_{\mathrm{r}}=1}^{u^{\mathrm{rect}}} \beta_{l_{\mathrm{r}}}^{\mathrm{rect}} X_{k,l_{\mathrm{r}}}^{\mathrm{rect}}, \qquad k = 1, \ldots, n, \\
X_{k}^{\mathrm{feed}} &= \sum_{l_{\mathrm{r}}=1}^{u^{\mathrm{rect}}} \beta_{l_{\mathrm{r}}}^{\mathrm{rect}} X_{k,l_{\mathrm{r}}+1}^{\mathrm{rect}}, \qquad k = 1, \ldots, n, \\
Y_{k}^{\mathrm{feed}+1} &= \sum_{l_{\mathrm{s}}=1}^{u^{\mathrm{strip}}} \beta_{l_{\mathrm{s}}}^{\mathrm{strip}} Y_{k,l_{\mathrm{s}}}^{\mathrm{strip}}, \qquad k = 1, \ldots, n, \\
X_{k}^{\mathrm{feed}} &= \sum_{l_{\mathrm{s}}=1}^{u^{\mathrm{strip}}} \beta_{l_{\mathrm{s}}}^{\mathrm{strip}} X_{k,l_{\mathrm{s}}+1}^{\mathrm{strip}}, \qquad k = 1, \ldots, n,
\end{aligned} \tag{18a}$$

$$l^{\mathrm{col}} = \sum_{l_{\mathrm{r}}=1}^{u^{\mathrm{rect}}} \beta_{l_{\mathrm{r}}}^{\mathrm{rect}} l_{\mathrm{r}} + \sum_{l_{\mathrm{s}}=1}^{u^{\mathrm{strip}}} \beta_{l_{\mathrm{s}}}^{\mathrm{strip}} l_{\mathrm{s}} + 1,$$

$$\sum_{l_{\mathrm{r}}=1}^{u^{\mathrm{rect}}} \beta_{l_{\mathrm{r}}}^{\mathrm{rect}} = 1, \quad \beta_{l_{\mathrm{r}}}^{\mathrm{rect}} \in \{0, 1\}, \ l_{\mathrm{r}} = 1, \ldots, u^{\mathrm{rect}}, \tag{18b}$$

$$\sum_{l_{\mathrm{s}}=1}^{u^{\mathrm{strip}}} \beta_{l_{\mathrm{s}}}^{\mathrm{strip}} = 1, \quad \beta_{l_{\mathrm{s}}}^{\mathrm{strip}} \in \{0, 1\}, \ l_{\mathrm{s}} = 1, \ldots, u^{\mathrm{strip}}.$$

– The *aggregated purity constraints*:

$$X_{\sigma}^{\mathrm{dist}} \geq \pi^{\mathrm{dist}} \quad \text{and} \quad (1 - X_{\sigma}^{\mathrm{bot}}) \geq \pi^{\mathrm{bot}}. \tag{19}$$

− The *objective function*:

$$
\begin{aligned}
\text{cost} \; = \; & \lambda_1 V \; + \; \lambda_2 l^{\mathrm{col}} \; + \; \lambda_3 (\lambda_4 V \; + \; \lambda_5 B) l^{\mathrm{col}} \\
& + \; \lambda_6 (\lambda_4 V \; + \; \lambda_5 B)^{\gamma_1} \, (\lambda_7 l^{\mathrm{col}} \; + \; \lambda_8)^{\gamma_2} \\
& + \; \lambda_9 (\lambda_4 V \; + \; \lambda_5 B)^{\gamma_3} l^{\mathrm{col}},
\end{aligned}
\tag{20}
$$

with coefficients as specified in Table 2.

The resulting reformulated mixed-integer nonlinear optimization problem is given by

$$
\min (20) \quad \text{s.t. } (13) - (19).
\tag{21}
$$

It turns out that the concentration variables of each aggregated component show the desired monotonic behavior, i.e., the overall molar fraction of all components above each possible split position $\sigma \in \{1, \ldots, n-1\}$ change monotonically throughout the distillation column. This is illustrated in Figure 2 (b), and will be proven in Section 4.

4 Monotonicity of the Aggregated Concentration Profiles

In this section we prove that for each aggregated component, the corresponding concentration variables introduced in Section 3 behave monotonic through the distillation column. We refer to a sequence of liquid or vapor phase concentration values of an aggregated component as a *(concentration) profile*.

In what follows, we investigate the restrictions of each such profile to the stripping section and to the rectifying section separately. Subsection 4.1 deals with the stripping section. We show that each profile is non-decreasing when considered from the bottom to the top. For the rectifying section discussed in Subsection 4.2, we first apply a suitable transformation. That transformation traces the profiles restricted to the rectifying section back to the case of profiles restricted a stripping section. We then conclude that each profile also behaves non-decreasing in the rectifying section from bottom to the top.

As the coupling conditions (18) ensure that, for each profile, the parts restricted to the stripping section and restricted to the rectifying section must coincide at the feed tray, we finally obtain that each profile of an aggregated component runs monotonically through the whole distillation column.

4.1 Stripping Section

We omit superscript "strip" and denote by $\boldsymbol{X} := (X_{k,l_{\mathrm{s}}})_{k=0,\ldots,n,\, l_{\mathrm{s}}=1,\ldots,u+1}$ and $\boldsymbol{Y} := (Y_{k,l_{\mathrm{s}}})_{k=0,\ldots,n,\, l_{\mathrm{s}}=1,\ldots,u+1}$ the matrices consisting of all liquid and vapor phase concentration variables w.r.t. aggregated components (including the dummy one) and restricted to the stripping section.

With this notation and combining the phase equilibrium equations (16) with the component mass balance equations (14), we obtain the following

subsystem that is satisfied by every feasible solution of our distillation column model from Section 3.

$$X_{k,l_s+1} = \nu_s \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_s} - X_{j-1,l_s})} + (1-\nu_s)X_{k,1}, \quad k = 0,\dots,n,$$
$$l_s = 1,\dots,u,$$

$$0 < X_{1,l_s} \le X_{2,l_s} \le \cdots \le X_{n,l_s}, \qquad\qquad l_s = 1,\dots,u+1, \quad (22)$$

$$X_{0,l_s} = 0, \ X_{n,l_s} = 1, \qquad\qquad\qquad l_s = 1,\dots,u+1,$$
$$\boldsymbol{X} \in \mathbb{R}^{(n+1)\times(u+1)}, \quad v_s \in [0,1].$$

For our analysis, the following remarks are worth to mention.

– In System (22), we impose that all variables X_{k,l_s}, $k \ge 1$, are strictly positive. This assumption can be made with out loss of generality. Indeed, when $X_{k,l_s} = 0$ holds, for some $k \ge 1$ and some l_s, the recursive formula already implies that the concentration of component k is zero at every position in the stripping section, and, hence, in the entire column. In that case we can exclude component k from our considerations.
– To keep the notation simple, we define the following expressions to denote the denominators appearing in System (22).

$$N_{l_s}(\boldsymbol{X}) := \sum_{j=1}^{n} \alpha_j (X_{j,l_s} - X_{j-1,l_s}), \quad l_s = 1,\dots,u+1.$$

Note that $N_{l_s}(\boldsymbol{X}) > 0$ holds for all $\boldsymbol{X}$ such that there is a $\nu_s \in [0,1]$ with $(\boldsymbol{X}, \nu_s)$ being feasible to System (22).
– Finally, we observe the identities

$$\sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) = \sum_{j=1}^{k-1} X_{j,l_s}(\alpha_j - \alpha_{j+1}) + X_{k,l_s}\alpha_k.$$

for $k = 1,\dots,n$ and $l_s = 1,\dots,u+1$, that we will frequently use throughout the proofs.

For a solution $(\boldsymbol{X}, \nu_s)$ feasible to System (22), we will next show that for each aggregated component $k = 1,\dots,n$, the sequence $\{X_{k,l_s}\}_{l_s=1}^{u+1}$ is non-decreasing. More precisely, we will prove a more general statement implying the desired property.

Theorem 1 *Let $\boldsymbol{X} \in \mathbb{R}^{(n+1)\times(u+1)}$ and $\nu_s \in [0,1]$ be feasible to System (22) for some $\alpha \in \mathbb{R}^n$ with $\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_n > 0$. Then,*

$$\frac{X_{k,l_s+1} - X_{k,l_s}}{X_{k,l_s}} \ge \frac{X_{k+1,l_s+1} - X_{k+1,l_s}}{X_{k+1,l_s}} \qquad (23)$$

holds for $k = 1,\dots,n-1$ and for $l_s = 1,\dots,u$.

Proof The statement is proven by induction on l_{s}. We first consider the case with $l_{\mathrm{s}} = 1$. For an arbitrary $k \in \{1, \ldots, n-1\}$, System (22) yields

$$X_{k,2} - X_{k,1} = \nu_{\mathrm{s}} X_{k,1} \Big(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1})}{N_1(\boldsymbol{X}) \, X_{k,1}} - 1 \Big) \quad \text{and}$$

$$X_{k+1,2} - X_{k+1,1} = \nu_{\mathrm{s}} X_{k+1,1} \Big(\frac{\sum_{j=1}^{k+1} \alpha_j (X_{j,1} - X_{j-1,1})}{N_1(\boldsymbol{X}) X_{k+1,1}} - 1 \Big),$$

or equivalently

$$\frac{X_{k,2} - X_{k,1}}{X_{k,1}} = \frac{\nu_{\mathrm{s}}}{N_1(\boldsymbol{X})} \left(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1})}{X_{k,1}} \right) - \nu_{\mathrm{s}},$$

$$\frac{X_{k+1,2} - X_{k+1,1}}{X_{k+1,1}} = \frac{\nu_{\mathrm{s}}}{N_1(\boldsymbol{X})} \left(\frac{\sum_{j=1}^{k+1} \alpha_j (X_{j,1} - X_{j-1,1})}{X_{k+1,1}} \right) - \nu_{\mathrm{s}}.$$

Thus, in order to prove our statement for $l_{\mathrm{s}} = 1$, we show that

$$\frac{\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1})}{X_{k,1}} \geq \frac{\sum_{j=1}^{k+1} \alpha_j (X_{j,1} - X_{j-1,1})}{X_{k+1,1}}.$$

Observing that $X_{k+1,1} \geq X_{k,1}$ and $\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1}) - \alpha_{k+1} X_{k,1} \geq 0$ hold, yields

$$\frac{\sum_{j=1}^{k+1} \alpha_j (X_{j,1} - X_{j-1,1})}{X_{k+1,1}} = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1}) - \alpha_{k+1} X_{k,1}}{X_{k+1,1}} + \alpha_{k+1}$$

$$\leq \frac{\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1}) - \alpha_{k+1} X_{k,1}}{X_{k,1}} + \alpha_{k+1}$$

$$= \frac{\sum_{j=1}^{k} \alpha_j (X_{j,1} - X_{j-1,1})}{X_{k,1}},$$

i.e., for $l_{\mathrm{s}} = 1$, the statement holds for $k = 1, \ldots, n-1$.

Now assume that, for some $l_{\mathrm{s}} \geq 1$, the statement is true for each $k = 1, \ldots, n-1$. We will show that for $l_{\mathrm{s}} + 1$ the statement is then true for each $k = 1, \ldots, n-1$ as well. For this, we define $m_k := \frac{X_{k,l_{\mathrm{s}}}}{X_{k,l_{\mathrm{s}}-1}}$ for each $k = 1, \ldots, n$. By our induction hypothesis, we have that $m_k \geq m_{k+1}$ holds for $k = 1, \ldots, n-1$.

Next, the values of the terms $X_{k,l_{\mathrm{s}}+1} - X_{k,l_{\mathrm{s}}}$ and $X_{k+1,l_{\mathrm{s}}+1} - X_{k+1,l_{\mathrm{s}}}$ are compared. Using the recursive formula and the definition of m_k, we obtain for $X_{k,l_{\mathrm{s}}+1} - X_{k,l_{\mathrm{s}}}$ that

$$\nu_{\mathrm{s}} \left(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} - \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}-1} - X_{j-1,l_{\mathrm{s}}-1})}{N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \right) =$$

$$\nu_{\mathrm{s}} \Big(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} - \frac{\sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_{\mathrm{s}}}}{m_j} - \frac{X_{j-1,l_{\mathrm{s}}}}{m_{j-1}} \right)}{N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \Big). \tag{24}$$

Moreover, we have that

$$\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{m_1} \leq \sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_{\mathrm{s}}}}{m_j} - \frac{X_{j-1,l_{\mathrm{s}}}}{m_{j-1}} \right)$$

$$\leq \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{m_k}.$$

By the intermediate value theorem, there must exist some $\tilde{m} \in [m_k, m_1]$ with

$$\sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_{\mathrm{s}}}}{m_j} - \frac{X_{j-1,l_{\mathrm{s}}}}{m_{j-1}} \right) = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{\tilde{m}}. \qquad (25)$$

Combining formula (24) with formula (25) gives rise to

$$X_{k,l_{\mathrm{s}}+1} - X_{k,l_{\mathrm{s}}}$$
$$= \nu_{\mathrm{s}} \sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}}) \left(\frac{1}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} - \frac{1}{\tilde{m} \, N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \right) \qquad (26)$$
$$= X_{k,l_{\mathrm{s}}} \nu_{\mathrm{s}} \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{X_{k,l_{\mathrm{s}}}} \left(\frac{1}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} - \frac{1}{\tilde{m} N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \right).$$

Again, using the recursive formula and the definition of m_k, the second term can be rewritten as follows.

$$X_{k+1,l_{\mathrm{s}}+1} - X_{k+1,l_{\mathrm{s}}}$$
$$= \nu_{\mathrm{s}} \left(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}}) + \alpha_{k+1}(X_{k+1,l_{\mathrm{s}}} - X_{k,l_{\mathrm{s}}})}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} \right.$$
$$\left. - \frac{\sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_{\mathrm{s}}}}{m_j} - \frac{X_{j-1,l_{\mathrm{s}}}}{m_{j-1}} \right) + \alpha_{k+1} \left(\frac{X_{k+1,l_{\mathrm{s}}}}{m_{k+1}} - \frac{X_{k,l_{\mathrm{s}}}}{m_k} \right)}{N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \right)$$
$$= \nu_{\mathrm{s}} \left(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} - \frac{\sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_{\mathrm{s}}}}{m_j} - \frac{X_{j-1,l_{\mathrm{s}}}}{m_{j-1}} \right)}{N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \right)$$
$$+ \nu_{\mathrm{s}} \left(\frac{(\alpha_{k+1}(X_{k+1,l_{\mathrm{s}}} - X_{k,l_{\mathrm{s}}})}{N_{l_{\mathrm{s}}}(\boldsymbol{X})} - \frac{\alpha_{k+1} \left(\frac{X_{k+1,l_{\mathrm{s}}}}{m_{k+1}} - \frac{X_{k,l_{\mathrm{s}}}}{m_k} \right)}{N_{l_{\mathrm{s}}-1}(\boldsymbol{X})} \right).$$

Using formula (25) and the fact that $m_{k+1} \le m_k \le \tilde{m}$ holds, we can further estimate

$$
\begin{aligned}
&X_{k+1,l_s+1} - X_{k+1,l_s} \\
&= \nu_s \sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) \Big(\frac{1}{N_{l_s}(\boldsymbol{X})} - \frac{1}{\tilde{m} N_{l_s-1}(\boldsymbol{X})} \Big) \\
&\quad + \nu_s \Big(\frac{(\alpha_{k+1}(X_{k+1,l_s} - X_{k,l_s})}{N_{l_s}(\boldsymbol{X})} - \frac{\alpha_{k+1}\big(\frac{X_{k+1,l_s}}{m_{k+1}} - \frac{X_{k,l_s}}{m_k}\big)}{N_{l_s-1}(\boldsymbol{X})} \Big) \\
&\le \nu_s \sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) \Big(\frac{1}{N_{l_s}(\boldsymbol{X})} - \frac{1}{\tilde{m} N_{l_s-1}(\boldsymbol{X})} \Big) \\
&\quad + \nu_s \Big(\frac{(\alpha_{k+1}(X_{k+1,l_s} - X_{k,l_s})}{N_{l_s}(\boldsymbol{X})} - \frac{\alpha_{k+1}(X_{k+1,l_s} - X_{k,l_s})}{\tilde{m} N_{l_s-1}(\boldsymbol{X})} \Big) \\
&= \nu_s \sum_{j=1}^{k+1} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) \Big(\frac{1}{N_{l_s}(\boldsymbol{X})} - \frac{1}{\tilde{m}\, N_{l_s-1}(\boldsymbol{X})} \Big).
\end{aligned}
$$

Finally, we exploit that $X_{k+1,l_s} \ge X_{k,l_s}$ and that $\sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) - \alpha_{k+1} X_{k,l_s} \ge 0$ holds. This yields

$$
\begin{aligned}
&X_{k+1,l_s+1} - X_{k+1,l_s} \\
&\le X_{k+1,l_s} \nu_s \Big(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) - \alpha_{k+1} X_{k,l_s}}{X_{k+1,l_s}} + \alpha_{k+1} \Big) \\
&\quad \cdot \Big(\frac{1}{N_{l_s}(\boldsymbol{X})} - \frac{1}{\tilde{m} N_{l_s-1}(\boldsymbol{X})} \Big) \\
&\le X_{k+1,l_s} \nu_s \Big(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s}) - \alpha_{k+1} X_{k,l_s}}{X_{k,l_s}} + \alpha_{k+1} \Big) \qquad (27) \\
&\quad \cdot \Big(\frac{1}{N_{l_s}(\boldsymbol{X})} - \frac{1}{\tilde{m} N_{l_s-1}(\boldsymbol{X})} \Big) \\
&= X_{k+1,l_s} \nu_s \Big(\frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_s} - X_{j-1,l_s})}{X_{k,l_s}} \Big) \Big(\frac{1}{N_{l_s}(\boldsymbol{X})} - \frac{1}{\tilde{m}\, N_{l_s-1}(\boldsymbol{X})} \Big).
\end{aligned}
$$

From formulas (26) and (27), we can deduce that

$$
\frac{X_{k,l_s+1} - X_{k,l_s}}{X_{k,l_s}} \ge \frac{X_{k+1,l_s+1} - X_{k+1,l_s}}{X_{k+1,l_s}}
$$

holds.   □

We are now able to prove that the concentration profiles behave non-decreasing through the stripping section.

Corollary 2 *Let $(\boldsymbol{X}, \nu_s)$ be a feasible solution to System (22) for some $\alpha \in \mathbb{R}^n$ with $\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_n > 0$. Let $\boldsymbol{Y}$ be the matrix consisting of all vapor*

phase concentration variables Y_{k,l_s} that are implied by $\boldsymbol{X}$ through the phase equilibrium equations (16). *Then, for each aggregated component $k = 1, \ldots, n$, both sequences $\{X_{k,l_s}\}_{l_s=1}^{u+1}$ and $\{Y_{k,l_s}\}_{l_s=1}^{u+1}$ are non-decreasing.*

Proof By definition, we have that $X_{n,l_s} = 1$ holds for $l_s = 1, \ldots, u+1$. Thus, the statement holds for $k = n$. For each fixed $l_s \in \{1, \ldots, u\}$, we obtain from Theorem 1 that

$$\frac{X_{1,l_s+1} - X_{1,l_s}}{X_{1,l_s}} \geq \frac{X_{2,l_s+1} - X_{2,l_s}}{X_{2,l_s}} \geq \cdots \geq \frac{X_{n,l_s+1} - X_{n,l_s}}{X_{n,l_s}} = 0.$$

By assumption $X_{k,l_s} > 0$, for $k = 1, \ldots, n$ and $l_s = 1, \ldots, u+1$, it follows that $X_{k,l_s+1} - X_{k,l_s} \geq 0$, for all $k = 1, \ldots, n$. This proves the statement for sequence $\{X_{k,l_s}\}_{l_s=1}^{u+1}$.

Using the equations (14), and $X_{k,l_s+2} \geq X_{k,l_s+1}$, for all k and l_s, we can further derive, for all $k = 1, \ldots, n$ and $l_s = 1, \ldots, u^{\text{strip}}$, that

$$\nu_s Y_{k,l_s+1} + (1 - \nu_s)X_{k,1} \geq \nu_s Y_{k,l_s} + (1 - \nu_s)X_{k,1} \quad \Leftrightarrow \quad \nu_s(Y_{k,l_s+1} - Y_{k,l_s}) \geq 0.$$

This implies that $Y_{k,l_s+1} - Y_{k,l_s} \geq 0$ holds for $\nu_s > 0$. Moreover, if $\nu_s = 0$, then we can deduce from the equations (14) that $X_{k,1} = X_{k,2} = \cdots = X_{k,u+1}$. By phase equilibrium equations (16), it follows that $Y_{k,1} = Y_{k,2} = \cdots = Y_{k,u+1}$. $\qquad \square$

4.2 Rectifying Section

Next, we prove monotonicity for the profiles when restricted to the rectifying section. We omit superscript "rect" and denote the matrices consisting of all liquid and vapor phase concentration variables of aggregated components (including the dummy one) and restricted to the rectifying section by $\boldsymbol{X} := (X_{k,l_r})_{k=0,\ldots,n,\, l_r=1,\ldots,u+1}$ and $\boldsymbol{Y} := (Y_{k,l_r})_{k=0,\ldots,n,\, l_r=1,\ldots,u+1}$.

Recap that in our model description the trays in the rectifying section are labeled from top to bottom. For this labeling, we show that the sequences $\{X_{k,l_r}\}_{l_r=1}^{u+1}$ and $\{Y_{k,l_r}\}_{l_r=1}^{u+1}$ are non-increasing for every $k = 1, \ldots, n$. Therefore, the profiles considered from the bottom to the top are non-decreasing.

In order to derive a system for the rectifying section that corresponds to System (22) of the stripping section, we need the well-known *inverses* of the phase equilibrium equations (16). For each aggregated component $k \geq 1$ and for each tray l_r, they are given as

$$X_{k,l_r} = \frac{\sum_{j=1}^{k} \alpha_j^{-1}(Y_{j,l_r} - Y_{j-1,l_r})}{\sum_{j=1}^{n} \alpha_j^{-1}(Y_{j,l_r} - Y_{j-1,l_r})}, \quad l_r = 1, \ldots, u+1. \tag{28}$$

Using (28), we obtain from the component mass balance equations (14) the following subsystem

$$Y_{k,l_r+1} = \nu_r \frac{\sum_{j=1}^{k} \alpha_j^{-1}(Y_{j,l_r} - Y_{j-1,l_r})}{\sum_{j=1}^{n} \alpha_j^{-1}(Y_{j,l_r} - Y_{j-1,l_r})} + (1 - \nu_r)Y_{k,1}, \quad k = 0, \ldots, n,$$

$$l_r = 1, \ldots, u,$$
$$0 < Y_{1,l_r} \leq Y_{2,l_r} \leq \cdots \leq Y_{n,l_r}, \quad l_r = 1, \ldots, u+1, \quad (29)$$
$$Y_{0,l_r} = 0, \quad Y_{n,l_r} = 1, \quad l_r = 1, \ldots, u+1,$$
$$\mathbf{Y} \in \mathbb{R}^{(n+1)\times(u+1)}, \quad v_r \in [0,1],$$

that must be satisfied by every feasible solution of our distillation column model.

Now, we make use of the following transformation rules.

$$\hat{u} := u, \quad \hat{\nu}_s := \nu_r,$$
$$\hat{Y}_{k,l} := (1 - X_{n-k,l}), \quad k = 0, \ldots, n, \quad l = 1, \ldots, u+1, \quad (30)$$
$$\hat{X}_{k,l} := (1 - Y_{n-k,l}),$$
$$\hat{\alpha}_k := \alpha_{n+1-k}^{-1}, \quad k = 1, \ldots, n.$$

These rules allow us to restate System (29) equivalently as follows (see Appendix A, for details).

$$\hat{X}_{k,l+1} = \hat{\nu}_s \frac{\sum_{j=1}^{k} \hat{\alpha}_j(\hat{X}_{j,l} - \hat{X}_{j-1,l})}{\sum_{j=1}^{n} \hat{\alpha}_j(\hat{X}_{j,l} - \hat{X}_{j-1,l})} + (1 - \hat{\nu}_s)\hat{X}_{k,1}, \quad k = 0, \ldots, n,$$

$$l = 1, \ldots, \hat{u},$$
$$0 < \hat{X}_{1,l} \leq \hat{X}_{2,l} \leq \cdots \leq \hat{X}_{n,l}, \quad l = 1, \ldots, \hat{u}+1, \quad (31)$$
$$\hat{X}_{0,l} = 0, \quad \hat{X}_{n,l} = 1, \quad l = 1, \ldots, \hat{u}+1,$$
$$\hat{\mathbf{X}} \in \mathbb{R}^{(n+1)\times(\hat{u}+1)}, \quad \hat{v}_s \in [0,1].$$

Note that index k appears in the variables $\hat{X}_{k,l}$ in reverse order compared with the variables Y_{k,l_r}. Therefore, we have that in reverse order

$$\hat{\alpha}_1 \geq \hat{\alpha}_2 \geq \cdots \geq \hat{\alpha}_n > 0$$

holds when $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n > 0$ in the original order. This means that System (31) satisfies all conditions of Theorem 1 and Corollary 2. We can, hence, conclude that, for every $k = 1, \ldots, n$, both sequences $\{\hat{X}_{k,l}\}_{l=1}^{\hat{u}+1}$ and $\{\hat{Y}_{k,l}\}_{l=1}^{\hat{u}+1}$ are non-decreasing. Using the transformations rules (30), again, we can draw the following conclusion.

Corollary 3 *Let (Y, ν_r) be a feasible solution to System (29) for some $\alpha \in \mathbb{R}^n$ with $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n > 0$. Let $\mathbf{X}$ be the matrix consisting of all liquid phase concentrations variables X_{k,l_r}, that are implied by $\mathbf{Y}$ through the phase equilibrium equations (28). Then, for each aggregated component $k \in \{1, \ldots, n\}$, both sequences $\{Y_{k,l_r}\}_{l_r=1}^{u+1}$ and $\{X_{k,l_r}\}_{l_r=1}^{u+1}$ are non-increasing.*

5 Domain Reduction Techniques

In this section, we use the results from Section 4 to derive a problem-specific bound tightening strategy for our distillation column model associated with the aggregated components (Subsection 5.1). For this, we adapt the arguments used in previous work [2] for binary mixtures to the multi-component case with aggregated components. In Subsection 5.2, we moreover restate a method to derive additional bounds on the aggregated concentration variables by computing the fixed points of the concentration profiles. This method has already been applied in [15] to the binary distillation case. Both techniques are implemented in global optimization software and their impact is computationally evaluated in Section 6.

5.1 Bound Tightening Strategy

The monotonic behavior of the aggregated concentration profiles together with the aggregated component mass balance equations allow us to propagate given bounds on the aggregated concentration variables at a certain tray to the aggregated concentration variables associated with an adjacent tray.

For this, recap that, for every $k = 1, \ldots, n$, the aggregated component mass balance equations (14)

$$
\begin{aligned}
Y_{k,l_\mathrm{r}+1}^\mathrm{rect} &= \nu_\mathrm{r}\, X_{k,l_\mathrm{r}}^\mathrm{rect} + (1-\nu_\mathrm{r})\, Y_{k,1}^\mathrm{rect}, \qquad l_\mathrm{r} = 1, \ldots, u^\mathrm{rect}, \\
X_{k,l_\mathrm{s}+1}^\mathrm{strip} &= \nu_\mathrm{s}\, Y_{k,l_\mathrm{s}}^\mathrm{strip} + (1-\nu_\mathrm{s})\, X_{k,1}^\mathrm{strip}, \qquad l_\mathrm{s} = 1, \ldots, u^\mathrm{strip}
\end{aligned}
$$

associated with the trays in the rectifying and stripping sections form two families of recursive functions (one for each section). By analyzing the partial derivatives, one can show that in both cases the recursive functions behave monotonic in each of their arguments (see also [2] for the binary case). The analysis is mainly straightforward. Only the partial derivatives

$$
\frac{\partial Y_{k,l_\mathrm{r}+1}^\mathrm{rect}}{\partial \nu_\mathrm{r}} = X_{k,l_\mathrm{r}}^\mathrm{rect} - Y_{k,1}^\mathrm{rect} \quad \text{and} \quad \frac{\partial X_{k,l_\mathrm{s}+1}^\mathrm{strip}}{\partial \nu_\mathrm{s}} = Y_{k,l_\mathrm{s}}^\mathrm{strip} - X_{k,1}^\mathrm{strip}
$$

need special attention. For these, we remark that the monotonicity of the concentration profiles ensures for each aggregated component $k = 1, \ldots, n$ that

$$
\begin{aligned}
X_{k,l_\mathrm{r}}^\mathrm{rect} &\le X_{k,1}^\mathrm{rect}, \quad l_\mathrm{r} = 1, \ldots, u^\mathrm{rect}, \\
Y_{k,l_\mathrm{s}}^\mathrm{strip} &\ge Y_{k,1}^\mathrm{strip}, \quad l_\mathrm{s} = 1, \ldots, u^\mathrm{strip},
\end{aligned}
\tag{32}
$$

hold. As $X_{k,l_\mathrm{s}} > 0$, we moreover observe that the phase equilibrium equations can be restated as follows:

$$
Y_{k,l_\mathrm{r}}^\mathrm{rect} = X_{k,l_\mathrm{r}}^\mathrm{rect}\, \frac{\sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_\mathrm{r}}^\mathrm{rect}}{X_{k,l_\mathrm{r}}^\mathrm{rect}} - \frac{X_{j-1,l_\mathrm{r}}^\mathrm{rect}}{X_{k,l_\mathrm{r}}^\mathrm{rect}} \right)}{X_{k,l_\mathrm{r}}^\mathrm{rect}\, \sum_{j=1}^{k} \alpha_j \left(\frac{X_{j,l_\mathrm{r}}^\mathrm{rect}}{X_{k,l_\mathrm{r}}^\mathrm{rect}} - \frac{X_{j-1,l_\mathrm{r}}^\mathrm{rect}}{X_{k,l_\mathrm{r}}^\mathrm{rect}} \right) + \sum_{j=k+1}^{n} \alpha_j \left(X_{j,l_\mathrm{r}}^\mathrm{rect} - X_{j-1,l_\mathrm{r}}^\mathrm{rect} \right)}.
\tag{33}
$$

Note further that the numerator in equation (33) is a convex combination of parameters $\alpha_1, \ldots, \alpha_k$. From $\alpha_k \leq \alpha_{k-1} \leq \cdots \leq \alpha_1$, it follows that the numerator is greater or equal to α_k. The denominator is a convex combination of the numerator and parameters $\alpha_{k+1}, \ldots, \alpha_n$. As $\alpha_n \leq \cdots \leq \alpha_{k+1} \leq \alpha_k$, we can conclude that the fractional term in the right-hand-side of equation (33) is greater or equal to one. This implies that $Y_{k,l_r}^{\text{rect}} \geq X_{k,l_r}^{\text{rect}}$. In a similar way, we can verify that $Y_{k,l_s}^{\text{strip}} \geq X_{k,l_s}^{\text{strip}}$ holds. Combining these results shows that $\partial Y_{k,l_r+1}^{\text{rect}}/\partial\nu_r \leq 0$ and $\partial X_{k,l_s+1}^{\text{strip}}/\partial\nu_s \geq 0$.

Thus, given bounds on the arguments in (14) can be used to compute bounds on the aggregated concentration variables associated with the consecutive tray via standard interval arithmetic (e.g., see also [12]). The resulting formulas are stated in the following two lemmas, where Lemma 4 addresses the stripping section and Lemma 5 deals with the rectifying section. In both lemmas, superscripts "strip" and "rect" are neglected in order to keep the notation simple.

Lemma 4 *(Stripping Section)*
Consider any $l_s \in \{1, \ldots, u\}$ and $k \in \{1, \ldots, n\}$. Assume further that ν_s ranges on $[\nu_s^{lo}, \nu_s^{up}]$, Y_{k,l_s} ranges on $[Y_{k,l_s}^{lo}, Y_{k,l_s}^{up}]$ and that $X_{k,1}$ ranges on $[X_{k,1}^{lo}, X_{k,1}^{up}]$. Then, lower and upper bounds X_{k,l_s+1}^{lo}, X_{k,l_s+1}^{up} on X_{k,l_s+1} are given by

$$X_{k,l_s+1}^{lo} = \nu_s^{lo} Y_{k,l_s}^{lo} + (1 - \nu_s^{lo}) X_{k,1}^{lo}, \quad X_{k,l_s+1}^{up} = \nu_s^{up} Y_{k,l_s}^{up} + (1 - \nu_s^{up}) X_{k,1}^{up}.$$

Lemma 5 *(Rectifying Section)*
Consider any $l_r \in \{1, \ldots, u\}$ and $k \in \{1, \ldots, n\}$. Assume further that ν_r ranges on $[\nu_r^{lo}, \nu_r^{up}]$, X_{k,l_r} ranges on $[X_{k,l_r}^{lo}, X_{k,l_r}^{up}]$ and that $Y_{k,1}$ ranges on $[Y_{k,1}^{lo}, Y_{k,1}^{up}]$. Then, lower and upper bounds Y_{k,l_r+1}^{lo}, Y_{k,l_r+1}^{up} on Y_{k,l_r+1} are given by

$$Y_{k,l_r+1}^{lo} = \nu_r^{up} X_{k,l_r}^{lo} + (1 - \nu_r^{up}) Y_{k,1}^{lo}, \quad Y_{k,l_r+1}^{up} = \nu_r^{lo} X_{k,l_r}^{up} + (1 - \nu_r^{lo}) Y_{k,1}^{up}.$$

It is worth to mention that the equations (32) do not hold for the variables used for the molar fractions in the original model formulation from Section 2, e.g., see molar fraction profile of second component (blue-colored dashed curve) in Figure 2 (a). Thus, Lemma 4 and Lemma 5 are not applicable in that case.

The purity constraints (19) already provide (strong) valid bounds on the concentration variables at the condenser and, hence, at the first tray of the rectifying section as well as on the concentration variables at the reboiler and the first tray (in our ordering) of the stripping section. Starting with these bounds, our next goal is to propagate bounds on the concentration variables tray by tray through each section by repeatedly applying the formulas for bound calculations from Lemma 4 and Lemma 5. This procedure defines the bound tightening strategy.

However, a re-use of the formulas from Lemma 4 and Lemma 5 will make it necessary to translate bounds on the aggregated concentration variables associated with the vapor or the liquid phase into valid bounds on the aggregated concentration variables in the other phase. This is achieved by exploiting

the phase equilibrium equations (16) and their inverses (28), respectively, and leads to the formulas as given in Lemma 6 (for the stripping section) and in Lemma 7 (for the rectifying section). Again, superscripts "strip" and "rect" are omitted to keep the statements easy to read. The proofs are postponed to Appendix B.

Lemma 6 *(Stripping Section)*
Let $l_s \in \{1, \ldots, u+1\}$ be fixed. Assume further that, for every $k = 1, \ldots, n$, lower and upper bounds $X_{k,l_s}^{lo}, X_{k,l_s}^{up}$ on X_{k,l_s} are given, where

$$X_{k,l_s}^{lo} \leq X_{k+1,l_s}^{lo} \quad and \quad X_{k,l_s}^{up} \leq X_{k+1,l_s}^{up} \quad hold \ for \ k = 1, \ldots, n-1.$$

Then, for each k, lower and upper bounds $Y_{k,l_s}^{lo}, Y_{k,l_s}^{up}$ on Y_{k,l_s} are given by

$$Y_{k,l_s}^{lo} = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_s}^{a_k} - X_{j-1,l_s}^{a_k})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_s}^{a_k} - X_{j-1,l_s}^{a_k})} \quad and \quad Y_{k,l_s}^{up} = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_s}^{b_k} - X_{j-1,l_s}^{b_k})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_s}^{b_k} - X_{j-1,l_s}^{b_k})},$$

where we define for $j = 1, \ldots, n$

$$X_{j,l_s}^{a_k} := \begin{cases} X_{j,l_s}^{lo}, & if \ j \leq k, \\ X_{j,l_s}^{up}, & if \ j > k, \end{cases} \quad and \quad X_{j,l_s}^{b_k} = \begin{cases} X_{j,l_s}^{up}, & if \ j \leq k, \\ \max\{X_{k,l_s}^{up}, X_{j,l_s}^{lo}\}, & if \ j > k. \end{cases}$$

Lemma 7 *(Rectifying Section)*
Let $l_r \in \{1, \ldots, u+1\}$ be fixed. Assume further that, for every $k = 1, \ldots, n$, lower and upper bounds $Y_{k,l_r}^{lo}, Y_{k,l_r}^{up}$ on Y_{k,l_r} are given, where

$$Y_{k,l_r}^{lo} \leq Y_{k+1,l_r}^{lo} \quad and \quad Y_{k,l_r}^{up} \leq Y_{k+1,l_r}^{up} \quad hold \ for \ k = 1, \ldots, n-1.$$

Then, for each k, lower and upper bounds $X_{k,l_r}^{lo}, X_{k,l_r}^{up}$ on X_{k,l_r} are given by

$$X_{k,l_r}^{lo} = \frac{\sum_{j=1}^{k} \alpha_j^{-1} (Y_{j,l_r}^{a_k} - Y_{j-1,l_r}^{a_k})}{\sum_{j=1}^{n} \alpha_j^{-1} (Y_{j,l_r}^{a_k} - Y_{j-1,l_r}^{a_k})} \quad and \quad X_{k,l_r}^{up} = \frac{\sum_{j=1}^{k} \alpha_j^{-1} (Y_{j,l_r}^{b_k} - Y_{j-1,l_r}^{b_k})}{\sum_{j=1}^{n} \alpha_j^{-1} (Y_{j,l_r}^{b_k} - Y_{j-1,l_r}^{b_k})},$$

where we define for $j = 1, \ldots, n$

$$Y_{j,l_r}^{a_k} := \begin{cases} \min\{Y_{j,l_r}^{up}, Y_{k,l_r}^{lo}\}, & if \ j < k, \\ Y_{j,l_r}^{up}, & if \ j \geq k, \end{cases} \quad and \quad Y_{j,l_r}^{b_k} = \begin{cases} Y_{j,l_r}^{lo}, & if \ j < k, \\ Y_{j,l_r}^{up}, & if \ j \geq k. \end{cases}$$

5.2 Fixed Points

A further way to tighten the model formulation from Section 3 is to determine the fixed points from the recursive functions given by the component mass balance equations (14) and the phase equilibrium equations (16). Corollaries 2 and 3 imply that, for each aggregated component k, the four infinite sequences

$$\{X_{k,l_s}^{strip}\}_{l_s \in \mathbb{Z}_{\geq 1}}, \ \{Y_{k,l_s}^{strip}\}_{l_s \in \mathbb{Z}_{\geq 1}} \quad and \quad \{X_{k,l_r}^{rect}\}_{l_r \in \mathbb{Z}_{\geq 1}}, \ \{Y_{k,l_r}^{rect}\}_{l_r \in \mathbb{Z}_{\geq 1}}$$

must converge since they are monotonic and range on the bounded interval $[0, 1]$. Due to the monotonic behavior, the limit of each sequence further provides either a lower or an upper bound valid for each element of the sequence. As done in [15] for the binary case, we can exploit that property by incorporating, for each $k = 1, \ldots, n$, the following (redundant) nonlinear constraints to the aggregated model formulation.

$$
\begin{aligned}
& X_k^{\text{strip},\star} = \nu_{\text{s}} Y_k^{\text{strip},\star} + (1 - \nu_{\text{s}}) X_{k,1}^{\text{strip}}, \\
& Y_k^{\text{strip},\star} = \frac{\sum_{j=1}^{k} \alpha_j (X_j^{\text{strip},\star} - X_{j-1}^{\text{strip},\star})}{\sum_{j=1}^{n} \alpha_j (X_j^{\text{strip},\star} - X_{j-1}^{\text{strip},\star})}, \\
& X_k^{\text{strip},\star} \geq X_{k,l_{\text{s}}}^{\text{strip}}, \quad Y_k^{\text{strip},\star} \geq Y_{k,l_{\text{s}}}^{\text{strip}}, \qquad l_{\text{s}} = 1, \ldots, u^{\text{strip}} + 1, \\
& Y_k^{\text{rect},\star} = \nu_{\text{r}} X_k^{\text{rect},\star} + (1 - \nu_{\text{r}}) Y_{k,1}^{\text{rect}}, \\
& X_k^{\text{rect},\star} = \frac{\sum_{j=1}^{k} \alpha_j^{-1} (Y_j^{\text{rect},\star} - Y_{j-1}^{\text{rect},\star})}{\sum_{j=1}^{n} \alpha_j^{-1} (Y_j^{\text{rect},\star} - Y_{j-1}^{\text{rect},\star})}, \\
& X_k^{\text{rect},\star} \leq X_{k,l_{\text{r}}}^{\text{rect}}, \quad Y_k^{\text{rect},\star} \leq Y_{k,l_{\text{r}}}^{\text{rect}}, \qquad l_{\text{r}} = 1, \ldots, u^{\text{rect}} + 1.
\end{aligned}
\tag{34}
$$

We however remark that, in general, the fixed points lead to weaker bounds than the bounds that can be obtained by the bound tightening strategy. This holds as the bounds obtained by the bound tightening strategy are monotonic with respect to the section lengths and also bounded by the fixed points.

6 Computational Studies

In this section, we computationally evaluate the impact of the presented techniques on the performance of global optimization software. For this, we consider several numerical test instances dealing with ideal multi-component distillation processes. The objective of all instances is to find an optimal column design w.r.t. cost function (20) that separates the more volatile components from the less volatile components.

6.1 Test Setting

We consider 16 test instances. The reference test instance *ref* consists of a mixture of $n = 4$ components with initial composition $x_i^{\text{in}} = \frac{1}{4}$ for $i = 1, 2, 3, 4$, and with the constant relative volatilities $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (6, 4, 1.2, 1)$. The number l^{col} of trays that can be used in the entire distillation column is bounded by 25. Every section consists of at least one tray, so that the upper bound on the number of trays used for the rectifying as well as for the stripping section is given by $u^{\text{rect}} = u^{\text{strip}} = 23$. Molar flows F, B, D and V are given in terms of $\text{mol}\,s^{-1}$. The feed molar flow F is fixed to 1, while the remaining molar flows are variable and may range as follows: $0\ \text{mol}\,s^{-1} \leq V \leq 20\ \text{mol}\,s^{-1}$, $0\ \text{mol}\,s^{-1} \leq B, D \leq 1\ \text{mol}\,s^{-1}$. We choose the split σ to be 2. Recap from

Section 3 that σ defines the more volatile single components $(1,\dots,\sigma)$ withdrawn from the condenser and the less volatile single components $(\sigma+1,\dots,n)$ withdrawn from the reboiler. With respect to split σ, we call the components σ and $\sigma+1$ *key* components, while the others are called *non-key* components. The purity requirements are given by $\pi^{\mathrm{dist}} = \pi^{\mathrm{bot}} = 0.99$.

The remaining test instances are defined by changing the values of several parameters, resulting in five groups of further test instances that are briefly explained, next.

The first group is defined by varying the constant relative volatilities for the *non-key* components from the reference instance. The specifications are given in Table 3.

Instance	*adis1*	*adis2*	*adis3*
$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$	$(12, 4, 1.2, 1)$	$(12, 8, 2.4, 1)$	$(24, 8, 2.4, 1)$

Table 3: Specification of test instances with change in the distribution of volatilities

The second group consists of two further instances for that the split σ is changed. Moreover, the constant relative volatilities are adapted in such a way that the ratios between the volatilities of the key components are the same as in the reference setting. Table 4 shows the concrete specifications.

Instance	*apos1*	*apos2*
σ	1	3
$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$	$(6.67, 2, 1.5, 1)$	$(8, 5, 3.33, 1)$

Table 4: Specification of test instances with changing σ

In the third group, we change the initial composition of the mixture as given in Table 5.

Instance	*con1*	*con2*	*con3*
$(x_1^{\mathrm{in}}, x_2^{\mathrm{in}}, x_3^{\mathrm{in}}, x_4^{\mathrm{in}})$	$\left(\frac{1}{10}, \frac{2}{5}, \frac{2}{5}, \frac{1}{10}\right)$	$\left(\frac{1}{4}, \frac{2}{5}, \frac{1}{10}, \frac{1}{4}\right)$	$\left(\frac{1}{4}, \frac{1}{10}, \frac{2}{5}, \frac{1}{4}\right)$

Table 5: Specification of test instances by varying the initial composition of the mixture

Group four consists of the test instances, for that we vary the purity requirements on condenser and reboiler. In addition, we adapt the constant relative volatilities in order to keep the separation processes approximately as

difficult as the separation process of the reference instance. Table 6 provides the specific setting for the changed parameters.

Instance	*pur1*	*pur2*	*pur3*
$(\pi^{\mathrm{dist}}, \pi^{\mathrm{bot}})$	$(0.95, 0.95)$	$(0.99, 0.95)$	$(0.95, 0.99)$
$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$	$(3.64, 2.42, 1.2, 1)$	$(4.62, 3.08, 1.2, 1)$	$(4.62, 3.08, 1.2, 1)$

Table 6: Specification of test instances with different purity requirements

Finally, we define a fifth group of test instances in which different numbers of components are considered. For each such instance, we adapt split σ, initial composition and relative volatilities accordingly, as summarized in Table 7.

Instance	*comp1*	*comp2*	*comp3*	*comp4*
n	2	3	5	5
σ	1	1	2	1
$(x_1^{\mathrm{in}}, \ldots, x_n^{\mathrm{in}})$	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{4}, \frac{1}{10}, \frac{3}{40}, \frac{3}{40})$
$(\alpha_1, \ldots, \alpha_n)$	$(3.33, 1)$	$(4, 1.2, 1)$	$(6, 4, 1.2, 1.1, 1)$	$(6.67, 2, 1.5, 1.2, 1)$

Table 7: Test instances where the number of components is changed

6.2 Problem Formulation

For each instance, two MINLP formulations are derived. The first formulation, called *MINLP-O*, is based on the original distillation column model (9) as presented in Section 2. The second formulation makes use of the model formulation (21) with aggregated concentration variables as introduced in Section 3 and is called *MINLP-A*. Additionally, we apply several different solution strategies to *MINLP-A*. The first strategy, indicated by *w/Mo*, adds the (redundant) conditions on monotonicity from Corollaries 2 and 3. The second strategy takes the fixed-point equations (34) into account and is labeled by *w/Fix*.

All formulations have been implemented using the following standard reformulation techniques. Due to their redundancy, all variables that are associated with the (corresponding) last component n as well as the variables $x_i^{\mathrm{dist}}, x_i^{\mathrm{bot}}$, $i = 1, \ldots, n$, and $X_k^{\mathrm{dist}}, X_k^{\mathrm{bot}}$, $k = 1, \ldots, n$, are eliminated. Moreover, each rational function appearing in one of the model constraints is expanded by its denominator and restated by polynomial constraints.

Note that the coupling conditions (6a) and (18a) involve quadratic terms including binary variables. We use a standard approach to linearize these types of equations. We applied several different linearization techniques and found,

based on preliminary computations, that the following one is best suited for our cases. Consider the first line of the equations (18a)

$$X_k^{\text{feed-1}} = \sum_{l_{\mathrm{r}}=1}^{u^{\text{rect}}} \beta_{l_{\mathrm{r}}}^{\text{rect}} X_{k,l_{\mathrm{r}}}^{\text{rect}},$$

for a fixed $k \in \{1, \ldots, n\}$. Using that $\sum_{l_{\mathrm{r}}=1}^{u^{\text{rect}}} \beta_{l_{\mathrm{r}}}^{\text{rect}} = 1$ holds, we can reformulate the equation as follows.

$$\beta_{l_{\mathrm{r}}}^{\text{rect}} X_k^{\text{feed-1}} = \beta_{l_{\mathrm{r}}}^{\text{rect}} X_{k,l_{\mathrm{r}}}^{\text{rect}}, \qquad l_{\mathrm{r}} = 1, \ldots, u^{\text{rect}}.$$

Next, we introduce a new variable $C_{l_{\mathrm{r}}}$ for $l_{\mathrm{r}} = 1, \ldots, u^{\text{rect}}$ and demand

$$C_{l_{\mathrm{r}}} = \beta_{l_{\mathrm{r}}}^{\text{rect}} X_k^{\text{feed-1}}, \quad C_{l_{\mathrm{r}}} = \beta_{l_{\mathrm{r}}}^{\text{rect}} X_{k,l_{\mathrm{r}}}^{\text{rect}}, \qquad l_{\mathrm{r}} = 1, \ldots, u^{\text{rect}}.$$

For every $l_{\mathrm{r}} = 1, \ldots, u^{\text{rect}}$, these two quadratic equations are linearized in the following well-known way.

$$\begin{aligned} &C_{l_{\mathrm{r}}} \geq \beta_{l_{\mathrm{r}}}^{\text{rect}} + X_k^{\text{feed-1}} - 1, \quad &&C_{l_{\mathrm{r}}} \geq \beta_{l_{\mathrm{r}}}^{\text{rect}} + X_{k,l_{\mathrm{r}}}^{\text{rect}} - 1, \\ &C_{l_{\mathrm{r}}} \leq \beta_{l_{\mathrm{r}}}^{\text{rect}}, \quad &&C_{l_{\mathrm{r}}} \leq X_k^{\text{feed-1}}, \quad &&&C_{l_{\mathrm{r}}} \leq X_{k,l_{\mathrm{r}}}^{\text{rect}}. \end{aligned}$$

The remaining equations in (18a) and (6a) are handled analogously.

All computations are carried out on a 3.00GHz Intel Xeon E5450 Processor with a limit of 30 GB memory space for each run. Moreover, running time is limited to 24 hours and the relative optimality gap is chosen to be 10^{-4}.

In order to compare different solving strategies for our test instances, we utilize a standard performance measure. It is given by the geometric mean of the solution times for each instance relative to a reference strategy. We will use this as an auxiliary tool in our analysis without relying solely on it, as the number of test instances is quite restricted. For calculating the average solution time, we use 24 hours for all instances that have not been solved during the time limit.

6.3 Results using SCIP

All MINLPs are implemented and solved with SCIP 3.2 [1] using CPLEX 12.6.0 [14] as LP-sub-solver and IPOPT 3.12.4 [27] (incl. HSL-routines MA27 and MC19 [13]) as NLP-sub-solver.

By using SCIP we are able to apply a third solution strategy to *MINLP-A*. Therein, the bound tightening procedure as described in Section 5 is used at every node in the branch and bound tree. This is achieved by implementing a domain propagation routine as an own `constraint_handler` in SCIP. The label *w/BT* indicates that the bound tightening strategy is switched on.

Our computational results are summarized in the Tables 8 to 11. Table 8 shows the running time in CPU minutes for all instances that have been solved

in the time limit of 24 hours. Table 9 displays the total number of Branch-and-Bound nodes needed in the solving process for these instances. Table 10 lists the relative gap between upper and lower bound in percentage after 24 hours for all instances that have not been solved in this time. Table 11 shows the geometric mean of the solution times for each instance relative to the standard formulation *MINLP-O* and to other formulations, respectively.

Ex.	MINLP-O	MINLP-A	MINLP-A w/Mo	MINLP-A w/Mo,Fix	MINLP-A w/BT	MINLP-A w/Mo,BT
ref	–	95	61	90	9	**6**
adis1	280	–	88	123	24	**13**
adis2	–	–	372	122	26	**8**
adis3	–	–	34	40	13	**12**
apos1	249	–	–	92	15	**12**
apos2	–	51	34	20	5	**4**
con1	232	15	30	74	6	**5**
con2	–	–	36	127	**25**	45
con3	1038	97	134	–	19	**10**
pur1	–	–	658	405	64	**32**
pur2	–	–	1331	110	21	**12**
pur3	1373	462	71	–	57	**13**
comp1	1.1	0.4	0.5	0.4	0.2	**0.1**
comp2	12	14	33	28	3	**1**
comp3	–	157	290	–	**41**	46
comp4	481	–	443	–	32	**20**

Table 8: Running time in CPU minutes using the SCIP framework. Label "–" means that the problem was not solved during the time limit of 24 hours. The lowest running time for every instance is highlighted.

First we compare the original formulation (*MINLP-O*) to the aggregated one (*MINLP-A*). By using formulation *MINLP-O*, eight of the sixteen instances are not solved to global optimality within the time limit (see column 2 of Table 8). The same holds when formulation *MINLP-A* is used, but with a different subset of unsolved instances (see column 3 of Table 8). Comparing the two columns, we see that among all instances that are solved by both formulations, using formulation *MINLP-A* leads to a lower running time for all but instance *comp2*, and to a lower number of Branch-and-Bound nodes needed for all these instances (see columns 2, 3 of Table 9). Among all instances that are not solved by both formulations, the remaining optimality gap is significantly lower when formulation *MINLP-A* is used (see columns 2, 3 of Table 10). On

Ex.	MINLP-O	MINLP-A	MINLP-A w/Mo	MINLP-A w/Mo,Fix	MINLP-A w/BT	MINLP-A w/Mo,BT
ref	–	148	100	119	10	**7**
adis1	393	–	118	153	31	**19**
adis2	–	–	558	217	31	**9**
adis3	–	–	61	64	**16**	16
apos1	392	–	–	123	19	**17**
apos2	–	91	53	34	**5**	5
con1	346	19	47	95	6	**5**
con2	–	–	61	180	**31**	47
con3	1539	125	216	–	21	**12**
pur1	–	–	920	514	60	**35**
pur2	–	–	2285	151	24	**13**
pur3	1526	623	110	–	55	**14**
comp1	18	4	5	4	1	**1**
comp2	49	48	133	96	8	**5**
comp3	–	111	219	–	28	**24**
comp4	327	–	417	–	22	**16**

Table 9: Branch-and-Bound nodes needed in 1000 for solving the problem using the SCIP framework. Label "–" means that the problem was not solved during the time limit of 24 hours. The lowest number of nodes needed for every instance is highlighted.

average, the running time is reduced to 51.3% as shown in row 2 of Table 11. These observations suggest an advantageous behavior of formulation MINLP-A during the solving process. However, we need to mention that applying the reformulation alone does not always lead to an improvement, since three of our test instances are solved by using formulation MINLP-O but not by using formulation MINLP-A.

Next, we discuss the influence of adding the redundant monotonicity constraints (w/Mo) to our model formulations by examining the differences between columns 3 and 4 of Tables 8 and 9. For five of our test instances, the influence is negative for both running time and number of nodes needed in the solving process, while the opposite holds for ten instances. Only instance apos1 is not solved by formulation MINLP-A w/Mo, but still the remaining optimality gap is lower than the one obtained by using formulation MINLP-A. The average running time is reduced to 42% by using MINLP-A w/Mo instead of MINLP-A (row 3 in Table 11). Thus we can conclude that the positive influence of the additional constraints for a wide subset of our instances dominates the negative influences. A similar result can be obtained by comparing the columns 6 and 7 of Tables 8 and 9. All but two instances have a lower running

Ex.	MINLP-O	MINLP-A	MINLP-A w/Mo	MINLP-A w/Mo,Fix	MINLP-A w/BT	MINLP-A w/Mo,BT
ref	18.36	–	–	–	–	–
adis1	–	0.33	–	–	–	–
adis2	13.02	8.15	–	–	–	–
adis3	13.44	0.26	–	–	–	–
apos1	–	0.27	0.02	–	–	–
apos2	18.25	–	–	–	–	–
con1	–	–	–	–	–	–
con2	6.11	0.11	–	–	–	–
con3	–	–	–	0.33	–	–
pur1	11.07	0.10	–	–	–	–
pur2	20.97	7.9	–	–	–	–
pur3	–	–	–	1.17	–	–
comp1	–	–	–	–	–	–
comp2	–	–	–	–	–	–
comp3	22.62	–	–	0.03	–	–
comp4	–	0.02	–	0.03	–	–

Table 10: Relative gap given in percentage after 24 hours using the SCIP framework. Label "–" means that the problem was solved with a gap lower than 0.01%.

relative to	MINLP-O	MINLP-A	MINLP-A w/Mo	MINLP-A w/Mo,Fix	MINLP-A w/BT	MINLP-A w/Mo,BT
-O	100%	51.3%	21.7%	27.1%	2.9%	**1.9%**
-A	–	100%	42.3%	–	5.6%	–
w/BT	–	–	–	–	100%	66.2%
w/Mo	–	–	100%	125.0%	–	8.8%

Table 11: Geometric mean of the running times relative to selected reference formulations using the SCIP framework.

time and all but three instances have a lower amount of nodes needed when the additional monotonicity constraints are added. The average running time is reduced to 66% in this case (row 4 in Table 11).

Now we analyze the influence of the fixed point equations (w/Fix). For this, we consider the columns 3 and 4 in the Table 8. For two instances (apos1, pur2), this influence is significantly positive while in four other cases (con3, pur3, comp3, comp4), the respective instances could not be solved in the time limit. For all other instances, the influence on running time and nodes needed

(Table 9) is very mixed and differences are not as significant. The average running time by adding the fixed point strategy is increased to 125% (row 5 in Table 11). We conclude an overall small, but rather negative influence of this solving strategy, with a huge impact in some special cases.

At last we focus on our main contribution in terms of algorithmic impact, which is the problem specific bound tightening strategy (w/BT). Note that the two solving strategies using this method (columns 6 and 7 in Tables 8 and 9) are the only ones able to solve all our instances to global optimality during the given time limit. Furthermore, one of these two strategies is always the best in terms of both, running time and Branch-and-Bound nodes needed. By comparing the columns 3 and 6 in Table 8, we can see a huge improvement in running time by applying the bound tightening strategy to the aggregated model formulation. On average, the running time is reduced to 5.6% (row 3 of Table 11). A similar result is obtained by analyzing the influence of the bound tightening on the model formulation already using the monotonicity constraints (columns 4 and 7 in Table 8). The average running time in this case is reduced to 8.8% (row 5 of Table 11). These results show a significant performance improvement by applying our developed bound tightening strategy during the optimization process of ideal multi-component distillation columns.

To summarize the analysis, three of our four proposed solution strategies have a positive influence on the performance of SCIP on our test set. These strategies are the aggregated reformulation, the monotonicity constraints and especially the bound tightening. If we add all three strategies and compare them to the original model formulation (columns 2 and 7 in Table 8) we can derive an average reduction of the running time to 1.9% (row 2 of Table 11).

6.4 Results using GAMS:BARON

We finally investigate the computational behavior of another global optimization solver on our MINLP formulations. For this, we chose the standard solver BARON 16.3.4 [24] as provided within the modeling system GAMS 24.7.1[10]. The solver is used with default settings, CPLEX as LP-subsolver and CONOPT as NLP-subsolver. We remark that our focus is on the question how the solver works as a black box on the different model formulations rather than on comparing the performance with SCIP.

In the following, we consider the model formulations *MINLP-O*, *MINLP-A*, *MINLP-A w/Mo* and *MINLP-A w/Mo,Fix* as defined in Subsection 6.2. We do not see a way to implement the bound-tightening strategy in the closed-source environment GAMS, so that this strategy is excluded from further considerations.

Table 12 displays the running time in CPU minutes and Table 13 displays the number of Branch-and-Bound nodes needed for the solving process. The single instance that has not been solved during the time limit using formulation *MINLP-O* is indicated by the symbol "–" and has a remaining optimality gap of $19, 46\%$ after 24 hours. Table 14 shows the geometric mean of the running

times relative to *MINLP-O*, *MINLP-A* and *MINLP-A w/Mo* respectively. Again we will analyze the influence of the different formulations and solving strategies one by one.

Instance	MINLP-O	MINLP-A	MINLP-A w/Mo	MINLP-A w/Mo,Fix
ref	240	201	**51**	80
adis1	**46**	103	59	95
adis2	–	307	**131**	160
adis3	110	220	114	**103**
apos1	46	68	**36**	53
apos2	131	**29**	83	41
con1	117	55	**41**	242
con2	43	185	**23**	34
con3	232	174	163	**121**
pur1	376	559	88	**78**
pur2	332	**126**	190	178
pur3	392	172	**36**	53
comp1	0.8	0.7	0.7	**0.4**
comp2	**4**	14	22	15
comp3	596	**156**	304	234
comp4	**56**	182	355	590

Table 12: Running time in CPU minutes using GAMS:BARON. Label "–" means that the problem was not solved during the time limit of 24 hours. The lowest running time for every instance is highlighted.

Using the reformulation *MINLP-A* instead of *MINLP-O* has a rather insignificant influence on the solution time of our test instances. Nine of the sixteen instances are solved faster while the other seven are solved slower (columns 2 and 3 in Table 12). On average, the running time is reduced to 91.4% as shown in row 2 of Table 14. However, it is important to note that all instances are solved by using *MINLP-A*, and that all but one instance needed a lower amount of Branch-and-Bound nodes.

Next we analyze the effect of adding the monotonicity constraints (*w/Mo*) to our model formulation (columns 3 and 4 in Table 12). Although there are six instances with a higher running time using this solution strategy, the overall influence is very positive. On average, the running time is reduced to 66% (row 3 in Table 14)

Further, adding the fixed point equations (*w/Mo,Fix*) to this formulation has a small negative influence on the performance of the solver on our test

Instance	*MINLP-O*	*MINLP-A*	*MINLP-A w/Mo*	*MINLP-A w/Mo,Fix*
ref	47	11	**6**	7
adis1	20	13	**12**	13
adis2	–	37	28	**24**
adis3	22	11	**5**	13
apos1	9	7	**7**	8
apos2	3	**0.2**	14	3
con1	24	15	**8**	19
con2	18	30	**3**	5
con3	53	16	28	**5**
pur1	25	14	5	**1**
pur2	35	**7**	19	17
pur3	46	3	5	**0.3**
comp1	1	0.3	2	**0.1**
comp2	3	**3**	17	5
comp3	18	**5**	9	8
comp4	12	**9**	17	12

Table 13: Branch-and-Bound nodes needed in 1000 for solving the problem using GAMS:BARON. Label "–" means that the problem was not solved during the time limit of 24 hours. The lowest number of nodes needed for every instance is highlighted.

relative to	*MINLP-O*	*MINLP-A*	*MINLP-A w/Mo*	*MINLP-A w/Mo,Fix*
-O	100%	91.4%	**60.4%**	69.2%
-A	–	100%	66.1%	–
w/Mo	–	–	100%	114.5%

Table 14: Geometric mean of the running times relative to selected reference formulations using GAMS:BARON.

set. Half of the instances perform better and the other half performs worse in terms on running time (columns 4 and 5 in Table 14). On average, the running time is increased to 114.5% (row 4 in Table 14).

The differences between the performance of our problem formulations are considerably smaller using GAMS: BARON as a solver instead of SCIP. Nevertheless, we can detect the same general tendencies. The aggregated formulation has a small positive influence and the monotonicity constraints a significantly higher one. The fixed point equations tend to reduce the performance of the

solvers. Unfortunately, we are not able to compare the influence of the bound tightening strategy on the solver GAMS:BARON.

We observe that some of our instances benefit a lot from our solution strategies while others are rather disturbed. We assume that this holds for the following reason. Adding additional constraints on the one hand tightens the model formulation, but on the other hand increases the problem size. The respective trade off in terms of solver performance varies among the instances and leads to the observed behaviour.

7 Conclusion

This article introduced a reformulation of the variables used for mole fractions of ideal multi-component distillation column models. The reformulated variables turned out to be monotonic, which allowed us to design a bound tightening strategy and to apply it within deterministic global optimization software. It was shown that a reduction in running time is already given by using the reformulated model. Further significant reduction was achieved with the proposed problem-specific bound tightening strategy. Focus in the present paper was on simple column configurations and ideal thermodynamics. Future work will focus on global optimization of more complex column configurations, including hybrid separation processes as well as non-ideal thermodynamics.

References

1. Achterberg, T.: SCIP: solving constraint integer programs. Mathematical Programming Computation **1**(1), 1–41 (2009). doi:10.1007/s12532-008-0001-1
2. Ballerstein, M., Kienle, A., Kunde, C., Michaels, D., Weismantel, R.: Deterministic global optimization of binary hybrid distillation/melt-crystallization processes based on relaxed MINLP formulations. Optimization and Engineering **16**(2), 409–440 (2015). doi:10.1007/s11081-014-9267-5
3. Belotti, P.: Bound reduction using pairs of linear inequalities. Journal of Global Optimization **56**(3), 787–819 (2013). doi:10.1007/s10898-012-9848-9
4. Belotti, P., Cafieri, S., Lee, J., Liberti, L.: On feasibility based bounds tightening. http://www.optimization-online.org/DB_HTML/2012/01/3325.html (2012). hal-00935464
5. Belotti, P., Kirches, C., Leyffer, S., Linderoth, J., Luedtke, J., Mahajan, A.: Mixed-integer nonlinear optimization. Acta Numerica **22**, 1–131 (2013). doi:10.1017/S0962492913000032
6. Belotti, P., Lee, J., Liberti, L., Margot, F., Wächter, A.: Branching and bounds tightening techniques for non-convex MINLP. Optimization Methods and Software **24**(4-5), 597–634 (2009). doi:10.1080/10556780903087124

7. Boukouvala, F., Misener, R., Floudas, C.A.: Global optimization advances in Mixed-Integer Nonlinear Programming, MINLP, and Constrained Derivative-Free Optimization, CDFO. European Journal of Operational Research **252**(3), 701 – 727 (2016). DOI 10.1016/j.ejor.2015.12.018

8. Caprara, A., Locatelli, M.: Global optimization problems and domain reduction strategies. Mathematical Programming **125**(1), 123–137 (2010). doi:10.1007/s10107-008-0263-4

9. Domes, F., Neumaier, A.: Constraint propagation on quadratic constraints. Constraints **15**(3), 404–429 (2010). doi:10.1007/s10601-009-9076-1

10. GAMS Development Corporation: General Algebraic Modeling System (GAMS) Release 24.7.1. Washington, DC, USA (2016). URL http://www.gams.com/

11. Gleixner, A.M., Berthold, T., Müller, B., Weltge, S.: Three enhancements for optimization-based bound tightening. Journal of Global Optimization pp. 1–27 (2016). doi:10.1007/s10898-016-0450-4

12. Hansen, P., Jaumard, B., Lu, S.H.: An analytical approach to global optimization. Mathematical Programming **52**(1-3), 227–254 (1991). doi:10.1007/BF01582889

13. HSL: A collection of Fortran codes for large-scale scientific computation. URL http://www.hsl.rl.ac.uk/

14. IBM CPLEX: (2014). URL http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

15. Kunde, C., Michaels, D., Micovic, J., Lutze, P., Górak, A., Kienle, A.: Deterministic global optimization in conceptual process design of distillation and melt crystallization. Chemical Engineering and Processing: Process Intensification **99**, 132 – 142 (2016). doi:10.1016/j.cep.2015.09.010

16. Mersmann, A., Kind, M., Stichlmair, J.: Thermal Separation Technology. Springer Heidelberg Dordrecht London New York (2011)

17. Mertens, N., Kunde, C., Kienle, A., Michaels, D.: A reformulation strategy for deterministic global optimization of ideal multi-component distillation processes. In: 26th European Symposium on Computer Aided Process Engineering, pp. 691–696 (2016)

18. Misener, R., Floudas, C.A.: GloMIQO: Global mixed-integer quadratic optimizer. Journal of Global Optimization **57**(1), 3–50 (2013). doi:10.1007/s10898-012-9874-7

19. Nallasivam, U., Shah, V.H., Shenvi, A.A., Huff, J., Tawarmalani, M., Agrawal, R.: Global optimization of multicomponent distillation configurations: 2. Enumeration based global minimization algorithm. AIChE Journal **62**(6), 2071–2086 (2016). doi:10.1002/aic.15204

20. Quesada, I., Grossmann, I.E.: Global optimization algorithm for heat exchanger networks. Industrial & Engineering Chemistry Research **32**(3), 487–499 (1993). doi:10.1021/ie00015a012

21. Ratschek, H., Rokne, J.: Interval Methods. In: R. Horst, P.M. Pardalos (eds.) Handbook of Global Optimization, pp. 751–828. Springer US (1995). doi:10.1007/978-1-4615-2025-2_14

22. Sahinidis, N.V.: Global Optimization and Constraint Satisfaction: The Branch-and-Reduce Approach. In: C. Bliek, C. Jermann, A. Neumaier (eds.) Global Optimization and Constraint Satisfaction. COCOS 2002, *Lecture Notes in Computer Science*, vol. 2861, pp. 1–16. Springer Berlin Heidelberg (2003). doi:10.1007/978-3-540-39901-8_1

23. Schichl, H., Neumaier, A.: Interval Analysis on Directed Acyclic Graphs for Global Optimization. Journal of Global Optimization **33**(4), 541–562 (2005). doi:10.1007/s10898-005-0937-x

24. Tawarmalani, M., Sahinidis, N.V.: A polyhedral branch-and-cut approach to global optimization. Mathematical Programming **103**(2), 225–249 (2005). doi:10.1007/s10107-005-0581-8

25. Vigerske, S.: Decomposition in multistage stochastic programming and a constraint integer programming approach to mixed-integer nonlinear programming. Ph.D. thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II (2013). doi:10.18452/16704

26. Vigerske, S., Gleixner, A.: Scip: global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. Optimization Methods and Software **0**(0), 1–31 (2017). doi:10.1080/10556788.2017.1335312

27. Wächter, A., Biegler, L.T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Mathematical Programming **106**(1), 25–57 (2006). doi:10.1007/s10107-004-0559-y

A Proof of Corollary 3

Consider System (29) with parameters $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n > 0$ and variables (Y, ν_r). We show that applying the transformation rules (30) to (Y, ν_r) and $\alpha_1, \ldots, \alpha_n$ leads to System (31) with parameters $\hat{\alpha}_1 \geq \hat{\alpha}_2 \geq \cdots \geq \hat{\alpha}_n > 0$ and with variables $\hat{X}, \hat{\nu}_\mathrm{s}$ whose feasible solutions satisfy the conditions of Theorem 1 (and Corollary 2). Note that feasible solutions to System (29) are in one-to-one correspondence to solutions feasible to (31) via the transformation rules (30). As we obtain from Corollary 2 that, for every feasible solution $(\hat{X}, \hat{\nu}_\mathrm{s})$, the sequences $\{\hat{X}_{k,\hat{l}_\mathrm{s}}\}_{\hat{l}_\mathrm{s}=1}^{\hat{u}+1}$ with $k = 1, \ldots, n$ are non-decreasing, the corresponding sequences $\{Y_{k,l_\mathrm{r}}\}_{l_\mathrm{r}=1}^{u+1}$, $k = 1, \ldots, n$ are non-increasing.

The first part of System (29) is given as a combination of the mass balance equations (14) and the inverted phase equilibrium equations (28).

$$Y_{k,l_\mathrm{r}+1} = \nu_\mathrm{r} X_{k,l_\mathrm{r}} + (1 - \nu_\mathrm{r}) Y_{k,1} \quad \text{and} \quad X_{k,l_\mathrm{r}} = \frac{\sum_{j=1}^{k} \alpha_j^{-1}(Y_{j,l_\mathrm{r}} - Y_{j-1,l_\mathrm{r}})}{\sum_{j=1}^{n} \alpha_j^{-1}(Y_{j,l_\mathrm{r}} - Y_{j-1,l_\mathrm{r}})},$$

for $k = 1, \ldots, n$ and $l_\mathrm{r} = 1, \ldots, u$. We apply the transformation rules (30) to each constraint, separately. Recap that after the transformation, the components appear in reverse order. To indicate that, we introduce a new index $m := n - k$.

For all $k = 1, \ldots, n$ and all $l = 1, \ldots, u$, we obtain

$$Y_{k,l+1} = \nu_\mathrm{r} X_{k,l} + (1 - \nu_\mathrm{r}) Y_{k,1} \quad \Leftrightarrow \quad (1 - Y_{k,l+1}) = 1 - \left(\nu_\mathrm{r} X_{k,l} + (1 - \nu_\mathrm{r}) Y_{k,1} \right)$$
$$\Leftrightarrow \quad (1 - Y_{k,l+1}) = \nu_\mathrm{r}(1 - X_{k,l}) + (1 - \nu_\mathrm{r})(1 - Y_{k,1}).$$

Thus, the transformation rules (30) yield

$$\hat{X}_{k,l+1} = \hat{\nu}_\mathrm{s} \hat{Y}_{k,l} + (1 - \hat{\nu}_\mathrm{s}) \hat{X}_{k,1}, \qquad m = 0, \ldots, n, \quad l = 1, \ldots, u. \tag{35}$$

For the inverted phase equilibrium equations, we further derive that

$$\hat{Y}_{n-k,l} = 1 - X_{k,l} = 1 - \frac{\sum_{j=1}^{k} \alpha_j^{-1}(Y_{j,l} - Y_{j-1,l})}{\sum_{j=1}^{n} \alpha_j^{-1}(Y_{j,l} - Y_{j-1,l})} = \frac{\sum_{j=k+1}^{n} \alpha_j^{-1}(Y_{j,l} - Y_{j-1,l})}{\sum_{j=1}^{n} \alpha_j^{-1}(Y_{j,l} - Y_{j-1,l})}$$

$$= \frac{\sum_{j=k+1}^{n} \alpha_j^{-1}(1 - \hat{X}_{n-j,l} - (1 - \hat{X}_{n-(j-1),l}))}{\sum_{j=1}^{n} \alpha_j^{-1}(1 - \hat{X}_{n-j,l} - (1 - \hat{X}_{n-(j-1),l}))}$$

$$= \frac{\sum_{j=k+1}^{n} \alpha_j^{-1}(\hat{X}_{n-(j-1),l} - \hat{X}_{n-j,l})}{\sum_{j=1}^{n} \alpha_j^{-1}(\hat{X}_{n-(j-1),l} - \hat{X}_{n-j,l_\mathrm{s}})} = \frac{\sum_{j=k+1}^{n} \hat{\alpha}_{n+1-j}(\hat{X}_{n+1-j,l} - \hat{X}_{n-j,l})}{\sum_{j=1}^{n} \hat{\alpha}_{n+1-j}(\hat{X}_{n+1-j,l} - \hat{X}_{n-j,l})}$$

holds for every $k = 1, \ldots, n$ and for every $l = 1, \ldots, u$. By an index shift $p := n + 1 - j$, we derive the equivalence to the non-inverted phase equilibrium equations (16).

$$\hat{Y}_{n-k,l} = \frac{\sum_{p=1}^{n-k} \beta_p(\hat{X}_{p,l} - \hat{X}_{p-1,l})}{\sum_{p=1}^{n} \beta_p(\hat{X}_{p,l} - \hat{X}_{p-1,l})}, \qquad \begin{array}{l} n = 0, \ldots, n \\[4pt] l = 1, \ldots, u+1 \end{array}$$

$$\Leftrightarrow \qquad \hat{Y}_{m,l} = \frac{\sum_{p=1}^{m} \beta_p(\hat{X}_{p,l} - \hat{X}_{p-1,l})}{\sum_{p=1}^{n+1} \beta_p(\hat{X}_{p,l} - \hat{X}_{p-1,l})}, \qquad \begin{array}{l} m = 0, \ldots, n \\[4pt] l = 1, \ldots, u+1 \end{array} \tag{36}$$

Combining the equations (35) and (36), we obtain the first line from System (31).

The second line of System (31) results from the following relation.

$$Y_{k+1,l} \geq Y_{k,l}, \ k = 0, \ldots, n-1 \quad \Leftrightarrow \quad \hat{X}_{m-1,l} \leq \hat{X}_{m,l}, \ m = 1, \ldots, n.$$

The third line of System (31) trivially holds.

It remains to argue that the transformed constant relative volatilities $\hat{\alpha}_m = \alpha_{(n+1)-m}^{-1}$, $m = 1, \ldots, n$, are strictly positive and monotonically non-decreasing in the new ordering of the components. This, however, holds as $\alpha_1 \geq \cdots \geq \alpha_n > 0$ implies that

$$0 < \alpha_1^{-1} \equiv \hat{\alpha}_n \leq \cdots \leq \alpha_n^{-1} \equiv \hat{\alpha}_1.$$

Now, we can conclude that any feasible solution to System (31) satisfies the conditions of Theorem 1 and Corollary 2. □

B Proof of Lemma 6 and Lemma 7

We only give a proof for Lemma 6 dealing with the stripping section. The correctness of Lemma 7 can be shown in a similar way.

In what follows, superscript "strip" is omitted, again.

Proof (Lemma 6)

We interpret the aggregated phase equilibrium equations (16)

$$Y_{k,l_{\mathrm{s}}}(\boldsymbol{X}) = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}$$

as functions in the liquid phase concentration variables. For all $k, q = 1, \ldots, n$ and for all $l_{\mathrm{s}} = 1, \ldots, u+1$, we consider the partial derivatives $\partial Y_{k,l_{\mathrm{s}}}(\boldsymbol{X})/\partial X_{q,l_{\mathrm{s}}}$ where we distinguish the three cases $q \leq k-1$, $q = k$ and $q \geq k+1$. To keep notation short, we introduce constant $\alpha_{n+1} := 0$.

For $q \leq k-1$, we obtain

$$\begin{aligned}
&\frac{\partial Y_{k,l_{\mathrm{s}}}(\boldsymbol{X})}{\partial X_{q,l_{\mathrm{s}}}} \\
&= \frac{(\alpha_q - \alpha_{q+1}) \sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{\left(\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})\right)^2} - \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})(\alpha_q - \alpha_{q+1})}{\left(\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})\right)^2} \\
&= \frac{(\alpha_q - \alpha_{q+1}) \sum_{j=k+1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{\left(\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})\right)^2}.
\end{aligned}$$

As $(\alpha_q - \alpha_{q+1}) \geq 0$ holds, this derivative is non-negative for all $k = 1, \ldots, n$ and all $l_{\mathrm{s}} = 1, \ldots, u+1$.

For $q = k$ we obtain

$$\begin{aligned}
\frac{\partial Y_{q,l_{\mathrm{s}}}(\boldsymbol{X})}{\partial X_{q,l_{\mathrm{s}}}} &= \frac{\alpha_q \sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}}) - \sum_{j=1}^{q} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})(\alpha_q - \alpha_{q+1})}{\left(\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})\right)^2} \\
&= \frac{\alpha_q \sum_{j=q+1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}}) + \alpha_{q+1} \sum_{j=1}^{q} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})}{\left(\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})\right)^2},
\end{aligned} \tag{37}$$

which is also non-negative for all $q = 1, \ldots, n$ and all $l_{\mathrm{s}} = 1, \ldots, u+1$.

For $q \geq k+1$ we obtain

$$\frac{\partial Y_{k,l_{\mathrm{s}}}(\boldsymbol{X})}{\partial X_{q,l_{\mathrm{s}}}} = \frac{-\sum_{j=1}^{k} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})(\alpha_q - \alpha_{q+1})}{\left(\sum_{j=1}^{n} \alpha_j (X_{j,l_{\mathrm{s}}} - X_{j-1,l_{\mathrm{s}}})\right)^2}, \tag{38}$$

which is non-positive for all $k = 1, \ldots, n$ and all $l_{\mathrm{s}} = 1, \ldots, u+1$ due to $(\alpha_q - \alpha_{q+1}) \geq 0$.

This shows that the phase equilibrium equations are component-wise monotonic. Therefore, we can apply simple interval arithmetic, again, leading to following lower and upper bounds on the vapor phase concentration variables $Y_{k,l_\mathrm{s}+1}$.

$$Y_{k,l_\mathrm{s}}^{\mathrm{lo}} = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_\mathrm{s}}^{a_k} - X_{j-1,l_\mathrm{s}}^{a_k})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_\mathrm{s}}^{a_k} - X_{j-1,l_\mathrm{s}}^{a_k})} \quad \text{and} \quad Y_{k,l_\mathrm{s}}^{\mathrm{up}} = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_\mathrm{s}}^{b_k} - X_{j-1,l_\mathrm{s}}^{b_k})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_\mathrm{s}}^{b_k} - X_{j-1,l_\mathrm{s}}^{b_k})},$$

where for $j = 1, \dots, n$

$$X_{j,l_\mathrm{s}}^{a_k} := \begin{cases} X_{j,l_\mathrm{s}}^{\mathrm{lo}}, & \text{if } j \le k, \\ X_{j,l_\mathrm{s}}^{\mathrm{up}}, & \text{if } j > k, \end{cases} \quad \text{and} \quad X_{j,l_\mathrm{s}}^{b_k} = \begin{cases} X_{j,l_\mathrm{s}}^{\mathrm{up}}, & \text{if } j \le k, \\ X_{j,l_\mathrm{s}}^{\mathrm{lo}}, & \text{if } j > k, \end{cases} \tag{39}$$

We remark that the upper bound $Y_{k,l_\mathrm{s}}^{\mathrm{up}}$ on Y_{k,l_s} is not tight when $X_{k,l_\mathrm{s}}^{\mathrm{up}} > X_{k',l_\mathrm{s}}^{\mathrm{lo}}$ holds for some $k' > k$. In those cases, we can compute an improved upper bound on Y_{k,l_s} by finding the maximum of

$$Y_{k,l_\mathrm{s}}(\boldsymbol{X}) = \frac{\sum_{j=1}^{k} \alpha_j (X_{j,l_\mathrm{s}} - X_{j-1,l_\mathrm{s}})}{\sum_{j=1}^{n} \alpha_j (X_{j,l_\mathrm{s}} - X_{j-1,l_\mathrm{s}})}$$

restricted to $X_{k',l_\mathrm{s}}^{\mathrm{lo}} \le X_{k,l_\mathrm{s}} \le X_{k',l_\mathrm{s}} \le X_{k,l_\mathrm{s}}^{\mathrm{up}}$. As $\partial Y_{k,l_\mathrm{s}}(\boldsymbol{X})/\partial X_{k,l_\mathrm{s}} \ge 0$ and $\partial Y_{k,l_\mathrm{s}}(\boldsymbol{X})/\partial X_{k',l_\mathrm{s}} \le 0$, it follows that $X_{k,l_\mathrm{s}} = X_{k',l_\mathrm{s}}$ must hold for a solution on that the maximum is attained. A comparison of the equations (37) and (38) gives rise to the following relation.

$$\frac{\partial Y_{k,l_\mathrm{s}}(\boldsymbol{X})}{\partial X_{k,l_\mathrm{s}}} + \sum_{j=k+1}^{n} \frac{\partial Y_{k,l_\mathrm{s}}(\boldsymbol{X})}{\partial X_{j,l_\mathrm{s}}} \ge 0.$$

This shows that the maximum is attained when $X_{k,l_\mathrm{s}} = X_{k',l_\mathrm{s}} = X_{k,l_\mathrm{s}}^{\mathrm{up}}$. Hence, we can replace the definition of $X_{j,l_\mathrm{s}}^{b_k}$ in the equations (39) by

$$X_{j,l_\mathrm{s}}^{b_k} = \begin{cases} X_{j,l_\mathrm{s}}^{\mathrm{up}}, & \text{if } j \le k, \\ \max \left\{ X_{k,l_\mathrm{s}}^{\mathrm{up}}, X_{j,l_\mathrm{s}}^{\mathrm{lo}} \right\}, & \text{if } j > k. \end{cases}$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$