

ADVANCES IN MINLP FOR OPTIMAL DISTILLATION COLUMN
SEQUENCING

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Radhakrishna Tumbalam Gooty

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2020

Purdue University

West Lafayette, Indiana

**THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL**

Dr. Rakesh Agrawal, Co-chair

Davidson School of Chemical Engineering

Dr. Mohit Tawarmalani, Co-chair

Krannert School of Management

Dr. Gintaras Reklaitis

Davidson School of Chemical Engineering

Dr. Joseph Pekny

Davidson School of Chemical Engineering

Approved by:

Dr. John Morgan

Head of the School Graduate Program

To my parents: Vani and Srinivas

ACKNOWLEDGMENTS

I take this opportunity to thank everyone who helped me directly or indirectly towards the completion of this thesis. First and foremost, I would like to express my deepest gratitude to my advisors, Prof. Rakesh Agrawal and Prof. Mohit Tawarmalani, for guidance and constant support. I would like to thank Prof. Agrawal for giving me freedom to pursue various aspects without any objection, and for teaching me engineering skills and various life lessons. I would like to thank Prof. Tawarmalani for introducing me to the field of global optimization, and for patiently teaching me difficult mathematical concepts which I had trouble in understanding. I cannot imagine doing any of the things in this thesis without their inputs. Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Gintaras Reklaitis and Prof. Joseph Pekny. Their feedback, insightful comments and hard questions were really valuable.

I would like to thank all the current and the previous members of the modeling group: Dr. Gautham Ramapriya, Dr. Zheyu Jiang, Dr. Parham Mobed, Dr. Jose Adrain Chavez Velasco, Dr. Taufik Ridha, Dr. Emre Gencer, Dr. Yiru Li, Tony Mathew, Zewei Chen, Edwin Rodriguez, and Peter Oladipupo. I enjoyed our (often long) discussions and learned a lot of things from you, and I hope my perspectives on various topics were helpful to you too.

I would like to thank Melissa LaGuire for scheduling meetings with Prof. Agrawal, and for assembling the nomination packet for various awards. I would like to thank the graduate program coordinator Bev Johnson for helping me with registrations, and for promptly responding to all the queries I had with completing the necessary paper work. I would like to thank the ECN staff for promptly responding to computer-related issues.

I would like to thank all my friends who made my stay here at Purdue a memorable one. Special thanks to Aaditya Candadai, Abhijit Talpade, Aditi Khot, Prabhu Venuthurumilli, Swapnil Deshmukh, Saaketh Desai, and Tony Mathew for patiently listening to me whenever I needed to vent out.

Last but not the least, I would like to thank my parents and my brother for their continued support throughout this journey.

TABLE OF CONTENTS

	Page
LIST OF TABLES	x
LIST OF FIGURES	xiii
ABSTRACT	xvi
1 Introduction	1
1.1 Intricacies of Designing Distillation-based Separation Units	2
1.2 Literature Review	5
1.2.1 State-of-the-art	9
1.3 Key Contributions	11
1.3.1 Chapter 3	11
1.3.2 Chapter 4	13
1.3.3 Chapter 5	14
2 Preliminaries	15
2.1 The Distillation Process	15
2.2 A Note on Underwood Constraints	16
2.3 Alternative Interpretation of Relative Volatilities	18
3 MINIMIZATION OF VAPOR DUTY	22
3.1 Problem Formulation	24
3.1.1 Objective Function	28
3.1.2 Search Space Formulation	28
3.1.3 Mass Balance Constraints	33
3.1.4 Underwood Constraints	37
3.1.5 Flow of Underwood Roots	42
3.1.6 Bounds on Variables	45
3.2 Case Studies	46

	Page
3.2.1 Heavy Crude Distillation	46
3.2.2 A Pseudocost Model	51
3.2.3 Case Studies from Literature	55
3.3 Concluding Remarks	56
4 Improved Formulation and Solution Approach	59
4.1 Problem Definition	59
4.2 Problem Formulation	61
4.2.1 Objective Function	61
4.2.2 Space of Admissible Configurations	62
4.2.3 Mass Balance Constraints	77
4.2.4 Underwood Constraints	83
4.2.5 Exploiting Monotonicity of Underwood Equations	84
4.3 Relaxation and Solution Procedure	85
4.3.1 Reformulation	87
4.3.2 Relaxation for $\alpha_2 < \theta^{\text{lo}}$ and $\theta^{\text{up}} < \alpha_1$	93
4.3.3 Valid Relaxation for $\theta^{\text{lo}} = \alpha_2$ and/or $\theta^{\text{up}} = \alpha_1$	100
4.3.4 Discretization and Solution Procedure	103
4.4 Effect of Individual Cuts on Relaxation	105
4.5 Computational Results	109
4.5.1 Comparison with Prior Approaches	111
4.5.2 Comparison with Nallasivam et al. [22]	115
4.6 Concluding Remarks	116
5 Minimization of Exergy Loss	118
5.1 Problem Formulation	121
5.1.1 Space of Admissible Distillation Configurations	122
5.1.2 Mass Balance Constraints	124
5.1.3 Underwood Constraints	127
5.1.4 Flow-of-roots Constraints	130

	Page
5.1.5 Objective Function and Exergy Constraints	131
5.1.6 Exploiting Monotonicity of Exergy Constraints	140
5.1.7 Relation Between Temperatures of Condensers and Reboilers .	141
5.2 Computational Studies	146
5.2.1 Case Study: Shale Gas Separation	146
5.2.2 Examples from Literature	151
5.3 Conclusion	154
6 Future Direction	155
6.1 Heat Integration	155
6.2 Improvement to the Proposed Algorithm	157
6.3 Application to Design and Operation	158
REFERENCES	159
A Test Set	165
B Missing Proofs and Derivations	167
B.1 Proof of Proposition 4.1	167
B.2 Proof of Remark 4.2	169
B.3 Proof of Proposition 4.6	170
B.4 Derivation of $\text{Conv}(\mathcal{F}_p)$	172
B.5 Proof of Corollary 4.2	173
B.6 Proof of Proposition 4.12	174
B.7 Relaxation of (4.39)	176
B.8 Derivation of MIP Representation of Piecewise Relaxation of $\mathcal{F}_1$. . .	179
B.9 Derivation of MIP representation of Piecewise Relaxation of $\mathcal{V}$	182
C Missing Identities and Derivations	185
C.1 Derivation of the Search Space Formulation	185
C.2 Gauss Quadrature Formula	189
C.3 Proof of Identity	189
C.4 Relative Volatilities for the Case Study	189

VITA	192
----------------	-----

LIST OF TABLES

Table	Page
1.1 Combinatorial explosion of the choice set. From [5].	5
3.1 Definition of sets	26
3.2 Domain of indices for the variables. For $\zeta_{i,j}$, $[i, j] \in \mathcal{T}$; For $\chi_{i,j}$ and $FC_{i,j}$, $(i, j) \in \mathcal{C}$; For $\rho_{i,j}$ and $FR_{i,j}$, $(i, j) \in \mathcal{R}$; For $U_{i,j+1}^{\text{rs}}$ and $K_{i,j+1}^{\text{rs}}$, $[i, j] \in \mathcal{S}_1 \cup \mathcal{S}_2$; For $U_{i-1,j}^{\text{ss}}$ and $K_{i-1,j}^{\text{ss}}$, $[i, j] \in \mathcal{S}_2 \cup \mathcal{S}_3$. For the remaining decision variables, $[i, j] \in \mathcal{P}$	26
3.3 Feed composition and relative volatility information for heavy crude mixture. Data is taken from [5]	47
3.4 Problem size for different systems	56
3.5 Case studies from literature. For all the cases, the feed is assumed to be in saturated liquid state ($\Phi = 1$). MINLPs are solved with BARON 18.5.8 [21, 40] on GAMS 25.1 with default values for all options except pDo (pDo is set to -1), and the relative tolerance for convergence (ϵ_r) is set to 1%. Further, the time limit is chosen to be 1200 s for four and five component systems, and 3600 s for six component system. TL corresponds to ‘Time Limit has been reached’, and the number in the parenthesis is the duality gap at the end of 1200 s. I/N/T: Baron Iterations/Maximum number of nodes in memory/Time in seconds. Case 1 is taken from [10], Cases 2, 3 and 4 from [45], Cases 6, 8 and 10 from [9], Cases 7, 9 and 11 from [17]. Relative volatility information for Cases 12 and 13 are taken from [46], and from [47] for Case 14.	57
4.1 Definition of sets.	61
4.2 CG06 for the space of admissible configurations. The first column indicates the constraint number in Table 1 of [9]. ‘Co.’, ‘Re.’ and ‘Pr.’ stand for Corollary, Remark and Proposition, respectively.	74
4.3 GA10 for space of admissible configurations. The first column indicates the section number in [10]. ‘Co.’, ‘Re.’ and ‘Pr.’ stand for Corollary, Remark and Proposition, respectively.	76
4.4 TAT19 for the space of admissible configurations. The first column indicates the constraint number in [23]. ‘Re.’ and ‘Pr.’ stand for Remark and Proposition, respectively.	77

Table	Page
4.5 Definition of continuous decision variables.	80
4.6 Variation of duality gap across the scenarios described in §4.4. Here, a gap value less than $10^{-4}\%$ is marked as 0%. For all the cases, $N = 5$, $\Phi_{1,N} = \Phi_{1,1} = \dots = \Phi_{N,N} = 1$. In Case-A , $F = \{20, 30, 20, 20, 10\}$ and $\alpha = \{4.1, 3.6, 2.1, 1.42, 1\}$; In Case-B , $F = \{5, 15, 25, 20, 35\}$ and $\alpha = \{7.98, 3.99, 3, 1.25, 1\}$; In Case-C , $F = \{25, 10, 25, 20, 20\}$ and $\alpha = \{13.72, 3.92, 3.267, 1.21, 1\}$; In Case-D , $F = \{42.5, 42.5, 5, 5\}$ and $\alpha = \{3.3275, 3.025, 1.21, 1.1, 1\}$; In Case-E , $F = \{30, 30, 5, 5, 30\}$ and $\alpha = \{1.4641, 1.331, 1.21, 1.1, 1\}$; In Case-F , $F = \{5, 5, 5, 42.5, 42.5\}$ and $\alpha = \{3.3275, 1.331, 1.21, 1.1, 1\}$; In Case-G , $F = \{5, 5, 5, 42.5, 42.5\}$ and $\alpha = \{17.1875, 6.875, 2.75, 1.1, 1\}$; In Case-H , $F = \{20, 20, 20, 20, 20\}$ and $\alpha = \{7.5625, 3.025, 1.21, 1.1, 1\}$. Data for Case-A is taken from [17], Case-B and Case-C from [46], and Case-D through Case-H from [39].	110
5.1 Definition of sets.	122
5.2 Material flowrates (in kmol/h) corresponding to the optimal operation of configuration in Figure. Material flowrates (in kmol/h) in condensers and reboilers at optimum operation: $FC_{1,1} = 63.659$, $FC_{2,2} = 20.912$, $FC_{1,3} = 1122.545$, $FR_{3,3} = 30.822$, $FR_{5,5} = 53.75$, $FR_{2,4} = 801.818$ and $FR_{4,5} = 57.5$. The remaining variables can be determined by performing mass balances, and are not shown in the table. The problem parameters are taken to be $N = 5$, $\{F_p\}_{p=1}^N = \{5, 5, 32.5, 25, 32.5\}$, $\{\alpha_p\}_{p=1}^N = \{11.761, 9.801, 3.63, 3, 1\}$, $\Phi_{1,n} = 1$ and $\{\Phi_{p,p}\}_{p=1}^N = \{1, 1, 1, 1, 1\}$. In the table, we write ‘-’ to indicate that the variable is not defined in the problem. The objective function value ($\Delta\mathcal{E}'_{\text{loss}}$) at optimum operation: -84.671 . 143	143
5.3 Composition of shale gas from Eagle Ford	147
5.4 Optimal material flowrates (in kmol/h) in rectifying and stripping sections for configurations in Figure 5.5. For Figure 5.5(a), material flowrates (in kmol/h) in condensers and reboilers at optimum operation: $FC_{1,1} = 45.54$, $FC_{2,2} = 17.64$, $FC_{3,3} = 6.23$, $FC_{4,4} = 4.03$, $FR_{2,5} = 41.73$, $FR_{3,5} = 17.68$, $FR_{4,5} = 8.77$ and $FR_{5,5} = 3.1$. For Figure 5.5(b), material flowrates (in kmol/h) in condensers and reboilers at optimum operation: $FC_{1,1} = 45.54$, $FC_{3,3} = 6.23$, $FC_{4,4} = 4.03$, $FR_{2,5} = 27.04$, $FR_{3,5} = 14.73$, $FR_{4,5} = 8.77$ and $FR_{5,5} = 3.1$. As before, we write ‘-’ to indicate that the variable is not defined in the problem.	150

5.5	Examples from literature. The remaining problem parameters are taken be $\Phi_{1,N} = 1$ and $\{\Phi_{p,p} = 1\}_{p=1}^N$. Case 1 is taken from [10], Cases 2, 3 and 4 from [45], Cases 6, 8 and 10 from [9], Cases 7, 9 and 11 from [17]. Relative volatility information for Cases 12, 13, 14 and 15 are taken from [46]. All BARON options except <code>pDo</code> are set at their defaults (<code>pDo</code> is set to -1). Relative tolerance for convergence (ϵ_r) is set to 1%.	153
-----	---	-----

LIST OF FIGURES

Figure	Page
1.1	Energy consumption of a configuration depends on the operating conditions. Here, NEC stands for normalized energy consumption. The energy consumption of (a) is used for the normalization. The composition (mole fraction) of the given four-component mixture is $\{z_A, z_B, z_C, z_D\} = \{0.45, 0.05, 0.05, 0.45\}$, and the relative volatilities of the constituent components w.r.t to the heaviest component are $\{\alpha_A, \alpha_B, \alpha_C, \alpha_D\} = \{3.025, 2.75, 1.1, 1\}$
3	
1.2	Basic configurations for the separation of a four-component mixture. Here, NE stands for normalized energy consumption, and the energy requirement for configuration (r) is used for normalization. The reported NE values are for the separation of an equimolar mixture of $A \equiv 1$ -pentene, $B \equiv$ n-pentane, $C \equiv 1$ -hexene, and $D \equiv$ n-hexane.
4	
1.3	A brief time-line and state-of-the-art
7	
1.4	(a) Performance profile for the MINLP model in [9] (b) Plot illustrating the remaining duality gap at the end of five hours.
10	
1.5	Performance profile for the nonlinear program proposed in [22]
12	
2.1	Schematic of a distillation column
16	
2.2	V_{min} diagram showing the feasible region (shaded region) for the split $ABCD/BCDE$. Here, the feed composition is $\{0.2, 0.2, 0.2, 0.2\}$, relative volatilities are $\{9, 6, 3.5, 2, 1\}$, and thermal quality = 1. The values are taken from [32]
19	
3.1	Five-component system. Here, filled and unfilled circles denote condensers and reboilers, red and blue colored arrows denote the rectifying and stripping sections.
23	
3.2	Schematic of four-component superstructure. The blue box shows a representative column of every other column in the superstructure. Variables used for vapor and liquid flows are shown in (b). Definition of the variables are described in Table 3.2
34	
3.3	Modeling of (a) Condenser, (b) Thermal coupling, (c) Combination of condenser and thermal coupling and (d) a partial condenser
36	

Figure	Page
3.4 Schematic showing component, vapor and liquid bypass	37
3.5 A five-component completely thermally coupled configuration	44
3.6 Candidate configurations for the heavy crude example	49
3.7 Configuration that minimizes the sloppy splits, maximizes the heat exchangers while being attractive from the perspective of vapor duty. (b) shows the heat integrated version of (a) where the reboiler associated with the side-stripper is replaced with a heat exchanger (thick square)	53
4.1 Figure depicting streams ($\zeta_{i,j}$), reboilers ($\rho_{i,j}$) and condensers ($\chi_{i,j}$) present in a four-component system. Section variables $\tau_{i,k,j}$ and $\beta_{i,l,j}$ are defined in (4.1).	60
4.2 Digraph for deriving conservation of components constraint in §4.2.2 . . .	67
4.3 Digraph for deriving presence of parent constraint in §4.2.2	69
4.4 Digraph for the proof of Proposition 4.5	72
4.5 (a) Representative column for splits of process feed i.e., $[i, j] \in \{[1, N]\}$, $\llbracket k \rrbracket_i^{j-1}$, $\llbracket l \rrbracket_{i+1}^{k+1}$ (b) Representative column for the remaining splits $[i, j] \in \mathcal{S}$, $\llbracket k \rrbracket_i^{j-1}$, $\llbracket l \rrbracket_{i+1}^{k+1}$ (c) Representative condenser for $(i, j) \in \mathcal{C} \setminus \{[i, i]\}_{i=1}^{N-1}$ (see (A12) for domain of indices m, n, k, l) (d) Representative reboiler for $(i, j) \in \mathcal{R} \setminus \{[i, i]\}_{i=2}^N$ (see (A13) for domain of indices m, n, k, l) (e) Representative arrangement for pure product withdrawals (see (A14) for domain of indices m, n , and (A15) for domain of indices m', n') (f) Representative arrangement for overall component mass balance for $[i, j] \in \mathcal{S}$ (see (A17) for domain of indices m, n, m', n', k, l)	79
4.6 (a) Projection of sets S , S_{std} , S_{RLT} and S_{RDLT} in Example 4.1 in $\Upsilon - \theta$ space. (b) Plots of nonlinear expression in Underwood constraint	90
4.7 (a) Plot showing percentage of cases solved to 1%–optimality against time. Here, Approach 1 corresponds to the current work, Approach 2 corresponds to the model proposed in [9] solved with BARON, after making the changes described in §4.5, and Approach 3 corresponds to the model proposed in [23] solved with BARON (b) Plot showing the remaining duality Gap at the end of five hours for all the three approaches.	113
4.8 Profiles showing remaining % Gap at the end of specific time instances for Approach 1 (A1) and Approach 3 (A3).	114
4.9 Plot showing percentage of cases solved to 1%–optimum against time, when discrete variables are fixed to fully thermally coupled configuration (see §4.5.2). Approach 4 corresponds to the model proposed in [22] solved with BARON.	116

Figure	Page
5.1 A four-component system	121
5.2 (a) Four-component superstructure and (b) A representative column in the superstructure. Reproduced from [23].	125
5.3 (a) Two-phase condenser (b) Two-phase reboiler	126
5.4 A five-component configuration motivating the need for constraints in section 2.7	141
5.5 (a) Conventional configuration for shale gas separation (b) A plausible retrofit to reduce exergy loss. Operating conditions at optimum operation are shown in Table 5.4.	148
5.6 (a) Conventional configuration for shale gas separation (b) A plausible retrofit to reduce exergy loss.	152
6.1 Four-component configurations. We refer the leftmost column as $Q1$, the middle column as $Q2$ and the rightmost column as $Q3$	156
B.1 (f_1, θ) domain for §B.6. The extreme points of $\text{Conv}(\mathcal{H}_1)$ are contained in points in red.	175
C.1 True VLE (solid blue curve) and the best surrogate (dashed red curve) in least square sense for (a) Methane-Ethane (b) Ethane-Propane (c) Propane-Butane and (d) Butane-Pentane pairs. For each pair, the surrogate model is taken to be $y = \alpha_{ij}x/(\alpha_{ij}x + 1 - x)$	191

ABSTRACT

Tumbalam Gooty, Radhakrishna PhD, Purdue University, December 2020. Advances in MINLP for Optimal Distillation Column Sequencing. Major Professors: Rakesh Agrawal and Mohit Tawarmalani.

Designing configurations for multicomponent distillation, a ubiquitous process in chemical and petrochemical industries, is often challenging. This is because, as the number of components increases, the number of admissible distillation configurations grows rapidly and these configurations vary substantially in their energy needs. Consequently, if a method could identify a few energy-efficient choices from this large set of alternatives, it would be extremely attractive to process designers. Towards this, we develop the first mixed-integer nonlinear programming (MINLP) based solution approach that successfully identifies the most energy-efficient distillation configuration for a given separation. Current sequence design strategies are largely heuristic. The rigorous approach presented here can help reduce the significant energy consumption and consequent greenhouse gas emissions by separation processes.

In addition to the combinatorial complexity, the challenge in solving this problem arises from the nonconvex fractional terms contained in the governing equations. We make several advances to enable solution of these problems.

1. We propose a novel search space formulation by embedding convex hulls of various important substructures. We prove that the resulting formulation dominates all the prior formulations in the literature.
2. We derive valid cuts to the problem by exploiting the monotonic nature of the governing equations.

3. We adapt the classical Reformulation-Linearization Technique (RLT) for fractional terms. Our RLT variant exploits the underlying mathematical structure of the governing equation, and yields a provably tighter convex relaxation.
4. We construct the simultaneous hull of multiple nonlinear terms that are constrained over a polytope obtained by intersecting a hypercube with mass balance constraints. This yields a tighter convex relaxation than the conventional approach where the nonlinear terms are convexified individually over a box.
5. A key challenge in constructing a valid convex relaxation has been that the denominator of certain fractions in the governing equation can approach arbitrarily close to zero. Using our RLT variant, we construct the first valid relaxation.
6. We leverage powerful mixed-integer programming (MIP) solvers by implementing a discretization-based solution procedure with an adaptive partitioning scheme.

With extensive computational experiments, we demonstrate that the proposed approach outperforms the state-of-the-art in the literature. The formulation can be tailored to other objectives by appending the relevant constraints. Here, we present an extension that identifies the distillation configuration that has the highest thermodynamic efficiency. Finally, we illustrate the practicality of the developed approaches with case studies on crude fractionation and natural gas liquid recovery.

1. INTRODUCTION

Separations are ubiquitous in all chemical, petrochemical and biochemical industries. Separations are required to meet product purity, to reduce the concentration of pollutants below their threshold limits, to separate unreacted reactants from the product stream for recycling and so forth. A plethora of unit operations are available for separation of mixtures. Well-known examples include distillation, permeation through membranes, absorption, adsorption, liquid-liquid extraction, etc. All of the foregoing *molecular separations* are energy intensive, and it is estimated that they can account for 40 – 60% of the total plant costs [1]. The ever increasing push to reduce the environmental impact makes process designers to strive for innovative solutions that improve the overall efficiency of chemical plants.

Of all the available unit operations, distillation has been the predominant choice in industry. The use of distillation can be traced back to as early as 3000 BC. Even today, distillation impacts directly our day to day lives; from gasoline in automobiles to roads on which they are driven, from breweries/wineries to plastic cups in which they are served, and so forth. It is estimated that 90 – 95% [1] of the liquid separations in industry are carried out using distillation. A few well-known and high throughput applications in chemical industry include crude fractionation, natural gas liquid recovery, air separation, etc. With the increased potential to harness shale reserves [2,3], the use of distillation is projected to increase further. Like every other molecular separation, distillation is energy intensive. To put it into perspective, just in the US, distillation consumes 2.4 Quads/year¹, which is roughly 2.5% of the overall energy consumption of the US. Even a 10% reduction in energy consumption would reduce CO₂ emissions by ~18 million tons/year. Given the ubiquity of appli-

¹1 Quad = 1 quadrillion (10¹⁵) BTU

cations and its energy-intensive nature, improving the efficiency of distillation-based separation units is a promising direction towards making chemical industries more sustainable.

1.1 Intricacies of Designing Distillation-based Separation Units

Even after practicing continuous distillation for over a century, designing energy-efficient distillation-based separation units continues to be a challenging task. This is especially true for the separation of multicomponent mixtures. To better understand the reasons, consider the separation of a four-component mixture $ABCD$. Here, A and D correspond to the most and the least volatile components, respectively, and the components are arranged in decreasing order of relative volatilities. Separation of a multicomponent mixture requires a sequence of distillation columns referred as a *configuration* (see Figure 1.1). The energy requirement of a configuration depends on the operating conditions. For example, Figure 1.1(a) shows a set of operating conditions that permit the separation of a given four-component feed (refer the caption for the properties of the feed). The numbers in red correspond to the molar flowrate of the vapor in different sections of the configuration per mole of the feed. By optimizing the operating conditions (see Figure 1.1(b)), the energy consumption of the configuration can be reduced by 14%. However, identifying the optimal operating conditions is challenging, since it requires the solution of a nonconvex mathematical program.

In addition, multiple distillation configurations exist for the separation of a multicomponent mixture. For example, Figure 1.2 shows the so-called *basic column configurations* for the separation of a four-component mixture. Even though all configurations perform the same separation, their energy consumption can differ substantially, as evident from the figure. This is due to the differences in mixing losses within each configuration. Moreover, one can derive additional configurations by replacing condensers and reboilers by two-way vapor-liquid transfer streams, known as *thermal*

couplings (compare Figure 1.2(c) and Figure 1.1). Thermal couplings provide a way to share vapor flows between two or more columns, which can help in reducing the energy consumption. A configuration with one or more thermal couplings is referred as a *thermally coupled configuration*. Therefore, a process designer needs to optimize the separation sequence, condensers and reboilers, and the operating conditions.

Even though powerful process simulators are available at our disposal, they are not suitable for the purpose. This is because, process simulators tend to get trapped in a local solution due to the nonconvex nature of the problem. The only way to ascertain the optimality is to perform an exhaustive sensitivity analysis, which is computationally intensive. Moreover, the computational time and effort needed to identify the optimal configuration can be prohibitive [4]. This is because, the number of admissible configurations (basic and thermally coupled configurations combined) explodes combinatorially with increase in number of components in the feed (see Table 1.1).

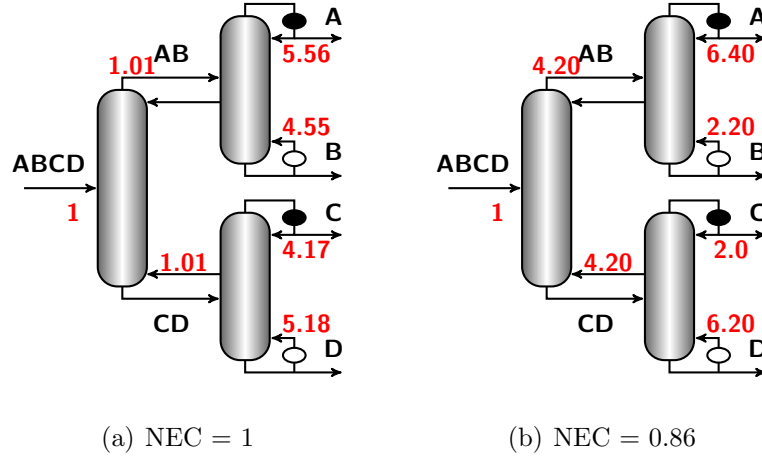


Fig. 1.1. Energy consumption of a configuration depends on the operating conditions. Here, NEC stands for normalized energy consumption. The energy consumption of (a) is used for the normalization. The composition (mole fraction) of the given four-component mixture is $\{z_A, z_B, z_C, z_D\} = \{0.45, 0.05, 0.05, 0.45\}$, and the relative volatilities of the constituent components w.r.t to the heaviest component are $\{\alpha_A, \alpha_B, \alpha_C, \alpha_D\} = \{3.025, 2.75, 1.1, 1\}$

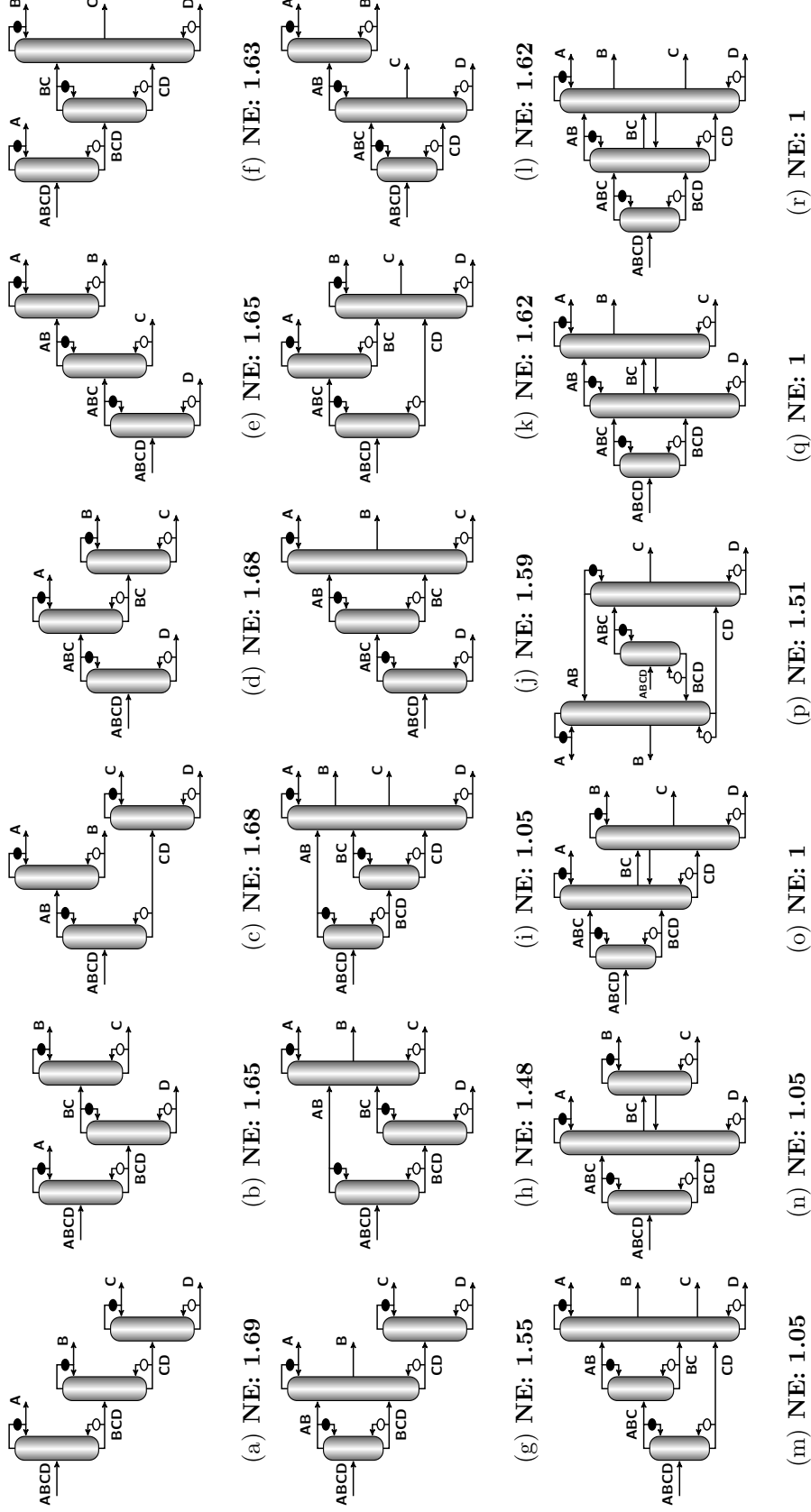


Fig. 1.2. Basic configurations for the separation of a four-component mixture. Here, NE stands for normalized energy consumption, and the energy requirement for configuration (r) is used for normalization. The reported NE values are for the separation of an equimolar mixture of $A \equiv 1$ -pentene, $B \equiv n$ -pentane, $C \equiv 1$ -hexene, and $D \equiv n$ -hexane.

For example, more than half a million and 85 million configurations are available for the separation of six and seven component mixtures. Such mixtures are quite common in petrochemical industries. Owing to this combinatorial complexity, industrial practitioners have always relied on heuristics, past experience and trial-and-error methods for designing distillation units. Quite often, the chosen configuration consumes much more energy than needed to perform the task at hand, but the practitioner has no way to assess this energy penalty because some configurations, inevitably, always remain unexplored. It is, therefore, essential to systematically identify the configuration that is most appropriate for an application.

Table 1.1.
Combinatorial explosion of the choice set. From [5].

Number of Components	Number of Admissible Configurations
4	152
5	6128
6	506,912
7	85,216,192

1.2 Literature Review

Figure 1.3 shows a brief time-line. There are two aspects to the design of distillation-based separation units: (1) Generate the *search space*: the space of admissible distillation configurations (2) Identify the configuration from the search space that optimizes the chosen objective. The enumeration of the entire search space remained an unsolved problem until early 2000s. Sargent and Gaminibandara [6] formulated an initial superstructure (see Figure 1.3) with the intent of capturing the entire search space. However, Agrawal [7] showed that the superstructure did not admit satellite-column

configurations (see Figure 1.3). Instead, the author proposed an alternate superstructure to capture all the configurations. Next, Agrawal [8] proposed a rule-based approach to generate the search space. Later, several formulations based on logical expressions in terms of split variables [9], top-product/bottom-product variables [10] and stream variables [5], have been proposed for enumerating these configurations. However, no clear dominance of the formulations has been established in the literature. Here, we use the tightness of the relaxed problem as the metric to determine the dominance. In this work, we propose two new formulations to generate the search space. The first formulation uses fewer variables, compared to the prior formulations [9, 10], to model the search space. The second formulation dominates every other formulation in the literature, including our first formulation. We prove that this is the best in a well-defined sense.

After modeling the search space, a systematic method is needed to identify the best distillation configuration along with its optimal operating conditions. Until the mid-1980's almost all column sequencing research was confined to sharp-split configurations due to the interest in industry. This makes the problem relatively easier to solve, since distillation calculations can be performed *a priori*. The design specifications, such as reflux ratio, condenser/reboiler duty and so forth, enter only as parameters in the optimization problem [11]. It is only after the successful demonstration of the energy saving potential of sloppy-splits and thermal couplings [12], that the need for inclusion of configurations with these arrangements in the search space became apparent. However, unlike sharp-split configurations, distillation calculations need to be performed online while solving the optimization problem. To make the problem computationally tractable, researchers replaced rigorous tray-by-tray models (also referred to as the MESH equations) involving hundreds of highly nonlinear-nonconvex equations with simplified models that serve as a surrogate. These simplified models provide reasonably accurate estimates of design specifications, but require only a handful of algebraic equations. Optimization problems using these simplified models were first developed for specific configurations [12–14] and later extended to spe-

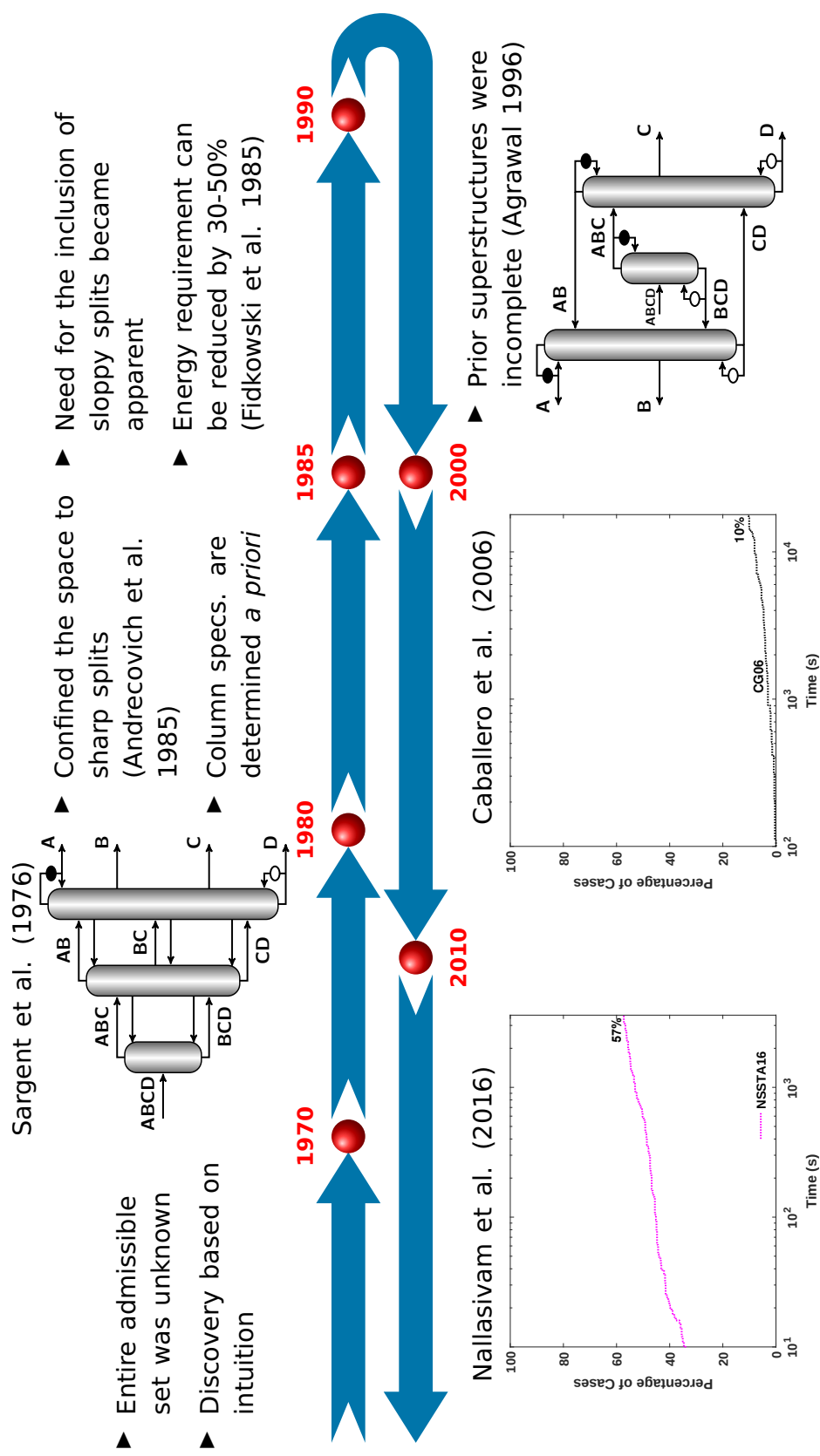


Fig. 1.3. A brief time-line and state-of-the-art

cific groups of configurations. For example, in [15], the authors considered a three-component Petlyuk-type configuration and formulated an optimization problem to determine optimal flows using Underwood’s method. They were able to derive an analytical solution for the vapor duty, and found that the optimal region was flat over a range of flows. Later, researchers have extended the approach to some non-FTC (non-Fully Thermally Coupled) configurations, especially the side-stripper and the side-rectifier configurations [12, 16]. Nevertheless, these studies either focused on specific configurations or considered certain regions of search space, leaving the structured search of the entire search space an open problem.

To the best of our knowledge, Caballero and Grossmann [17] were the first to perform a structured search over the entire search space. The authors proposed a new superstructure, where the component, vapor and liquid flows are bypassed when the corresponding pseudo-column is absent. This superstructure forms the building block in their subsequent contributions. The authors used a combination of State Task Network (STN) and State Equipment Network (SEN) to formulate an MINLP to identify the configuration that minimizes the total cost [17]. Unfortunately, the resulting MINLP could not be solved to global optimality. Moreover, even the local NLP solvers often failed to converge to a feasible solution due to singularities that arise in the model when certain splits are no longer a part of the configuration. When the local solvers found a feasible point, it was often a local optimum that was significantly inferior relative to the global solution [17]. To circumvent these issues, the authors architected an algorithm by suitably modifying the logic-based outer-approximation scheme. Although the resulting algorithm successfully found feasible configurations, it could still not guarantee that a global optimum was found. Later, [9] introduced a two-step method, where first they determined a completely thermally coupled configuration and then located the heat exchangers optimally. Clearly, this hierarchical approach does not guarantee global optimality either. In subsequent works, the authors proceeded to explore possible heat integration [9], rearrangement of column sections [18], and dividing wall columns [19]. More recently, inspired by

[11], [20] formulated a Mixed-Integer Linear Program, where they used FUG method to, a priori, optimize vapor traffic and heat loads for each split.

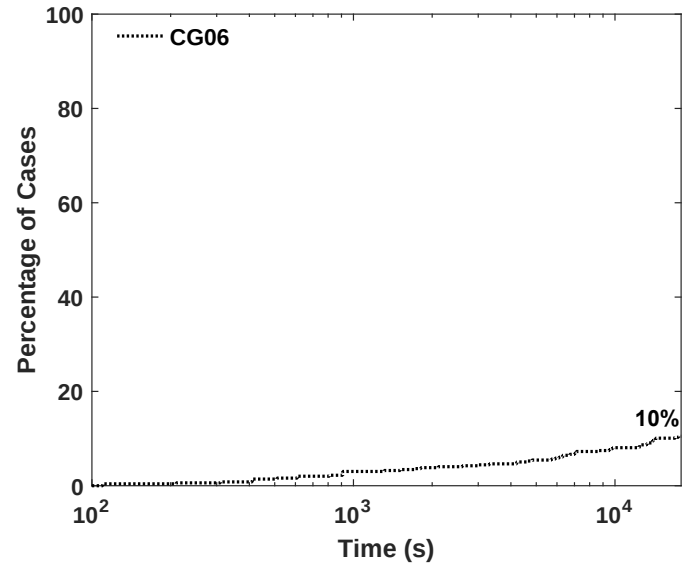
In the meanwhile, [10] introduced a new MINLP formulation to determine the configuration with the lowest vapor duty. They have succeeded in solving three and four component test cases to global optimality using BARON [21]. Although, their formulation was not able to solve a five-component MINLP directly, it was able to globally optimize the vapor flow for certain configurations, once their structure was specified by fixing the binary variables. This observation led to the enumeration-based approach of [22]. Here, the authors formulated a Nonlinear Program (NLP) for each configuration in the search space that was then solved using the global solver, BARON. The proposed approach is amenable to parallelization, but it remains computationally inefficient for six and more number of components in the process feed.

1.2.1 State-of-the-art

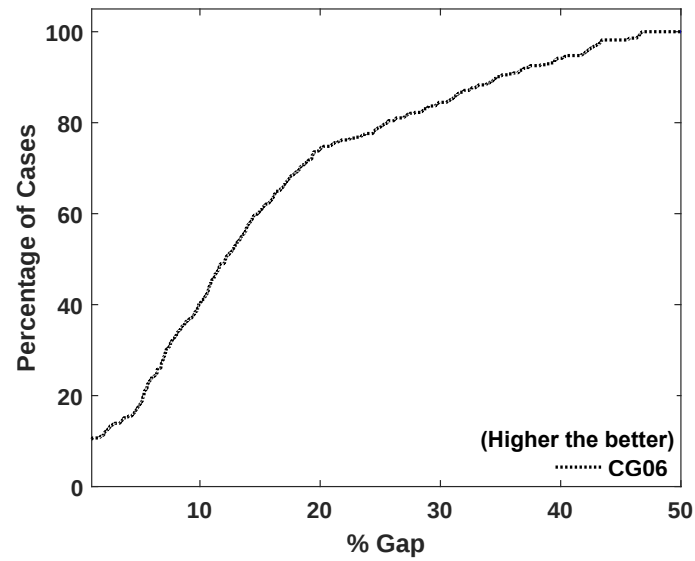
Here, we show that the state-of-the-art approaches cannot reliably identify the optimal distillation configuration. To the best of our knowledge, the MINLP formulation in [9] remains the state-of-the-art. To test the efficacy of their model, we considered a test set of 496 cases (see §A) that are representative of different types of separations. We solve their model for all the 496 test cases using BARON. The results are shown in the *performance profile* in Figure 1.4(a). Here, we plot the cumulative percentage of cases that have been solved to 1%-optimality gap as a function of time. We observe that only 10% of the cases could be solved within five hours. A large duality gap, defined as

$$\% \text{ Gap} = \left(1 - \frac{\text{Best known lower bound}}{\text{Best known upper bound}} \right) \times 100, \quad (1.1)$$

remain in several unsolved cases (see Figure 1.4(b)). This demonstrates that the problem is challenging, since only 10% of the cases can be solved with the best formulation in the literature and with a state-of-the-art global solver. While this



(a)



(b)

Fig. 1.4. (a) Performance profile for the MINLP model in [9] (b) Plot illustrating the remaining duality gap at the end of five hours.

work made substantial contribution towards developing a systematic approach, more work is needed to reliably solve this hard problem.

Next, we show that the state-of-the-art cannot reliably identify the optimal operating conditions even for a given configuration. We consider the formulation in Nallasivam et al. [22]. Here, the authors formulate a nonlinear program for each configuration in the search space. Their approach overcomes the challenge due to the nonconvexity associated with discrete variables. Nevertheless, the problem remains challenging due to the presence of nonconvex fractional terms. Indeed, they [22] showed that local solvers can get trapped in a sub-optimal solution that requires a much higher energy than the global optimum. Therefore, we use the global solver BARON to solve the nonlinear program. We consider the five-component fully thermally coupled configuration (the five-component analog of Figure 1.2(r)). We determine the optimal operating conditions of this configuration for each of the 496 cases in the test set (see §A), by solving the nonlinear program. The results are shown in the performance profile in Figure 1.5. As before, we plot the cumulative percentage of cases that have been solved to 1%-optimality gap as a function of time. We observe that, even for a given configuration, the current approach can solve only 57% of the cases in the test in one hour. Clearly, there is a need to develop better formulations and solution procedures to reliably identify the best distillation configuration for a given application. This thesis develops such a formulation and solution procedure. The rigorous approach presented here can reduce the energy consumption and consequent greenhouse gas emissions by the separation process.

1.3 Key Contributions

The main contributions are organized as chapters, and the highlights of each chapter are summarized below.

1.3.1 Chapter 3

This chapter is based on [23]. Here, we propose a novel MINLP that is formulated to identify the distillation configuration requiring least vapor duty. The highlights of

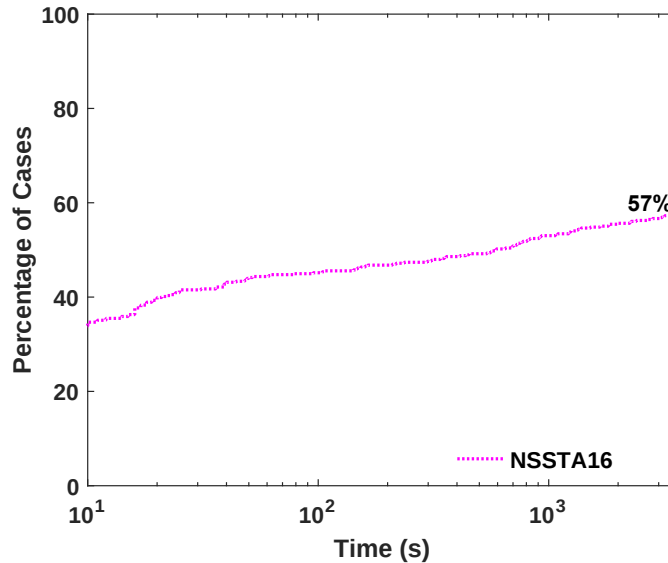


Fig. 1.5. Performance profile for the nonlinear program proposed in [22]

this chapter are as follows. First, we introduce a new notation to represent streams and heat exchangers, which allows for a more succinct and cleaner presentation. Using this notation, we describe a new formulation for the search space of configurations detailed in Shah and Agrawal [5]. Second, our formulation restricts the composition of distillate and residue to more closely represent the reality. For example, consider the model presented in [9] and, in particular, the example in the MINLP library [24]. In the optimal solution for this example, the recovery of component B in the distillate of split $ABCD/BCDE$ is higher than that of the heavier components, C and D , while, in reality, the recovery of a lighter component in the distillate must be at least as much as that of a heavier component. To address this issue, we propose a new formulation of Underwood constraints to avoid such occurrences. Third, we derive additional cuts by exploiting properties of Underwood equations. These cuts play a vital role yielding tighter relaxations, which help expedite the convergence of branch-and-bound. Finally, we demonstrate the efficacy of our approach by presenting various four, five, and six component examples.

1.3.2 Chapter 4

This chapter is based on [25]. Here, we describe a novel formulation, relaxation and discretization based solution procedure to reliably solve the MINLP. The highlights of this chapter are as follows. First, we introduce a new formulation, to model the space of admissible configurations. We show that this formulation is provably tighter than previous formulations in the literature [9, 10, 23]. More importantly, we show that our formulation is contained in the convex hull of various important substructures. Second, we adapt the classical Reformulation-Linearization Technique (RLT) [26] to obtain a family of cuts for fractional terms. These cuts are especially useful, as they exploit the mathematical structure of the governing equations. The techniques we propose are general, and can be used to relax other optimization problems involving fractional terms. Third, we employ simultaneous convexification techniques to construct convex hull of multiple nonlinear terms that are constrained over a polytope obtained by intersecting a hyperrectangle with material balance equations. This approach results in relaxations that are strictly tighter than conventional relaxations, which are obtained by relaxing nonconvex terms individually over a box. Fourth, a key challenge in deriving valid relaxations for the distillation configuration design problem has been that certain terms that appear in the denominator cannot be away from zero. Prior works have imposed arbitrary lower bounds on this denominator [23, 24]. The cuts derived using our RLT variant enable us to infer finite upper bounds on fractions, which allow us to construct provably valid relaxations. Fifth, we leverage powerful Mixed-Integer Linear/Second-order Cone Programming (MIL/SOCP) solvers by replacing convex relaxations with piecewise relaxations, obtained by adaptively discretizing the domain of specific decision variables and constructing convex relaxation within each partition using disjunctive programming techniques. Lastly, through computational experiments, we demonstrate the efficacy of our approach by comparing its performance with other state-of-the-art techniques.

1.3.3 Chapter 5

This chapter is based on [27]. Here, we propose a novel MINLP that is formulated to identify the distillation configuration that has the highest thermodynamic efficiency. The highlights of this chapter are listed in the following. First, we describe a formulation to model the space of admissible configurations that is tighter than the prior formulations in the literature. This formulation is obtained by projecting the formulation proposed in Chapter 4 onto a lower dimensional space. Second, we use the ideas introduced by [28] and [29] to formulate the model in a way that it does not depend on temperature of streams explicitly. This way, the results obtained from the model hold for any mixture, irrespective of sub-ambient or above-ambient operation, provided the mixture has the same composition and relative volatilities for constituent components. However, the model in its default form has several nonlinear nonconvex equations. To simplify the model, we reformulate it using a simple variable elimination technique. The reformulated model simplifies exergy loss calculations substantially by reducing a system of nonlinear equations that model material balance and vapor-liquid equilibrium to a single equation, which we refer as *exergy constraint*. Third, we describe the properties satisfied by exergy constraints, and exploit them in deriving additional cuts to the problem. Fourth, we use the MINLP formulation to identify attractive configurations for recovery of Natural Gas Liquids (NGLs) from shale gas. Finally, we investigate the efficacy of the model to changes in process parameters on several four and five component mixtures taken from the literature.

2. PRELIMINARIES

2.1 The Distillation Process

When a mixture of components with different volatilities is boiled, the vapor produced is rich in more volatile (or *light*) components, while the residual liquid is enriched in less volatile (or *heavy*) components. Distillation exploits this characteristic to separate mixtures of two or more components. Industrial distillation is carried out in a staged-tower/column (see Figure 2.1), where each stage establishes liquid-vapor contact for mass transfer. The *feed* (mixture of components) is introduced at an intermediate location of the column. The *sections* above and below the feed stream are known as *rectifying* and *stripping sections*, respectively. Conventional columns have a condenser (*resp.* reboiler) at the top (*resp.* bottom) which condenses (*resp.* vaporizes) the vapor (*resp.* liquid), and feeds a portion of it back to the column, known as *liquid* (*resp.* *vapor*) *reflux*. The liquid flowing from the top to bottom strips away heavy components from the vapor, while the vapor flowing from bottom to top gets enriched with lighter components. The net outflow from the rectifying and stripping sections, respectively, are known as *distillate* and *residue*. In short, distillation enriches the distillate with light components, and the residue with heavy components.

Remark 2.1. The recovery of a lighter component in distillate (ratio of component flowrate in distillate to flowrate in feed) is higher than the recovery of a heavier component, and the converse is true for residue [30].

Finally, for a given product composition, there is a threshold vapor flow in each section, below which the products do not reach the desired composition. To determine this threshold vapor flow, we use the classic Underwood method [31], which is described in the next section.

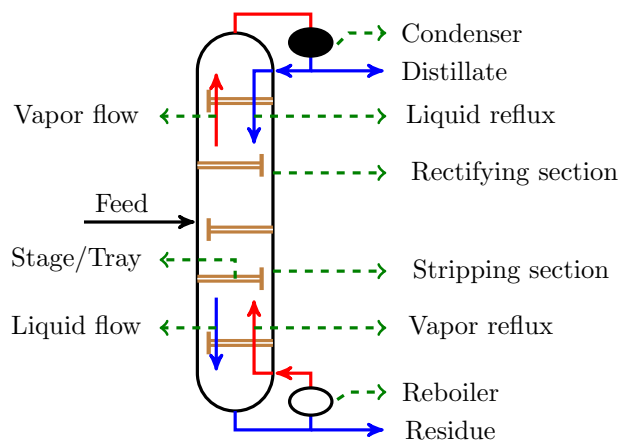


Fig. 2.1. Schematic of a distillation column

2.2 A Note on Underwood Constraints

In his seminal paper, [31] derived a shortcut method to compute the minimum vapor requirement of a distillation column. Before we describe the constraints, we emphasize a few key concepts like the definition of key components and the degrees of freedom. These concepts are crucial for the problem formulation. The optimization problem can yield physically infeasible solutions when modeled flexibility exceeds actual flexibility in how a system operates. For example, recovery of a heavier component in the distillate cannot be higher than that of a lighter component. A solution that does not satisfy this property violates the physics of the distillation process.

A *light key* (resp. *heavy key*) is defined as the lightest (resp. heaviest) component that distributes, or would distribute between distillate and residue if vapor flow was decreased slightly. Components lighter (resp. heavier) than the light key (resp. heavy key) are completely recovered in the distillate (resp. residue), with an insignificant amount in the residue (resp. distillate). Regardless of how many components are in the feed stream or which product streams are produced by the split, a distillation column has exactly two degrees of freedom. Therefore, the operation of the column is completely determined by specifying two values, often the recovery of light and heavy keys in one of the product streams. Specifying these quantities fixes the other

process variables such as minimum vapor requirement, and recoveries of remaining components in the product streams. Alternatively, the vapor flow and recovery of either of the keys can be specified to completely describe the system. In either case, no more than two values may be specified independently. We refer the reader to [32] for a detailed discussion on the degrees of freedom.

We now describe the Underwood method that can be used to limit the degrees of freedom to two and, thus, remove excess degrees of freedom from the model. The first step in Underwood method determines Underwood roots (θ_q) by solving the feed equation.

$$\sum_{p=1}^n \frac{\alpha_p f_p^{in}}{\alpha_p - \theta_q} = V^{in} \quad 1 \leq q \leq n - 1 \quad (2.1)$$

where α_p is the relative volatility of component p with respect to the heaviest component in the process feed and, it can be shown that the q^{th} Underwood root θ_q satisfies $\alpha_{q+1} < \theta_q < \alpha_q$. Here, f_p^{in} and V^{in} correspond to the net inflow of component p and vapor, respectively. In the second step, the roots that lie between the relative volatilities of the key components, called *active roots*, are determined. These roots are then used to compute minimum vapor requirement. The remaining roots are called *inactive roots*.

Halvorsen and Skogestad [33, 34] introduced a V_{min} diagram to visualize the minimum vapor requirement. In Figure 2.2 we show such a V_{min} diagram for feed conditions detailed in the figure's caption. The horizontal axis graphs the net distillate flowrate, while the vertical axis graphs the minimum vapor required in the rectifying section Υ^{rs} , for various splits. Note that, as long as a point is in or on the boundary of the curve (i) $0 - P_{AB} - 1$, θ_1 is active (ii) $0 - P_{BC} - 1$, θ_2 is active (iii) $0 - P_{CD} - 1$, θ_3 is active and (iv) $0 - P_{DE} - 1$, θ_4 is active. As an example, we consider the split $ABCD/BCDE$. The region shaded in Figure 2.2 shows the feasible region for the split. Clearly, θ_2 and θ_3 are active throughout the shaded region, while θ_1 (resp. θ_4) is active only along the line joining $P_{AD} - P_{AE}$ (resp. $P_{AE} - P_{BE}$). When θ_1 (resp. θ_4) is active, component A (resp. E) is the light (resp. heavy) key for the split.

Everywhere else, components B and C are the light and heavy keys. Therefore, the recoveries of the distributing components are constrained by the following relation:

$$\sum_{p=1}^4 \frac{\alpha_p f_p^{\text{rs}}}{\alpha_p - \theta_2} = \sum_{p=1}^4 \frac{\alpha_p f_p^{\text{rs}}}{\alpha_p - \theta_3} = \Upsilon^{\text{rs}} \quad (2.2)$$

or, given the vapor balance, in terms of component flows in the stripping section

$$-\sum_{p=2}^5 \frac{\alpha_p f_p^{\text{ss}}}{\alpha_p - \theta_2} = -\sum_{p=2}^5 \frac{\alpha_p f_p^{\text{ss}}}{\alpha_p - \theta_3} = \Upsilon^{\text{ss}} \quad (2.3)$$

where f_p^{rs} , f_p^{ss} , Υ^{rs} , and Υ^{ss} denote, respectively, the net flow rate of component p in the rectifying and stripping sections, minimum vapor requirement in the rectifying and stripping sections. Observe that, altogether there are four unknowns: three unknown recoveries, namely those for B , C , and D and the minimum vapor flow in the rectifying section. The two equalities in (2.2) or (2.3) ensure that, as desired, the system has only two degrees of freedom. If these are relaxed, as in [9], the system has more flexibility than allowable and the optimal solution may not be physically feasible. Finally, we mention that the minimum vapor constraints can also be imposed for inactive roots, albeit after relaxing the equality to inequality as below:

$$\begin{aligned} \sum_{p=1}^4 \frac{\alpha_p f_p^{\text{rs}}}{\alpha_p - \theta_q} &\leq \Upsilon^{\text{rs}} \quad q = \text{Inactive root} \\ -\sum_{p=2}^5 \frac{\alpha_p f_p^{\text{ss}}}{\alpha_p - \theta_q} &\leq \Upsilon^{\text{ss}} \quad q = \text{Inactive root} \end{aligned} \quad (2.4)$$

Validity of these constraints can be shown mathematically. They can also be inferred from the $V_{\min}$ diagram (cf. Figure 2.2).

2.3 Alternative Interpretation of Relative Volatilities

Shortcut models, such as Underwood method, require relative volatilities of all components for computing minimum vapor reflux and recoveries of distributing components in distillate/residue. Conventionally, relative volatilities are defined as the ratio of equilibrium constants (K -values), or alternatively as the ratio of saturation

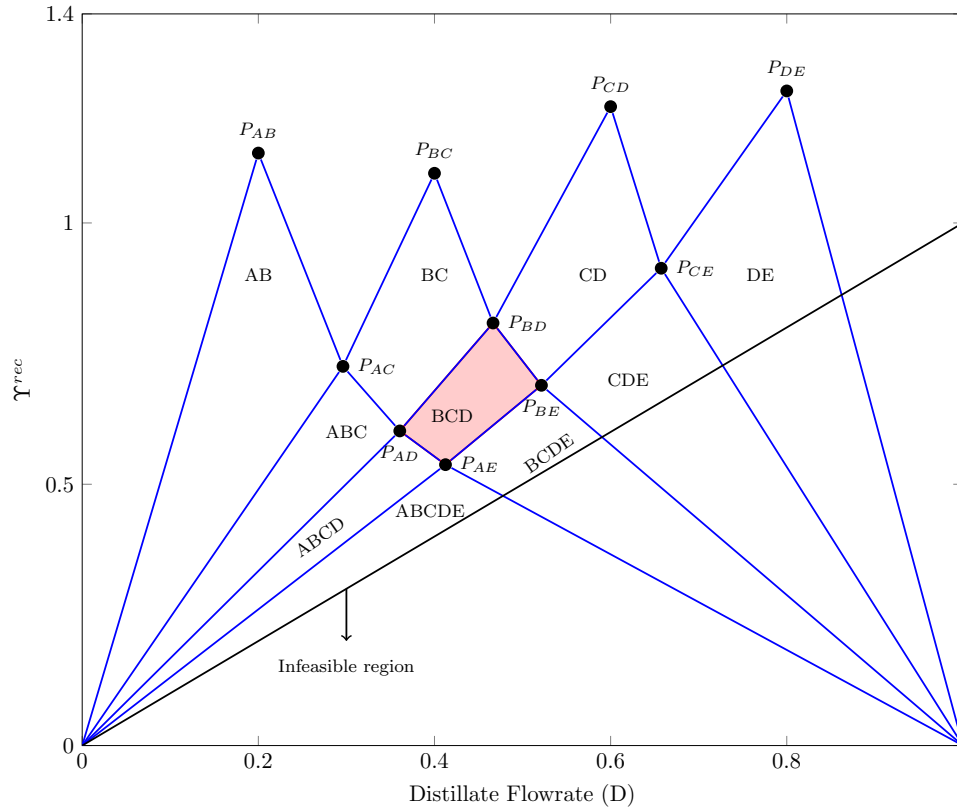


Fig. 2.2. V_{min} diagram showing the feasible region (shaded region) for the split $ABCD/BCDE$. Here, the feed composition is $\{0.2, 0.2, 0.2, 0.2, 0.2\}$, relative volatilities are $\{9, 6, 3.5, 2, 1\}$, and thermal quality = 1. The values are taken from [32]

pressures, of a more volatile component w.r.t a less volatile component. In a distillation column, the temperature varies substantially from top to bottom, in particular, for *easy separations*. Consequently, K -values, and in turn relative volatilities, vary substantially across the column. This behavior challenges the assumption of constant relative volatility inherent in Underwood method [31]. Nevertheless, our experience suggests that the minimum vapor reflux and pinch compositions predicted by Underwood method agrees well with those from solving MESH equations on Aspen Plus. This arises two questions: (i) why and when does the constant relative volatility assumption hold? (ii) how to choose relative volatilities required for Underwood

method? We propose a plausible explanation to the first question, and suggest a few ways to choose relative volatilities.

To answer the first question, we describe an alternative interpretation for relative volatilities $\{\alpha_p\}_{p=1}^N$. In his seminal paper, [31] considered the following model for vapor-liquid equilibrium (VLE):

$$y_p = \frac{\alpha_p x_p}{\sum_r \alpha_r x_r}, \quad \forall \quad p \in \{1, \dots, N\} \quad (2.5)$$

where x_p and y_p denote the molar fraction of component p in liquid and vapor phases, and α_p is the relative volatility of component p w.r.t to the least volatile component. Note that, by definition, $\alpha_N = 1$. The VLE model in (2.5) is derived from the conventional definition of relative volatility

$$\alpha_p = \frac{K_p}{K_N} = \frac{y_p/x_p}{y_N/x_N} \quad (2.6)$$

We interpret (2.5) as a *surrogate model* for the true VLE, and $\{\alpha_p\}_{p=1}^N$ as parameters for the surrogate model. The choice of parameters $\{\alpha_p\}_{p=1}^N$, which will be discussed shortly, must be such that the surrogate is as close to the true VLE as possible. In such cases, the estimates of vapor composition provided by the surrogate model agree well with the true vapor composition. As a result, for a given reflux ratio and distillate flowrate, the simplified tray-by-tray calculations considered by [31] will lead to a closer approximation of rigorous MESH equations. This is a plausible explanation why Underwood method is reliable despite the assumption of constant relative volatility. In contrast, when the surrogate model in (2.5) fails to capture the true VLE accurately, like in azeotropic systems, the estimates from the Underwood method will not be reliable, especially when the range of compositions fall in the region where there is a substantial deviation between the surrogate and the true VLE.

Now, we address the second question concerning the choice of $\{\alpha_p\}_{p=1}^N$. We determine $\{\alpha_p\}_{p=1}^N$ such that the surrogate in (2.5) closely approximates the true VLE in least-square error sense. Towards this, we consider a sample of points $\mathcal{M}$ spanning

the entire liquid composition space. We determine the respective vapor composition using the true VLE. Then, we formulate a least-squares error problem shown below

$$\text{Min.}_{\{\alpha_p\}_{p=1}^{N-1}} \sum_{m=1}^{|\mathcal{M}|} \sum_{p=1}^{N-1} \left(y_{p,m} - \frac{\alpha_p x_{p,m}}{\sum_{r=1}^n \alpha_r x_{r,m}} \right)^2 \quad (2.7)$$

Note that we did not include α_N as a decision variable, since $\alpha_N = 1$. After computing the optimal $\{\alpha_p\}_{p=1}^n$, it is necessary to check if the surrogate closely approximates the true VLE. In some cases, it is possible that the unconstrained optimization might lead to a solution which gives a closer approximation in some regions of the composition space, while a poorer approximation in some other regions. In such cases, constraints restricting the relative error in vapor composition to not exceed a specified value may be added to the optimization problem in (2.7).

3. MINIMIZATION OF VAPOR DUTY

In this chapter, we formulate a novel Mixed-Integer Nonlinear Program (MINLP) to identify the top $\mathbb{K}$ configurations that require least total vapor duty for the separation of a given non-azeotropic mixture. The vapor duty, defined as the sum of vapor flows generated by all the reboilers, serves as a proxy for energy consumption. Thus, by minimizing the vapor duty, we find the configuration that requires the least amount of utilities, and thus the operating cost. In addition, the vapor duty indirectly affects the capital cost, as the column diameter depends on the internal vapor flows. Given our focus on non-azeotropic mixtures, we use Underwood’s method to obtain a reasonable estimate of the minimum vapor requirement for each split. Despite the resulting simplifications relative to a detailed tray-by-tray calculation, the MINLP remains challenging to solve to global optimality because of the nonconvexities in the Underwood equations and combinatorial complexity of the search space. The current state-of-the-art in this area is the explicit enumeration based algorithm of Nallasivam et al. [22]. However, compared to implicit enumeration employed in this work, the former approach is computationally inefficient and does not scale well with increase in number of components in the mixture. The existing implicit enumeration based implementations [17] do not guarantee global optimality due to problem decomposition. The approach presented in this chapter can be used to solve the problem to ϵ -global optimality with off-the-shelf solvers.

The rest of the chapter is organized as follows: In §3.1, we present our MINLP formulation for identifying the configuration with minimum overall vapor duty. The optimal solution of the MINLP may be hard to implement for reasons that are not modeled and, although optimal to the model, the solution may not even be truly optimal once the simplifying assumptions that were made to derive a tractable formulation are relaxed. Moreover, the designers often have other considerations in mind

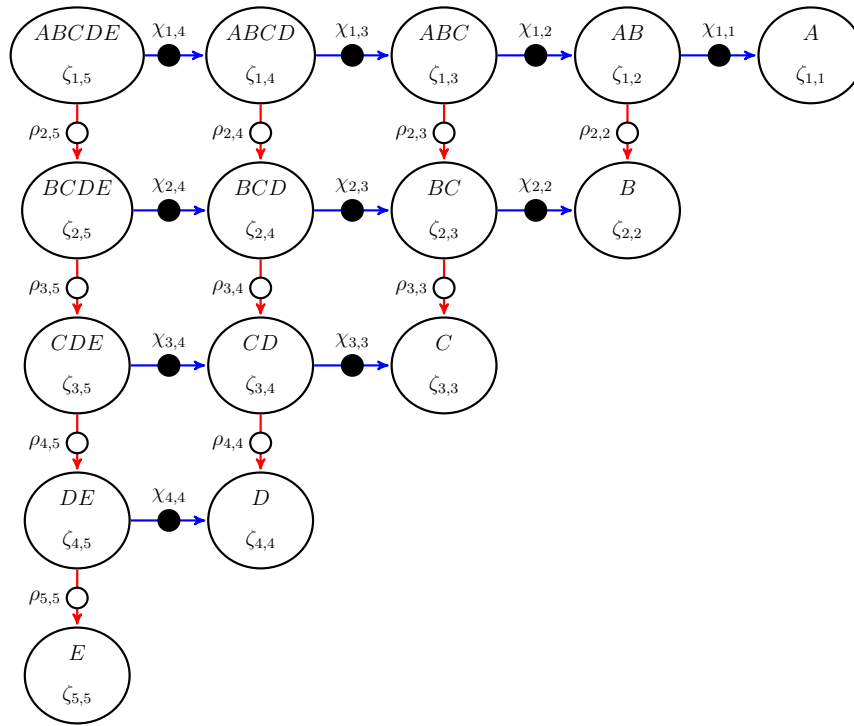


Fig. 3.1. Five-component system. Here, filled and unfilled circles denote condensers and reboilers, red and blue colored arrows denote the rectifying and stripping sections.

such as heat-integration with the remaining plant and environmental impact. Therefore, it is desirable to have a handful of configurations, on which rigorous tray-by-tray calculations are performed before the appropriate configuration is identified. To provide sufficiently many alternatives for downstream evaluation, in §3.2, we describe a simple method to systematically determine the top $\mathbb{K}$ solutions for the problem. Further, in §3.2, we consider a few case studies and solve them to global optimality. We conclude the chapter with a few remarks on the applicability and extensibility of the model in §3.3.

3.1 Problem Formulation

Before formulating the problem, we give a few definitions and discuss the notation we use to describe our constructions. To allow for a general zeotropic mixture, we denote streams as $ABCDE$, where each letter denotes a component. Alternatively, we will also use numbers to denote components: $A \equiv 1, B \equiv 2, \dots, E \equiv 5$. In a stream, components are arranged in a decreasing order of relative volatility with the first and the last components being, respectively, the most and the least volatile components. It is these extreme components that we use to denote a stream. In particular, if the most (resp. least) volatile component is component i (resp. j), then the stream is denoted as the couplet $[i, j]$. We remark that Shah and Agrawal [5] introduced a similar, albeit slightly different, notation, where they placed each stream in an upper triangular matrix, and used the coordinates of this position to refer to the stream. Since this placement depends on the number of components of the feed stream, the translation from their notation to stream labels requires this number. Instead, the translation with our notation is clearly independent of this number and can be done in a straightforward fashion. For example, $[2, 3]$ always represents the stream BC , regardless of the number of components n in the process feed $[1, n]$. In a similar manner, we denote heat exchangers by a couplet (i, j) whose coordinates are obtained from the extreme components in the associated stream. For example, the condenser and the reboiler of the column that splits $[1, 5]$ into submixtures $[1, 4]$ and $[2, 5]$ are respectively associated with these product streams (see Figure 3.1). Therefore, by condenser (i, j) (resp. reboiler (i, j)), we mean the heat exchanger through which stream $[i, j]$ is withdrawn as distillate (resp. residue). Splits will be represented as $ABCD/BCDE$ which signifies that $ABCD$ and $BCDE$ are produced as distillate and residue from $ABCDE$.

A *submixture* is defined as any intermediate stream that arises during the separation of the process feed into pure components/desired products. As an example, we refer to the five component system shown in Figure 3.1. Here, all the streams except

$\{ABCDE, A, B, C, D, E\}$ are submixtures. Further, we define *top parents* (resp. *bottom parents*) of a submixture as the streams, which if fed to a column could produce this submixture through the rectifying (resp. stripping) section. Analogously, we define *top children* (resp. *bottom children*) of a submixture as the streams which could be produced from the rectifying (resp. stripping) section of a column with this submixture feed. For any submixture, the top (resp. bottom) parents can be identified by traversing horizontally to the left (resp. vertically upwards) from this submixture, and the top (resp. bottom) children can be identified by traversing horizontally to the right (resp. vertically downwards). For any stream $[i, j]$, the top parents are obtained by increasing j : $\{[i, j + 1], [i, j + 2], \dots, [i, n]\}$, while the top children are obtained by reducing j : $\{[i, j - 1], [i, j - 2], \dots, [i, i]\}$. Similarly, the bottom parents of $[i, j]$ are obtained by reducing i : $\{[i - 1, j], [i - 2, j], \dots, [1, j]\}$, while the bottom children are obtained increasing i : $\{[i + 1, j], [i + 2, j], \dots, [j, j]\}$. As an example, the top (resp. bottom) parents of $[3, 4]$ are $\{[3, 5]\}$ (resp. $\{[1, 4], [2, 4]\}$), and the top (resp. bottom) children are $\{[3, 3]\}$ (resp. $\{[4, 4]\}$) (see Figure 3.1).

We assume that the final products, which are the pure components, are drawn in saturated liquid state. Nevertheless, by slightly modifying the superstructure (see Figure 3.2), we may withdraw final products in any desired state. We also assume that the molar overflows, the latent heat of vaporization, and the relative volatility remain constant throughout each column section of the configuration. These assumptions are implicit in our use of the Underwood method. The definition of the sets and variables used in the formulation are summarized in Tables 3.1 and 3.2, respectively. We split the set of submixtures into three sets, where $\mathcal{S}_1$ (resp. $\mathcal{S}_3$) denote the set of submixtures in the top row (resp. first column) of the superstructure representation in Figure 3.1, while $\mathcal{S}_2$ denotes the set of remaining submixtures. We represent each split using its feed stream, which is either a submixture or the process feed.

Table 3.1.
Definition of sets

Set	Symbol	Definition	5-component example
Submixtures	$\mathcal{S}_1$	$\{[1, j]\}_{j=2}^{n-1}$	$\{[1, 2], [1, 3], [1, 4]\}$
	$\mathcal{S}_2$	$\{[i, j]\}_{i=2, j=i+1}^{i=n-2, j=n-1}$	$\{[2, 3], [2, 4], [3, 4]\}$
	$\mathcal{S}_3$	$\{[i, n]\}_{i=2}^{n-1}$	$\{[2, 5], [3, 5], [4, 5]\}$
	$\mathcal{U}$	$\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$	$\{[1, 2], [1, 3], [1, 4], [2, 3], [2, 4], [2, 5], [3, 4], [3, 5], [4, 5]\}$
Splits	$\mathcal{P}$	$\{[1, n] \cup \mathcal{U}\}$	$\{[1, 5]\} \cup \mathcal{U}$
Streams	$\mathcal{T}$	$\mathcal{P} \cup \{[i, i]\}_{i=1}^n$	$\mathcal{P} \cup \{[1, 1], [2, 2], [3, 3], [4, 4], [5, 5]\}$
Condensers	$\mathcal{C}$	$\{(i, j) : 1 \leq i \leq j \leq n-1\}$	$\{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4), (4, 4)\}$
Reboilers	$\mathcal{R}$	$\{(i, j) : 2 \leq i \leq j \leq n\}$	$\{(2, 2), (2, 3), (2, 4), (2, 5), (3, 3), (3, 4), (3, 5), (4, 4), (4, 5), (5, 5)\}$

Table 3.2.: Domain of indices for the variables. For $\zeta_{i,j}$, $[i, j] \in \mathcal{T}$; For $\chi_{i,j}$ and $FC_{i,j}$, $(i, j) \in \mathcal{C}$; For $\rho_{i,j}$ and $FR_{i,j}$, $(i, j) \in \mathcal{R}$; For $U_{i,j+1}^{\text{rs}}$ and $K_{i,j+1}^{\text{rs}}$, $[i, j] \in \mathcal{S}_1 \cup \mathcal{S}_2$; For $U_{i-1,j}^{\text{ss}}$ and $K_{i-1,j}^{\text{ss}}$, $[i, j] \in \mathcal{S}_2 \cup \mathcal{S}_3$. For the remaining decision variables, $[i, j] \in \mathcal{P}$.

Parameters	
n	Number of components in the feed
α_p	Relative volatility of component p w.r.t the heaviest component

continued on next page

Table 3.2.: *continued*

F_p	Flowrate of component p in the feed
Φ	Liquid fraction in the process feed (thermal quality)
V^U	Upper bound on vapor flow rate
Notation	
$[i, j]$	Submixture with i and j as the lightest and heaviest components
$\{[i, l]\}_{l=j+1}^n$	Top parents of submixture $[i, j]$
$\{[k, j]\}_{k=1}^{i-1}$	Bottom parents of submixture $[i, j]$
$\{[i, r]\}_{r=i}^{j-1}$	Top children of submixture $[i, j]$
$\{[s, j]\}_{s=i+1}^j$	Bottom children of submixture $[i, j]$
$Q_{i,j}$	Pseudocolumn performing the separation of stream $[i, j]$
Decision Variables	
$\zeta_{i,j}$	Binary variable indicating the presence/absence of stream $[i, j]$
$\chi_{i,j}$	Binary variable indicating the presence/absence of condenser (i, j)
$\rho_{i,j}$	Binary variable indicating the presence/absence of reboiler (i, j)
$\{f_{i,j,p}^{\text{rs}}\}_{p=i}^{j-1}$	Net molar flow of component p in the rectifying section of $Q_{i,j}$
$\{f_{i,j,p}^{\text{ss}}\}_{p=i+1}^j$	Net molar flow of component p in the stripping section of $Q_{i,j}$
$V_{i,j}^{\text{rs}}$	Vapor flowrate in the rectifying section of $Q_{i,j}$
$V_{i,j}^{\text{ss}}$	Vapor flowrate in the stripping section of $Q_{i,j}$
$L_{i,j}^{\text{rs}}$	Liquid flowrate in the rectifying section of $Q_{i,j}$
$L_{i,j}^{\text{ss}}$	Liquid flowrate in the stripping section of $Q_{i,j}$
$FC_{i,j}$	Molar flowrate in condenser (i, j)
$FR_{i,j}$	Molar flowrate in reboiler (i, j)
$U_{i,j+1}^{\text{rs}}$	Vapor in-flow into $Q_{i,j}$ from $Q_{i,j+1}$
$K_{i,j+1}^{\text{rs}}$	Liquid out-flow from $Q_{i,j}$ to $Q_{i,j+1}$
$U_{i-1,j}^{\text{ss}}$	Vapor out-flow from $Q_{i,j}$ to $Q_{i-1,j}$

continued on next page

Table 3.2.: *continued*

$K_{i-1,j}^{ss}$	Liquid in-flow into $Q_{i,j}$ from $Q_{i-1,j}$
$\{\theta_{i,j,q}\}_{q=i}^{j-1}$	Underwood root of $Q_{i,j}$ satisfying $\alpha_{q+1} < \theta_{i,j,q} < \alpha_q$
$\Upsilon_{i,j}^{rsz}$	Minimum vapor flow required in the rectifying section of $Q_{i,j}$
$\Upsilon_{i,j}^{ssz}$	Minimum vapor flow required in the stripping section of $Q_{i,j}$
$\{S_{i,j,q}^z\}_{q=i+1}^{j-2}$	Slack variable associated with $\theta_{i,j,q}$

3.1.1 Objective Function

The objective is to minimize the net vapor generated by all the reboilers that leads to the following formulation:

$$(H): \text{Minimize } \sum_{i=2}^n \sum_{j=i}^n FR_{i,j}, \quad (H1)$$

where $FR_{i,j}$ denotes the vapor flow generated by reboiler (i, j) . The minimization in (H) is subject to constraints that enforce mass-balance, compute vapor flows using Underwood equations, and define the search-space of configurations.

3.1.2 Search Space Formulation

We begin by describing the constraints that define the space of regular-column configurations. For this, we rely on the method proposed by Shah and Agrawal [5] (SA method) and adapt it to our notation. We will describe these constraints in two sets; constraints that (i) govern the presence/absence of streams and (ii) govern the presence/absence of heat exchangers.

Constraints on stream variables

We use binary variables $\zeta_{i,j}$, $[i, j] \in \mathcal{T}$ to denote the presence/absence of a stream $[i, j]$, i.e., $\zeta_{i,j} = 1$ denotes stream $[i, j]$ is present while zero denotes its absence. SA

method requires three checks to ensure correct streams are present in the formulation. These ensure (i) the presence of process feed and final products (pure components), (ii) the presence of a parent, and (iii) that components do not vanish during a split. We now translate these checks into logical expressions.

Presence of feed and products constraint: The process feed $([1, n])$ and the final products (pure components) must be present in every feasible configuration.

$$\left. \begin{array}{l} \zeta_{1,n} = 1 \\ \zeta_{1,1} = \zeta_{2,2} = \dots = \zeta_{n,n} = 1 \end{array} \right\} \quad (\text{H2})$$

In addition, specific submixtures are forced to be present (resp. absent) by requiring that the corresponding $\zeta_{i,j} = 1$ (resp. $\zeta_{i,j} = 0$).

Presence of parent constraint: A submixture must be absent when all of its top $(\{[i, l]\}_{l=j+1}^n)$ and bottom $(\{[k, j]\}_{k=1}^{i-1})$ parents are absent. Mathematically,

$$\zeta_{i,j} \leq \sum_{l=j+1}^n \zeta_{i,l} + \sum_{k=1}^{i-1} \zeta_{k,j} \quad \forall \quad [i, j] \in \mathcal{S}_2 \quad (\text{H3})$$

Note that this constraint is not imposed for submixtures of the form $[1, j]$, which belong to $\mathcal{S}_1$, and $[i, n]$, which belong to $\mathcal{S}_3$, since it would be redundant given that process feed $[1, n]$ is always present.

Constraint for feasibility of split: Components must be conserved during a split i.e., every component must be recovered either in distillate or in residue or both. For example, feasible splits of ABC are $\{A/BC, AB/C, AB/BC\}$. However, split A/C is infeasible since component B is absent from both distillate and residue. To impose this constraint, SA method identifies the distillate and residue from each split, and verifies that the sum of number of components in the products is at least as high as the number of components in the feed. To impose the constraint in this form, additional variables would be needed to identify unique components in the product streams. Instead, we formulate this constraint in an alternate way.

Consider the split of submixture $[i, j]$. The top and bottom children of $[i, j]$ are $\{[i, r]\}_{r=i}^{j-1}$ and $\{[s, j]\}_{s=i+1}^j$. Let $[i, r]$, where $i \leq r \leq j-1$, be produced as the distillate from $[i, j]$, i.e., $\zeta_{i,j}(1 - \zeta_{i,j-1}) \dots (1 - \zeta_{i,r+1})\zeta_{i,r} = 1$. Here, we use the fact that, for regular-column configurations [5, 9, 10], submixtures $\{[i, j-1], \dots, [i, r+1]\}$ are absent if $[i, r]$ is produced as the distillate from $[i, j]$. In this case, for conservation of components, the bottom product must be one of $\{[i+1, j], [i+2, j], \dots, [r+1, j]\}$. This constraint is imposed as:

$$\zeta_{i,j}(1 - \zeta_{i,j-1}) \dots (1 - \zeta_{i,r+1})\zeta_{i,r} \leq \sum_{s=i+1}^{r+1} \zeta_{s,j} \quad (3.1)$$

The LHS of (3.1) involves a product of binary variables, which can be relaxed to the LHS of (H4) below:

$$\zeta_{i,j} - \zeta_{i,j-1} - \dots - \zeta_{i,r+1} + \zeta_{i,r} - 1 \leq \sum_{s=i+1}^{r+1} \zeta_{s,j} \quad \forall \quad i \leq r \leq j-2, \quad [i, j] \in \mathcal{P} \quad (H4)$$

Equivalently to (3.1), (H4) is activated only if $[i, r]$ is produced as the distillate from $[i, j]$ and is trivially satisfied otherwise. Further, (H4) would be redundant if imposed for $r = j-1$, since the RHS would be $\sum_{s=i+1}^j \zeta_{s,j} \geq \zeta_{j,j} = 1$, where the equality is from (H2). Physically, when $[i, j-1]$ is produced as the distillate from $[i, j]$, any of the bottom children $\{[s, j]\}_{s=i+1}^j$ can be withdrawn as residue. Thus, it is sufficient to impose (H4) for top children $\{[i, r]\}_{r=i}^{j-2}$. In a similar manner, component conservation constraints can be derived based on the bottom children. Let $[s, j]$, where $i+1 \leq s \leq j$, be produced as the residue from $[i, j]$, then one of $\{[i, r]\}_{r=s-1}^{j-1}$ must be produced as the distillate for components to be conserved. For a regular-column configuration, $[s, j]$ is a residue only if the submixtures $\{[i+1, j], \dots, [s-1, j]\}$ are absent. The constraint in nonlinear form, and its relaxed version are given below:

$$\begin{aligned} \zeta_{i,j}(1 - \zeta_{i+1,j}) \dots (1 - \zeta_{s-1,j})\zeta_{s,j} &\leq \sum_{r=s-1}^{j-1} \zeta_{i,r} \\ \zeta_{i,j} - \zeta_{i+1,j} - \dots - \zeta_{s-1,j} + \zeta_{s,j} - 1 &\leq \sum_{r=s-1}^{j-1} \zeta_{i,r} \quad \forall \quad i+2 \leq s \leq j; \quad [i, j] \in \mathcal{P} \end{aligned} \quad (H5)$$

As before, (H5) is redundant for $s = i+1$ and it suffices to impose (H5) for bottom children $\{[s, j]\}_{s=i+2}^j$.

Constraints on heat exchanger variables

We use binary variables $\chi_{i,j}$, $(i,j) \in \mathcal{C}$ and $\rho_{i,j}$, $(i,j) \in \mathcal{R}$ to model whether condensers and reboilers are present or absent. When a heat exchanger is absent, the material flow between two consecutive columns takes place via two-way vapor-liquid transfers (see Figure 3.2). Thus, when a heat exchanger is absent, we automatically assume that a thermal coupling is present instead. We now describe the constraints that govern the presence/absence of heat exchangers:

1. From the definition, heat exchangers must be absent when the associated submixtures are absent in the configuration. This constraint can be expressed in linear form as

$$\chi_{i,j} + \rho_{i,j} \leq \zeta_{i,j} \quad \forall [i,j] \in \mathcal{U} \quad (\text{H6})$$

Submixtures in $\mathcal{S}_1$ (resp. $\mathcal{S}_3$) are only associated with condensers (resp. reboilers). Thus, (H6) must be modified to $\chi_{1,j} \leq \zeta_{1,j}$ (resp. $\rho_{i,n} \leq \zeta_{i,n}$) for the corresponding submixtures.

2. Condenser (resp. reboiler) associated with a stream must be absent if all the top (resp. bottom) parents of the stream are absent. This constraint is modeled as

$$\left. \begin{aligned} \chi_{i,j} &\leq \sum_{l=j+1}^n \zeta_{i,l} \\ \rho_{i,j} &\leq \sum_{k=1}^{i-1} \zeta_{k,j} \end{aligned} \right\} \quad \forall [i,j] \in \mathcal{S}_2 \cup \{[m,m]_{m=2}^{n-1}\} \quad (\text{H7})$$

and is redundant if imposed for submixtures in $\mathcal{S}_1$ and $\mathcal{S}_3$, as they can always be produced from the process feed stream $([1,n])$.

3. Heat exchangers must be absent if the associated stream is withdrawn as a sidedraw. For a stream to be withdrawn as sidedraw, it must have at least one

of the top parents (i.e., $\min\{(1 - \zeta_{i,l})\}_{l=j+1}^n = 0$), and one of the bottom parents (i.e., $\min\{(1 - \zeta_{k,j})\}_{k=1}^{i-1} = 0$). Therefore, for every $[i, j] \in \mathcal{S}_2 \cup \{[m, m]\}_{m=2}^{n-1}$

$$\chi_{i,j} + \rho_{i,j} \leq \min\{(1 - \zeta_{i,l})\}_{l=j+1}^n + \min\{(1 - \zeta_{k,j})\}_{k=1}^{i-1} \quad (3.2)$$

By definition, $\min\{(1 - \zeta_{i,l})\}_{l=j+1}^n \leq (1 - \zeta_{i,l})$ for $j+1 \leq l \leq n$, and $\min\{(1 - \zeta_{k,j})\}_{k=1}^{i-1} \leq (1 - \zeta_{k,j})$ for $1 \leq k \leq i-1$. This allows min function to be linearized as follows:

$$\chi_{i,j} + \rho_{i,j} \leq (1 - \zeta_{i,l}) + (1 - \zeta_{k,j}) \quad \forall \quad 1 \leq k \leq i-1, j+1 \leq l \leq n, \left. \begin{array}{l} [i, j] \in \mathcal{S}_2 \cup \{[m, m]\}_{m=2}^{n-1} \end{array} \right\} \quad (H8)$$

The linearized constraints check every top parent-bottom parent combination to determine if the submixture is withdrawn as sidedraw. This constraint is, of course, not applicable for the remaining streams, because they can not be produced as sidedraw.

4. Pure components must be withdrawn through a heat exchanger, when they are produced by only one parent. Since the lightest and the heaviest components always have one parent, we have

$$\chi_{1,1} = 1 \quad \& \quad \rho_{n,n} = 1 \quad (H9)$$

For the remaining products ($\{[i, i]\}_{i=2}^{n-1}$), if they only have a top parent, they must be withdrawn through a condenser. In other words, if all the bottom parents ($\{[k, j]\}_{k=1}^{j-1}$) are absent, the condenser associated with the intermediate product must be present.

$$\chi_{j,j} \geq 1 - \sum_{k=1}^{j-1} \zeta_{k,j} \quad \forall \quad 2 \leq j \leq n-1 \quad (H10)$$

Similarly, when all the top parents ($\{[j, l]\}_{l=j+1}^n$) are absent, the intermediate product must be withdrawn from a reboiler.

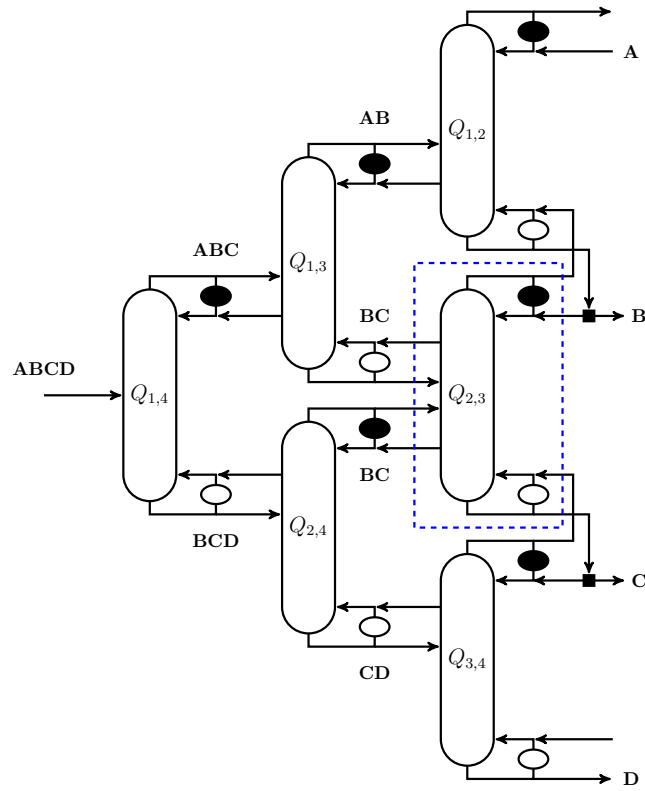
$$\rho_{j,j} \geq 1 - \sum_{l=j+1}^n \zeta_{j,l} \quad \forall \quad 2 \leq j \leq n-1 \quad (H11)$$

This concludes our discussion of the search-space formulation. The above constraints are sufficient to eliminate all infeasible regular-column configurations. The formulation does not require all variables to be declared as binary. As it is, the MINLP has $n(n+1)/2$ binary variables associated with streams $(\zeta_{i,j})$ and $n(n-1)$ binary variables for heat exchangers $(\chi_{i,j}, \rho_{i,j})$. However, due to (H2), the binary variables associated with the process feed and the pure components are already fixed at 1. Further, heat exchanger variables associated with pure products $\{[i, i]\}_{i=1}^n$ automatically take integral values. This is because if the product i does not have a top parent, (H7) (resp. (H11)) ensures that $\chi_{i,i}$ (resp. $\rho_{i,i}$) is zero (resp. one). The case without bottom parent is similar. On the other hand, when it has both parents, (H8) ensures that $\chi_{i,i}$ and $\rho_{i,i}$ are both zero. The total number of binary variables is thus $n(n+1)/2 + n(n-1) - (n+1) - 2(n-1) = (n-2)(3n-1)/2$. For example, the number of binary variables for four, five, and six component systems are 11, 21, and 34, respectively.

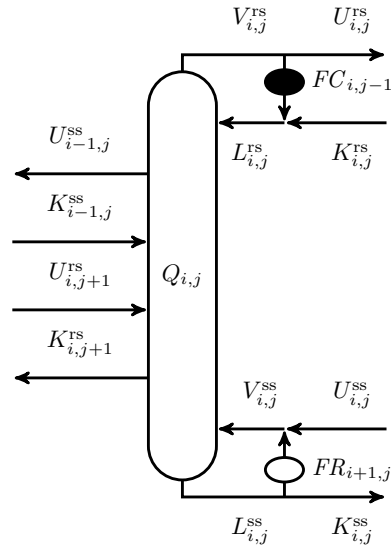
3.1.3 Mass Balance Constraints

Figure 3.2(a) shows the four-component superstructure used for modeling mass balance constraints. We observe that, by bypassing appropriate component, vapor, and liquid flows [17], the entire search space is embedded in Figure 3.2(a). We denote the *pseudocolumn* performing the split of stream $[i, j]$ as $Q_{i,j}$. We treat the system as a multicommodity network flow problem where a pseudocolumn, which redistributes the flow between the rectifying and stripping sections, is akin to a node. Whereas, the sections carrying the material from one node to the other are akin to edges.

Figure 3.2(b) shows a pseudocolumn that is a representative of every other pseudocolumn in the superstructure. Definition of variables used in Figure 3.2(b) are



(a) Four-component superstructure



(b) Representative column

Fig. 3.2. Schematic of four-component superstructure. The blue box shows a representative column of every other column in the superstructure. Variables used for vapor and liquid flows are shown in (b). Definition of the variables are described in Table 3.2

summarized in Table 3.2. The component and vapor mass balance equations enveloping a pseudocolumn $Q_{i,j}$ are given by:

$$\left. \begin{aligned} \{f_{i,j+1,p}^{\text{rs}} + f_{i-1,j,p}^{\text{ss}} = f_{i,j,p}^{\text{rs}} + f_{i,j,p}^{\text{ss}}\}_{p=i}^j \\ U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}} = V_{i,j}^{\text{rs}} - V_{i,j}^{\text{ss}} \end{aligned} \right\} \quad \forall [i,j] \in \mathcal{U} \quad (\text{H12})$$

For $Q_{1,n}$, the following constraints are imposed:

$$\left. \begin{aligned} \{F_p = f_{1,n,p}^{\text{rs}} + f_{1,n,p}^{\text{ss}}\}_{p=1}^n \\ (1 - \Phi) \sum_{p=1}^n F_p = V_{1,n}^{\text{rs}} - V_{1,n}^{\text{ss}} \end{aligned} \right\} \quad (\text{H13})$$

Here, Φ is the fraction of the process feed in liquid phase, typically referred to as thermal quality of the feed. For each column, as is typical, we assume that the lightest and the heaviest components of the feed appear only in the distillate and residue respectively. The intermediate components may, however, distribute between the distillate and the residue. In other words, the components in the distillate belong to $[i, j-1]$ (i.e., $f_{i,j}^{\text{rs}} = 0$), while those in residue belong to $[i+1, j]$ (i.e., $f_{i,j,i}^{\text{ss}} = 0$). Note that, for submixtures in $\mathcal{S}_1$ (resp. $\mathcal{S}_3$), $U_{i-1,j}^{\text{ss}}$ and $f_{i-1,j,p}^{\text{ss}}$ (resp. $U_{i,j+1}^{\text{rs}}$ and $f_{i,j+1,p}^{\text{rs}}$) are not defined and are to be regarded as zero in (H12). The difference between the vapor and liquid flows in the rectifying (resp. stripping) sections is the net distillate (resp. residue) flow.

$$\left. \begin{aligned} V_{i,j}^{\text{rs}} - L_{i,j}^{\text{rs}} &= \sum_{p=i}^{j-1} f_{i,j,p}^{\text{rs}} \\ L_{i,j}^{\text{ss}} - V_{i,j}^{\text{ss}} &= \sum_{p=i+1}^j f_{i,j,p}^{\text{ss}} \end{aligned} \right\} \quad \forall [i,j] \in \mathcal{P} \quad (\text{H14})$$

Mass balance equations around the condenser (H15) and reboiler (H16) are given by

$$\left. \begin{aligned} V_{i,j+1}^{\text{rs}} &= FC_{i,j} + U_{i,j+1}^{\text{rs}} \\ FC_{i,j} + K_{i,j+1}^{\text{rs}} &= L_{i,j+1}^{\text{rs}} \end{aligned} \right\} \quad \forall (i,j) \in \mathcal{C} \quad (\text{H15})$$

$$\left. \begin{aligned} L_{i-1,j}^{\text{ss}} &= FR_{i,j} + K_{i-1,j}^{\text{ss}} \\ FR_{i,j} + U_{i-1,j}^{\text{ss}} &= V_{i-1,j}^{\text{ss}} \end{aligned} \right\} \quad \forall (i,j) \in \mathcal{R} \quad (\text{H16})$$

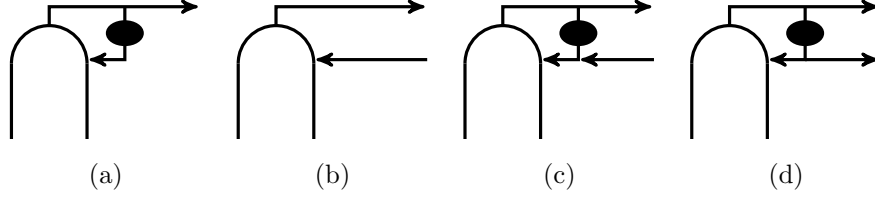


Fig. 3.3. Modeling of (a) Condenser, (b) Thermal coupling, (c) Combination of condenser and thermal coupling and (d) a partial condenser

see Figure 3.2(b). We point out that the liquid mass balance equation around each pseudocolumn is implied from (H12) – (H16). To facilitate the production of pure components in saturated liquid state, we add (see Figure 3.2a)

$$\left. \begin{aligned} U_{i,i+1}^{\text{rs}} &= U_{i-1,i}^{\text{ss}} \\ K_{i-1,i}^{\text{ss}} - K_{i,i+1}^{\text{rs}} &= F_p \end{aligned} \right\} \quad 2 \leq i \leq n-1 \quad (\text{H17})$$

Here, we used the fact that the net outflow of any intermediate pure product is the same as its flow in the process feed F_p . These constraints ensure vapor and liquid continuity among pseudocolumns $\{Q_{i,i+1}\}_{i=1}^{n-1}$. The superstructure in Figure 3.2 can model a heat exchanger (Figure 3.3a), a thermal coupling (Figure 3.3b), a combination of thermal coupling and heat exchanger (Figure 3.3c), which is a plausible alternative for modeling pump-around, and a partial condenser/reboiler (Figure 3.3d) by enabling/shunting flow in appropriate branches. In this work, we allow only two of these options. More specifically, we allow either a heat exchanger or a thermal coupling, but not both (Figure 3.3c). In addition, we will only consider total condenser and reboiler i.e., two-phase transfer streams are not allowed (Figure 3.3d). When a condenser (resp. reboiler) is present, we feed the submixture as vapor (resp. liquid) to the next column [35]. These constraints are modeled as follows:

$$\left. \begin{aligned} FC_{i,j} &\leq V^U \chi_{i,j} \\ K_{i,j+1}^{\text{rs}} &\leq V^U (1 - \chi_{i,j}) \end{aligned} \right\} \quad \forall \quad (i,j) \in \mathcal{C} \quad (\text{H18})$$

$$\left. \begin{aligned} FR_{i,j} &\leq V^U \rho_{i,j} \\ U_{i-1,j}^{ss} &\leq V^U (1 - \rho_{i,j}) \end{aligned} \right\} \quad \forall \quad (i,j) \in \mathcal{R} \quad (\text{H19})$$

Here, V^U is the upper bound on the vapor flow. The choice of V^U will be discussed in section 3.1.6.

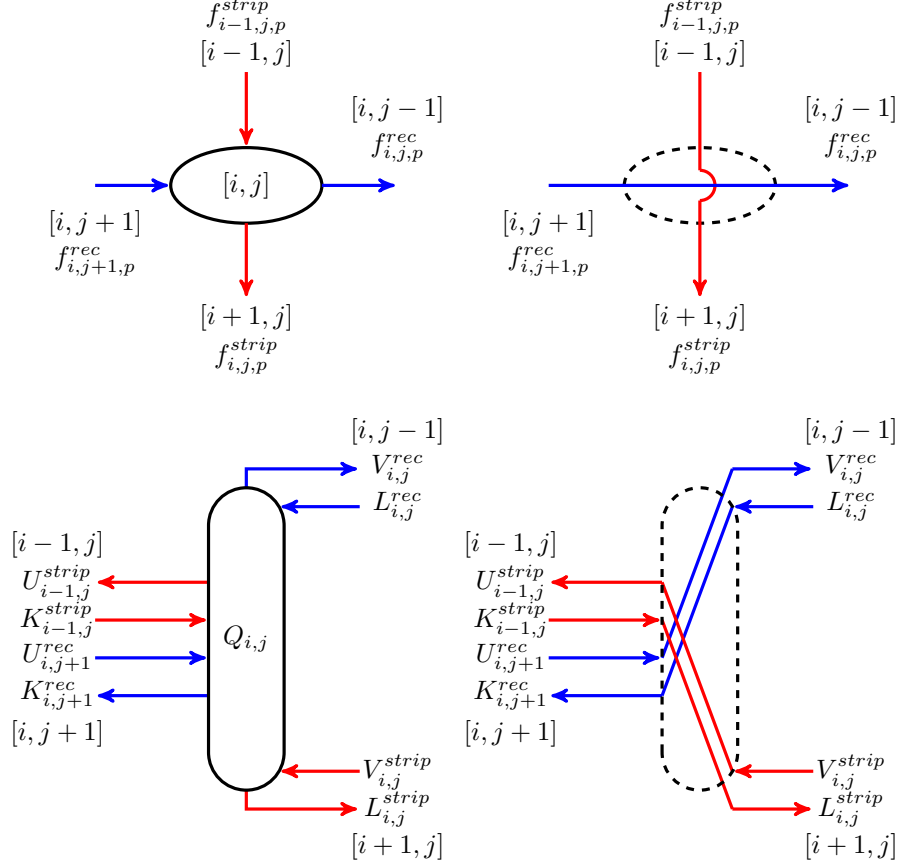


Fig. 3.4. Schematic showing component, vapor and liquid bypass

3.1.4 Underwood Constraints

Here, we describe the constraints that govern the minimum vapor flow required in each section of the configuration to perform the desired separation. These constraints also determine the extent to which component flows are distributed between the distillate and the residue, when a submixture undergoes sloppy split. All the sloppy split

configurations have at least one instance where two pseudocolumns stacked, one on top of the other, with the common product stream drawn as sidedraw. As a first-step approximation, despite the stacking, we model the split in each pseudocolumn independently of other pseudocolumns. We compute the minimum vapor requirement for each pseudocolumn using Underwood method. The minimum vapor flow requirement of a stacked column is then obtained as the maximum of vapor flow requirements at each of the sections. This procedure has been used in the literature by [9, 13, 15, 22]. Very recently, it was found that the results obtained with this approximation agree excellently with rigorous Aspen Plus simulations [4].

When stream $[i, j]$ is present, pseudocolumn $Q_{i,j}$ separates $[i, j]$ so that the component distribution and the threshold vapor flow satisfy Underwood constraints. On the other hand, if stream $[i, j]$ is absent in a configuration, pseudocolumn $Q_{i,j}$ bypasses the flow from the rectifying (resp. stripping) section of the parent pseudocolumn $Q_{i,j+1}$ (resp. $Q_{i-1,j}$) to the rectifying (resp. stripping) section of the pseudocolumn $Q_{i,j}$. A schematic of the bypass is shown in Figure 3.4. The relevant constraints are shown in the disjunction below, which is imposed for all $[i, j] \in \mathcal{P}$

$$\begin{aligned}
& \left[\begin{aligned}
& \zeta_{i,j} = 1 \\
& \left\{ \sum_{p=i}^j \frac{\alpha_p (f_{i,j+1,p}^{\text{rs}} + f_{i-1,j,p}^{\text{ss}})}{\alpha_p - \theta_{i,j,q}} = U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}} \right\}_{q=i}^{j-1} \\
& \sum_{p=i}^{j-1} \frac{\alpha_p f_{i,j,p}^{\text{rs}}}{\alpha_p - \theta_{i,j,q}} \leq \Upsilon_{i,j}^{\text{rs}} \text{ for } q = \{i, j-1\} \\
& \left\{ \sum_{p=i}^{j-1} \frac{\alpha_p f_{i,j,p}^{\text{rs}}}{\alpha_p - \theta_{i,j,q}} + S_{i,j,q} = \Upsilon_{i,j}^{\text{rs}} \right\}_{q=i+1}^{j-2} \\
& - \sum_{p=i+1}^j \frac{\alpha_p f_{i,j,p}^{\text{ss}}}{\alpha_p - \theta_{i,j,q}} \leq \Upsilon_{i,j}^{\text{ss}} \text{ for } q = \{i, j-1\} \\
& \left\{ - \sum_{p=i+1}^j \frac{\alpha_p f_{i,j,p}^{\text{ss}}}{\alpha_p - \theta_{i,j,q}} + S_{i,j,q} = \Upsilon_{i,j}^{\text{ss}} \right\}_{q=i+1}^{j-2} \\
& S_{i,j,q} \leq 2V^U \left[\min\{(1 - \zeta_{i,r})\}_{r=q+1}^{j-1} + \min\{(1 - \zeta_{s,j})\}_{s=i+1}^q \right] \\
& \Upsilon_{i,j}^{\text{rs}} \leq V_{i,j}^{\text{rs}} \\
& \Upsilon_{i,j}^{\text{ss}} \leq V_{i,j}^{\text{ss}} \\
& U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}} = \Upsilon_{i,j}^{\text{rs}} - \Upsilon_{i,j}^{\text{ss}}
\end{aligned} \right] \bigvee \left[\begin{aligned}
& \zeta_{i,j} = 0 \\
& \{f_{i,j+1,p}^{\text{rs}} = f_{i,j,p}^{\text{rs}}\}_{p=i}^{j-1} \\
& \{f_{i-1,j,p}^{\text{ss}} = f_{i,j,p}^{\text{ss}}\}_{p=i+1}^j \\
& U_{i,j+1}^{\text{rs}} = V_{i,j}^{\text{rs}} \\
& U_{i-1,j}^{\text{ss}} = V_{i,j}^{\text{ss}}
\end{aligned} \right] \tag{3.3}
\end{aligned}$$

Consider the set corresponding to $\zeta_{i,j} = 1$. The second constraint is the Underwood feed equation that ensures that the Underwood roots satisfy $\alpha_{q+1} \leq \theta_{i,j,q} \leq \alpha_q$ for $i \leq q \leq j-1$. The next four constraints determine the minimum vapor requirement of the rectifying and stripping sections $\Upsilon_{i,j}^{\text{rs}}$ and $\Upsilon_{i,j}^{\text{ss}}$, respectively. The fourth constraint, which applies to $q \in \{i+1, \dots, j-2\}$, is converted into an equality constraint by introducing a slack variable $S_{i,j,q}$. Since the nonlinear term on the LHS of this constraint is bounded from below by $-V^U$ [25] and the RHS is bounded from above by V^U , we can bound this variable from above by $2V^U$, as in the seventh constraint. Moreover, if the product streams of $Q_{i,j}$, $[i, r]$ and $[s, j]$, are such that $i < s < r < j$ i.e., at least two components distribute between the distillate and residue, then the Underwood roots satisfying $s \leq q < r$ are active (see Appendix A). The seventh constraint ensures that the slack in the fourth constraint is zero, as desired. The

min function is linearized just as it was for (H8). Similarly, the sixth constraint is converted into an equality by adding a slack variable. From the feed equation and the minimum vapor constraint in the rectifying section, it can be readily shown that the slack variable for the constraint in stripping section is also $S_{i,j,q}$. Although the constraints in stripping section are implied from that for the rectifying section and the feed equation, we have retained so they help strengthen the linear relaxation. These constraints, thus ensure that the solutions that are physically infeasible (see Chapter 1) are not admissible to our formulation. The eighth and ninth constraints ensure that the actual vapor flow is at least as high as the threshold vapor flow. It holds at equality when the column is operated at pinch, while a strict inequality holds when the column is operated above pinch. An inherent assumption for sloppy-splits is that, it is always possible to achieve the desired distribution by adjusting the trays in the column, when it is operated above pinch. Finally, the last constraint ensures vapor balance, requiring that the minimum vapor flow in the rectifying section differs from that in stripping section by the net vapor inflow.

On the other hand, the constraints corresponding to $\zeta_{i,j} = 0$ model component and vapor bypass. The liquid bypass constraints can be derived using linear arguments from vapor and component bypass constraints using (H14), (H15), and (H16). Thus, they are not included in the disjunction. Besides Underwood constraints, the disjunction in (3.3) differs from that in [9] in the following way. Overall mass balance equations across each pseudocolumn are part of the disjunction in [9]. These constraints hold regardless of the presence/absence of the pseudocolumn. However, the authors impose these constraints in BigM form, which relaxes them in the absence of pseudocolumn. On the contrary, we do not include the mass balance constraints in the disjunction and, therefore, are not relaxed in the absence of pseudocolumn. Computational evidence suggests that the current approach leads to linear relaxations that are significantly tighter.

As mentioned in the Introduction, for tightness, we rely on reformulation based techniques for the treatment of disjunctions. Inspired from disjunctive programming

techniques in [36], we transform (3.3), for $[i, j] \in \mathcal{P}$, into the following set of equalities and inequalities.

$$\left\{ \sum_{p=i}^j \frac{\alpha_p (f_{i,j+1,p}^{\text{rsI}} + f_{i-1,j,p}^{\text{ssI}})}{\alpha_p - \theta_{i,j,q}} = U_{i,j+1}^{\text{rsz}} - U_{i-1,j}^{\text{ssz}} \right\}_{q=i}^{j-1} \quad (\text{H20})$$

$$\left\{ \sum_{p=i}^{j-1} \frac{\alpha_p f_{i,j,p}^{\text{rsO}}}{\alpha_p - \theta_{i,j,q}} \leq \Upsilon_{i,j}^{\text{rsz}} \text{ for } q = \{i, j-1\} \right. \\ \left. \left\{ \sum_{p=i}^{j-1} \frac{\alpha_p f_{i,j,p}^{\text{rsO}}}{\alpha_p - \theta_{i,j,q}} + S_{i,j,q}^\zeta = \Upsilon_{i,j}^{\text{rsz}} \right\}_{q=i+1}^{j-2} \right\} \quad (\text{H21})$$

$$\left\{ - \sum_{p=i+1}^j \frac{\alpha_p f_{i,j,p}^{\text{ssO}}}{\alpha_p - \theta_{i,j,q}} \leq \Upsilon_{i,j}^{\text{ssz}} \text{ for } q = \{i, j-1\} \right. \\ \left. \left\{ - \sum_{p=i+1}^j \frac{\alpha_p f_{i,j,p}^{\text{ssO}}}{\alpha_p - \theta_{i,j,q}} + S_{i,j,q}^\zeta = \Upsilon_{i,j}^{\text{ssz}} \right\}_{q=i+1}^{j-2} \right\} \quad (\text{H22})$$

$$S_{i,j,q}^\zeta \leq 2V^U \left[\min\{(1 - \zeta_{i,r})\}_{r=q+1}^{j-1} + \min\{(1 - \zeta_{s,j})\}_{s=i+1}^q \right] \quad (\text{H23})$$

$$\left\{ \begin{aligned} \Upsilon_{i,j}^{\text{rsz}} &\leq V_{i,j}^{\text{rsz}} \\ \Upsilon_{i,j}^{\text{ssz}} &\leq V_{i,j}^{\text{ssz}} \end{aligned} \right\} \quad (\text{H24})$$

$$U_{i,j+1}^{\text{rsz}} - U_{i-1,j}^{\text{ssz}} = \Upsilon_{i,j}^{\text{rsz}} - \Upsilon_{i,j}^{\text{ssz}} \quad (\text{H25})$$

$$\left\{ \begin{aligned} \{f_{i,j+1,p}^{\text{rs}} - f_{i,j+1,p}^{\text{rsI}} - f_{i,j,p}^{\text{rs}} + f_{i,j,p}^{\text{rsO}} = 0\}_{p=i}^{j-1} \\ \{f_{i-1,j,p}^{\text{ss}} - f_{i-1,j,p}^{\text{ssI}} - f_{i,j,p}^{\text{ss}} + f_{i,j,p}^{\text{ssO}} = 0\}_{p=i+1}^j \\ U_{i,j+1}^{\text{rs}} - U_{i,j+1}^{\text{rsz}} - V_{i,j}^{\text{rs}} + V_{i,j}^{\text{rsz}} = 0 \\ U_{i-1,j}^{\text{ss}} - U_{i-1,j}^{\text{ssz}} - V_{i,j}^{\text{ss}} + V_{i,j}^{\text{ssz}} = 0 \end{aligned} \right\} \quad (\text{H26})$$

The above constraints are derived by multiplying the constraints corresponding to $\zeta_{i,j} = 1$ (resp. $\zeta_{i,j} = 0$) by $\zeta_{i,j}$ (resp. $(1 - \zeta_{i,j})$), expanding the product and linearizing the bilinear terms by introducing auxiliary variables $a_{ijp} = b_{ijp} \cdot \zeta_{i,j}$, where $a_{ijp} = [f_{i,j+1,p}^{\text{rsI}}, f_{i-1,j,p}^{\text{ssI}}, f_{i,j,p}^{\text{rsO}}, f_{i,j,p}^{\text{ssO}}, U_{i,j+1}^{\text{rsz}}, U_{i-1,j}^{\text{ssz}}, V_{i,j}^{\text{rsz}}, V_{i,j}^{\text{ssz}}]$ and $b_{ijp} = [f_{i,j+1,p}^{\text{rs}}, f_{i-1,j,p}^{\text{ss}}, f_{i,j,p}^{\text{rs}}, f_{i,j,p}^{\text{ss}}, U_{i,j+1}^{\text{rs}}, U_{i-1,j}^{\text{ss}}, V_{i,j}^{\text{rs}}, V_{i,j}^{\text{ss}}]$.

$f_{i,j,p}^{\text{ss}}, U_{i,j+1}^{\text{rs}}, U_{i-1,j}^{\text{ss}}, V_{i,j}^{\text{rs}}, V_{i,j}^{\text{ss}}$. The bilinear equality constraints are relaxed using McCormick constraints [37].

$$\left. \begin{aligned} a_{ijp} &\leq \min\{b_{ijp}, \zeta_{i,j} b_{ijp}^U\} \\ a_{ijp} &\geq \max\{0, b_{ijp} + b_{ijp}^U(\zeta_{i,j} - 1)\} \end{aligned} \right\} p = \{i, \dots, j\} \quad \forall \quad [i, j] \in \mathcal{P} \quad (\text{H27})$$

Because $\zeta_{i,j}$ is binary, it is always at its bound: either 0 or 1 and, as such, McCormick constraints exactly model the bilinear product. As before, undefined terms are dropped for submixtures in $\mathcal{S}_1$ and $\mathcal{S}_3$. This concludes the discussion on minimum vapor constraints, completing the formulation of the problem. In the next subsection, we exploit monotonicity of Underwood equations to derive additional valid cuts that strengthen the relaxation.

3.1.5 Flow of Underwood Roots

Carlberg and Westerberg [38] observed that if two columns are thermally coupled, the Underwood roots carry from one to another. This would, for example, be the case for the pseudocolumns $Q_{1,4}$ and $Q_{1,3}$ in Figure 3.2 if their interconnection is a thermal coupling. The flow of roots is because the minimum vapor constraint for the rectifying section of $Q_{1,4}$ becomes the feed equation for the next column $Q_{1,3}$. Mathematically, (H21), (H24), (H15), and (H20), in the specified order, lead to:

$$\sum_{p=1}^3 \frac{\alpha_p f_{1,4,p}^{\text{rs}}}{\alpha_p - \theta_{1,4,q}} \leq \Upsilon_{1,4}^{\text{rs}} \leq V_{1,4}^{\text{rs}} = U_{1,4}^{\text{rs}} = \sum_{p=1}^3 \frac{\alpha_p f_{1,4,p}^{\text{rs}}}{\alpha_p - \theta_{1,3,q}}; \quad q = \{1, 2\} \quad (3.4)$$

In deriving (3.4), we used $FC_{1,3} = 0$, as the columns are thermally coupled. Since $\Upsilon_{1,4}^{\text{rs}}$ dominates the nonlinear expression in (H21) for both active and inactive roots (see Appendix A), (3.4) holds for all the Underwood roots. Moreover, by recognizing that the nonlinear expressions in (3.4) are monotonically increasing functions of Underwood roots, [38] showed that $\theta_{1,4,q} \leq \theta_{1,3,q}$ for $q = \{1, 2\}$. Using similar arguments, it can be shown that $\theta_{1,4,q} \geq \theta_{2,4,q}$ for $q = \{2, 3\}$, when the pseudocolumns $Q_{1,4}$ and $Q_{2,4}$ are thermally coupled. This concept was used in the derivation of an analytical solution for the vapor duty of an n component FTC by [34].

Although these constraints are implicit and redundant in the presence of Underwood equations, imposing the constraints explicitly helps improve the LP/MIP relaxation, where Underwood equations only appear in relaxed form. We impose these constraints as:

$$\begin{aligned}
& \{\theta_{i,j+1,q} - \theta_{i,j,q} \leq M_q \chi_{i,j}\}_{q=i}^{j-1} \quad \forall \quad [i,j] \in \mathcal{S}_1 \\
& \left\{ \left\{ \theta_{i,l,q} - \theta_{i,j,q} \leq M_q \left[\chi_{i,j} + (1 - \zeta_{i,j}) + \sum_{k=1}^{i-1} \zeta_{k,j} + \sum_{m=j+1}^{l-1} \zeta_{i,m} \right] \right\}_{l=j+1}^n \right\}_{q=i}^{j-1} \quad \forall \quad [i,j] \in \mathcal{S}_2 \\
& \left\{ \left\{ \theta_{i,j,q} - \theta_{k,j,q} \leq M_q \left[\rho_{i,j} + (1 - \zeta_{i,j}) + \sum_{l=j+1}^n \zeta_{i,l} + \sum_{m=k+1}^{i-1} \zeta_{m,j} \right] \right\}_{k=1}^{i-1} \right\}_{q=i}^{j-1} \quad \forall \quad [i,j] \in \mathcal{S}_2 \\
& \{\theta_{i,j,q} - \theta_{i-1,j,q} \leq M_q \rho_{i,j}\}_{q=i}^{j-1} \quad \forall \quad [i,j] \in \mathcal{S}_3
\end{aligned} \tag{H28}$$

where the upper bound on the difference between the roots $M_q = \alpha_q - \alpha_{q+1}$. For $[i,j] \in \mathcal{S}_1$ (resp. $[i,j] \in \mathcal{S}_3$), the constraints flow the roots from the parent pseudocolumn $Q_{i,j+1}$ (resp. $Q_{i-1,j}$) to $Q_{i,j}$, when the associated condenser (resp. reboiler) is absent. Furthermore, when the submixture is absent, implies $\chi_{i,j} = 0$ (resp. $\rho_{i,j} = 0$) from (H6), these constraints bypass the roots. Whereas, for $[i,j] \in \mathcal{S}_2$, the constraints are modified for two reasons. First, the roots may not flow when the submixture is withdrawn as sidedraw and its parent pseudocolumns are operated above pinch. (H28) ensures that the flow constraints are deactivated if $[i,j] \in \mathcal{S}_2$ has both top and bottom parents. The second reason is better explained with an example. Consider the *completely thermally coupled configuration* shown in Figure 3.5. Here, $[2,3]$ (resp. $[3,4]$) is produced by a single parent $[2,5]$ (resp. $[1,4]$). Since their interconnection is a thermal coupling, the roots are allowed to flow to $Q_{2,3}$ (resp. $Q_{3,4}$) from $Q_{2,5}$ (resp. $Q_{1,4}$). However, without introducing additional variables, the flow cannot take place via $Q_{2,4}$, since the roots received from its parents ($Q_{2,5}$ and $Q_{1,4}$) can be different. (H28) resolves this issues by directly flowing the root from the parent pseudocolumn. The second constraint in (H28) flows the root to $Q_{i,j}$ from its parent $Q_{i,l}$ for $j+1 \leq l \leq n$, when $[i,j]$ is present and the associated condenser, all of its bottom parents $\{[k,j]\}_{k=1}^{i-1}$ and the submixtures $\{[i,m]\}_{m=j+1}^{l-1}$ are absent. This constraint enables

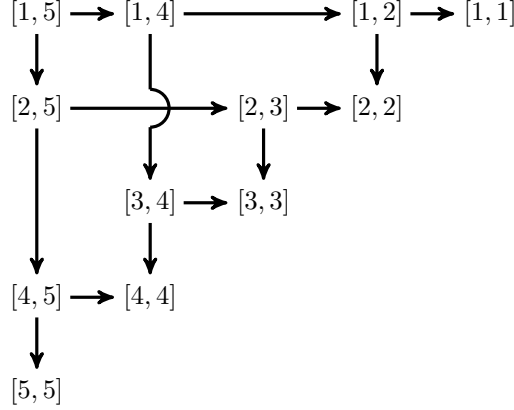


Fig. 3.5. A five-component completely thermally coupled configuration

the root to flow from $Q_{2,5}$ to $Q_{2,3}$ in the example under consideration. The third constraint is interpreted in a similar way, and it enables the root to flow from $Q_{1,4}$ to $Q_{3,4}$. To summarize, provided a stream is not drawn as a sidedraw, (H28) flows the roots to $Q_{i,j}$ from its parent if their interconnection is a thermal coupling.

We remark that the Underwood roots of the pseudocolumn $Q_{1,n}$ can be calculated *a priori*, thus, fixing $\{\theta_{1,n,q}\}_{q=1}^{n-1}$. Then (H28) propagates these roots to the subsequent pseudocolumns in the configuration when they are thermally coupled. Whenever the global solver decides to choose thermal coupling, it can use (H28) to improve bounds on Underwood roots using optimality and feasibility based range reduction techniques, thus improving the quality of relaxations and consequently converging to the optimal solution faster. We demonstrate the improvement in relaxation due to (H28) with a numerical example. We consider a mixture of alcohols (Case 9 in Table 5): $n = 5$, $\Phi = 1$, $F = \{20, 30, 20, 20, 10\}$ and $\alpha = \{4.1, 3.6, 2.1, 1.42, 1\}$. We fix the binary variables associated with streams and heat exchangers to Fully Thermally Coupled (FTC) configuration. For this case, the optimal objective function value is 402.703 kmol/hr. The root node relaxation reported by BARON without (H28) is 283.842 (42% gap w.r.t lower bound). Whereas, when (H28) is included in the model, BARON inferred a lower bound of 402.703 in the preprocessing step, thereby solving the problem to global optimum.

Inclusion of (H28) in optimization models is an important contribution of this Chapter. Before proceeding further, we comment on how the state-of-the-art explicit enumeration algorithm [22] fares in the absence of these cuts. As an example, we considered a five-component mixture of aromatics (see Case 10 in Table 3.5). Without (H28), the explicit enumeration based algorithm could not solve a few configurations to 1% gap within 1200 seconds. However, when (H28) is included in the model, BARON could solve all the unconverged configurations in less than five seconds! This example demonstrates the importance of these cuts, and how they improve the state-of-the-art. It will be even more evident when we discuss case studies in subsection 3.3.

3.1.6 Bounds on Variables

The bounds on binary variables are obtained by definition. Clearly, the net component flow in any column section cannot exceed the component flow rate in the feed. Thus, $0 \leq f_{i,j,p}^{rs}, f_{i,j,p}^{ss} \leq F_p$. The bounds on the Underwood roots are chosen to be $\alpha_{q+1} + 10^{-4} \leq \theta_{i,j,q} \leq \alpha_q - 10^{-4}$. A small portion is chopped at both the ends of the interval to prevent singularity issues. Although we show rigorous ways of avoiding this issue in [25], when using off-the-shelf solvers, the assumed strict separation is the simplest way to avoid numerical difficulties.

An arbitrarily large value can be chosen for the bounds on vapor and liquids flows. This, however, leads to weaker LP/MIP relaxations. To overcome this challenge we propose a heuristic. We have observed on a test set which spans over a wide range of feed cases [35, 39], that 3.5 times the FTC vapor duty is always higher than the total vapor duty of all configurations in the search space. Thus, 3.5 times the FTC vapor duty is chosen to be the upper bound on all vapor flow variables. Observe that using such a bound is reasonable, since we are only interested in configurations with low

vapor duty. To implement this bound, we calculate vapor duty of the FTC (VD_{FTC}) analytically (see [15, 34]) as follows:

$$VD_{FTC} = \max \left\{ \sum_{p=1}^q \frac{\alpha_p F_p}{\alpha_p - \theta_{1,n,q}} \right\}_{q=1}^{n-1} \quad (3.5)$$

The bounds on liquid flows are computed using mass balance equations, and bounds on vapor and component flows. All variables except $\{K_{i,i+1}\}_{i=1}^{n-1}$ are positive. The lower bound on these liquid flows is set to $-V^U$ to withdraw products in liquid state (see Figure 3.2a).

Altogether, the binary network, mass balances, Underwood constraints and bounds constitute the MINLP model (H), which can be solved using standard MINLP solvers such as BARON [21, 40] or SCIP [41].

3.2 Case Studies

In this section, we present a few case studies from literature that were solved to ϵ -global optimality. First, we study the five-component heavy crude example from [5]. Using this example, we illustrate how additional constraints can be added to customize the search space of configurations. Later, we will present four, five and six component instances from the literature.

3.2.1 Heavy Crude Distillation

Crude oil distillation is an important separation process that helps meet various energy and transportation needs of society. As mentioned in the introduction, the crude separation consumes energy equivalent of ~ 1.8 million bbl of crude oil per day. Clearly, a small reduction in the energy consumption for this process can save a lot of energy over time.

In our example, crude oil is regarded as a mixture of naphtha (A), kerosene (B), diesel (C), gas oil (D) and residue (E). The composition of components in the feed and the relative volatility of the components is borrowed from [5] and summarized in

Table 3.3.
Feed composition and relative volatility information for heavy crude mixture. Data is taken from [5]

	Component	Feed Composition	Relative Volatility
A	Naphtha	0.144	45.3
B	Kerosene	0.093	14.4
C	Diesel	0.101	4.7
D	Gas Oil	0.039	2
E	Residue	0.623	1

Table 3.3. We assume the flowrate of crude to be 100 kmol/h. Further, we assume that 90% of E in the process feed is in liquid phase, while the rest is in vapor phase. To prevent fouling, we require that the residue (E) be separated from the rest of the components in the first column. In other words, none of the submixtures will contain E i.e., $BCDE$, CDE and DE are absent from the configuration. To ensure that these submixtures are absent, we add

$$\zeta_{2,5} = \zeta_{3,5} = \zeta_{4,5} = 0 \quad (3.6)$$

Before solving the MINLP over the entire search space, we limit the search to sharp-split configurations.

Sharp-split configurations: Due to their simplicity and ease of operation, sharp-split configurations were preferred historically. This preference sustains and, even today many chemical plants employ sharp-split configurations. Sharp-split configurations are those regular-column configurations that have exactly $n - 2$ submixtures. This is because for sharp-split configurations a submixture is not produced at two different locations. Therefore, splits can be rearranged as binary trees that have $n - 1$ internal nodes, one of them corresponds to the feed and the rest to submixtures. When a sloppy-split is introduced, the trees following each product stream grow by the number of overlapping components, c , and at most $c - 1$ of these internal internal

nodes can be pruned, which happens when both product streams separate to produce a product downstream that consists of just the c overlapping components. Therefore, sloppy-split configurations have at least $n - 1$ submixtures. Thus, for the heavy crude example, we add

$$\sum_{[i,j] \in \mathcal{U}} \zeta_{i,j} = (n - 2) \quad (3.7)$$

We solve the MINLP (H) with (3.6) and (3.7) added. The optimal solution is shown in Figure 3.6(a). This configuration is known in the literature as indirect split configuration, and *for decades, has been the configuration of choice for crude oil separation!* Our model, despite several simplifying assumptions, identifies this optimal configuration in a few seconds among several sharp-split configurations. This result reaffirms that our simplifying assumptions are reasonable and our algorithms generate worthwhile results. However, is the optimal sharp-split configuration the best configuration? To determine the answer to this question, we compare the total vapor duty of FTC, obtained analytically as 69.96 kmol/h, with that of the optimal sharp-split configuration. The indirect split configuration has a vapor duty requirement of 84.402 kmol/hr, which is 20.6% higher than that for FTC. Thus, there is a substantial potential to reduce the total vapor duty. Since FTC is hard to implement, we will expand our search space to include sloppy-split configurations to see if a similar vapor duty is also achievable by a configuration that is easy to implement.

Sloppy-split configurations: Every-sloppy split configuration has at least two columns stacked one on top of the other. By construction, sloppy-split configurations withdraw at least one sidedraw by stacking columns, that produce the same streams one from the stripping section and the other from the rectifying section. There are four possible arrangements for a sidedraw: vapor-only-transfer, liquid-only-transfer, two-way vapor-liquid transfer and two-phase sidedraw. The formulation described in the previous section, captures all the four arrangements. Nevertheless, industrial

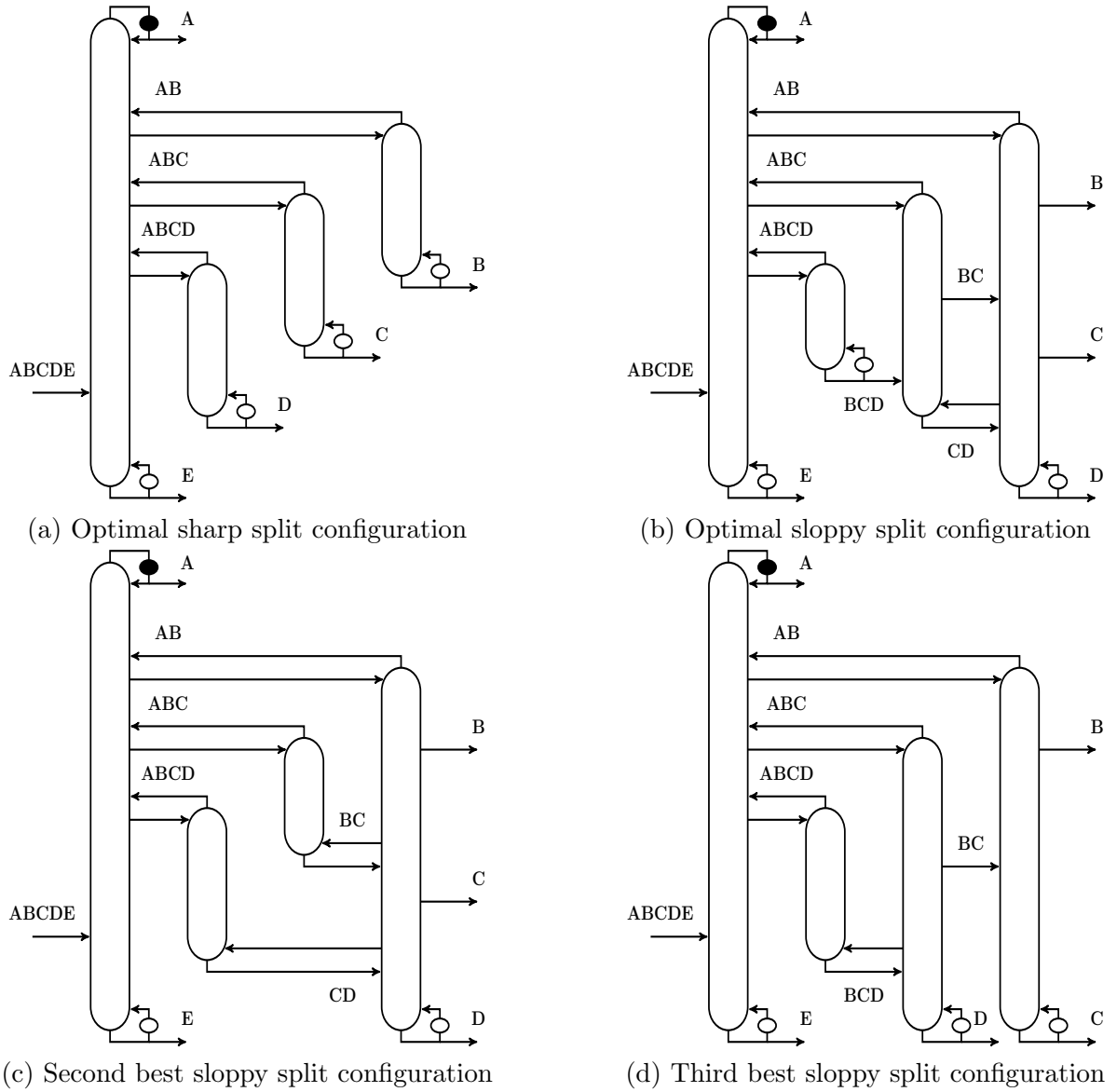


Fig. 3.6. Candidate configurations for the heavy crude example

practitioners prefer that the sidedraw is a liquid-only-transfer because such transfers are easier to operate. To maintain sidedraws as liquid transfer, we add

$$\begin{aligned}
 -V^U[(1 - \zeta_{i,j}) + (1 - \zeta_{i,l}) + (1 - \zeta_{k,j})] &\leq U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}} \leq \\
 V^U[(1 - \zeta_{i,j}) + (1 - \zeta_{i,l}) + (1 - \zeta_{k,j})], & \\
 \forall \quad j+1 \leq l \leq n; \quad 1 \leq k \leq i-1; \quad [i, j] \in \mathcal{S}_2 &
 \end{aligned} \tag{3.8}$$

Here, $(U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}})$ denotes the net vapor inflow into the pseudocolumn $Q_{i,j}$. The constraint forces the net vapor inflow to zero when three streams are present, namely stream $[i, j]$, one of the top parents $\{[i, l]\}_{l=j+1}^n$, and one of the bottom parents $\{[k, j]\}_{k=1}^{i-1}$. We now solve the MINLP (H), restricted to also satisfy (3.6) and (3.8) in order to determine the optimal sloppy-split configuration that produces E as the residue from the first column and draws sidedraws as liquid-only transfers. The optimal configuration has been drawn in an operable form in Figure 3.6(b). The optimal vapor duty of the configuration is 76.76 kmol/h, which is 9% lower than that of indirect split configuration.

If the optimal solution turns out to not be desirable for a reason not captured by the model, the practitioner would like to generate alternative solutions. Towards this end, we organize configurations into *families*, where configurations from each family have same structural arrangement of columns, but differ with each other in whether the connection between adjacent columns is a thermal coupling or via a heat exchanger. Although thermal couplings are known to benefit by conserving energy, such benefits are not always realized [42]. In other words, various configurations within a family have identical vapor duty. Moreover, a given configuration may perform optimally over a wide range of operating conditions. For these reasons, we find that the ‘Numsol’ option of BARON is not reliable in determining $\mathbb{K}$ -best families. BARON often finds the same configuration, among the best solutions, albeit with different operating conditions. Instead, we use binary cuts to systematically determine the $\mathbb{K}$ -best families.

Configurations from different families have a different set of submixtures. Using binary cuts on submixtures, we eliminate previously seen families of configurations.

There are various ways of achieving this result. One way is to consider each ζ vector as a binary encoding of a number and eliminate those solutions where $\zeta_{i,j}$ for $[i, j] \in B\zeta_1$ are at one and $\zeta_{i,j}$ for $[i, j] \in B\zeta_0$ are at zero. This is achieved by adding

$$\sum_{[i,j] \in B\zeta_0} (1 - \zeta_{i,j}) + \sum_{[i,j] \in B\zeta_1} \zeta_{i,j} \leq |B\zeta_0| + |B\zeta_1| - 1 \quad (3.9)$$

We refer the reader to [43] for an alternative approach. Solving the MINLP after adding this cut gives a configuration from the second-best family. The operable version of the solution is shown in Figure 3.6(c). The optimal vapor duty of the configuration is 77.39 kmol/h, which is slightly higher than that for the configuration in Figure 3.6(b), but requires one less transfer stream and two less column sections. To determine the third best configuration, we add a cut similar to the one described in (3.9). The operable version of the third best configuration is shown in Figure 3.6(d). The optimal vapor duty is 1.86% higher than the second best configuration. Interestingly, all the three configurations in Figure 3.6(b)-(d) exhibit similarities with the indirect-split configuration, and are thus amenable to retrofitting. A similar analysis has been performed on the light crude example from [5], and the results can be found in our recent publication [44].

3.2.2 A Pseudocost Model

We have illustrated that energy can be saved using sloppy-splits. In fact, the configuration with the maximum number of sloppy-splits (FTC) is known to require the least vapor duty [13, 34]. However, the energy savings come at a price. The introduction of sloppy-splits increases the number of column sections, which in turn may increase the capital cost. For example, Figure 3.6(b) requires six more column sections compared to Figure 3.6(a). Thus, the sloppy-split configuration may not be more economical since the increase in capital cost outweighs the reduction in operating cost (vapor duty). It is often desirable to identify configurations which reduce vapor duty (measure of operating cost) while keeping the number of column sections reasonably close to that of sharp-split configurations. Although a detailed

cost model can also be formulated, we describe two simple heuristics for identifying cost-effective configurations.

Procedure 1

To identify such configurations, we propose the use of the following objective function:

$$\text{Minimize } \sum_{[i,j] \in \mathcal{T}} \zeta_{i,j} - \sum_{(i,j) \in \mathcal{C}} \chi_{i,j} - \sum_{(i,j) \in \mathcal{R}} \rho_{i,j} \quad (3.10)$$

This objective minimizes the total number of streams (a measure of the number of column sections), and for better controllability maximizes the heat exchangers. Different weights may be used to tilt the emphasis between the two objectives. When two solutions have the same objective, we choose the configuration with least number of streams. Clearly, solving the MINLP (H) with (3.10) as the objective could yield any basic sharp split configuration as the optimal solution since, for these configurations, each submixture stream is associated with a heat exchanger. However, such a solution would be unattractive in terms of vapor duty. To eliminate such solutions, we also require that the vapor duty does not exceed that of FTC by more than a specified percentage, say 10%.

$$\sum_{(i,j) \in \mathcal{R}} FR_{i,j} \leq 1.1 VD_{FTC} \quad (3.11)$$

where VD_{FTC} is the optimal vapor duty of FTC (refer (3.5)). Thus, we propose to solve (H) after appending (3.11) and modifying the objective to (3.10). The $\mathbb{K}$ -best families are determined as before, by iteratively adding binary cuts (3.9).

Procedure 2

We now suggest an alternative procedure to find attractive configurations, which is summarized in Algorithm 1.

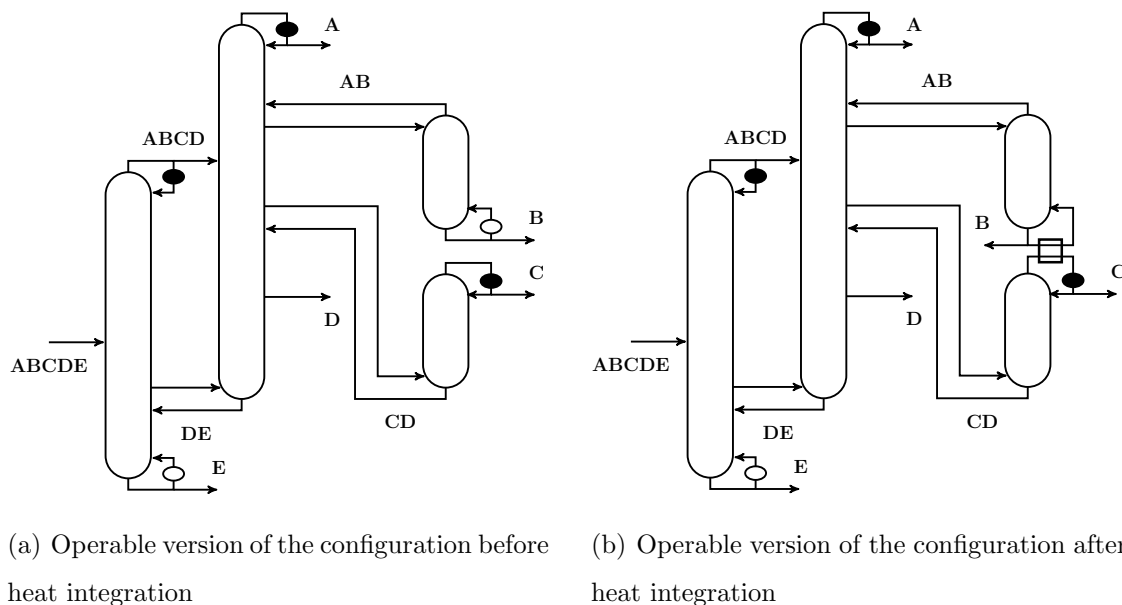


Fig. 3.7. Configuration that minimizes the sloppy splits, maximizes the heat exchangers while being attractive from the perspective of vapor duty. (b) shows the heat integrated version of (a) where the reboiler associated with the side-stripper is replaced with a heat exchanger (thick square)

We now use Procedure 1 on the heavy crude example and describe the results. For illustration, we will also include submixtures $BCDE$, CDE , and DE in the search space. It turns out that the resulting MINLP has multiple solutions with the same objective function value. One of the solutions is shown in Figure 3.7. The vapor duty requirement for this configuration is 10% less than that of indirect split configuration, and is within 10% of that of FTC. This solution has other attractive features as well. First, the configuration has only one sloppy split and just two extra column sections relative to indirect split configuration. As a result, the capital cost is expected to be not too far off from that of indirect split configuration (Figure 3.6(a)). Second, each column has at least one heat exchanger. This allows for better control of reflux in each column and make the startup as well as shut down process easier. Third, heat-integration of side-stripper and side-rectifier is possible because boiling point of C is higher than that of B . The reboiler associated with the side-stripper

Step 1: Formulate MINLP (H), and initialize number of submixtures
 $NS = n - 2$

Step 2: Append $\sum_{[i,j] \in \mathcal{U}} \zeta_{i,j} = NS$ to (H), and solve the MINLP with (H1) as the objective

if *Vapor duty* $< 1.1 VD_{FTC}$ **then**
 | go to Step 3;
else
 | $NS = NS + 1$;
 | go to Step 2;
end

Step 3: A configuration which minimizes number of column sections is found. If the next best solution is desired, append binary cut (3.9) to (H) and go to Step 2 else, stop.

Algorithm 1: Algorithm for identifying cost-effective distillation configurations

generates vapor reflux that contains B , whereas the condenser associated with the side-rectifier generates liquid reflux that contains C . The latent heat released during the condensation of C can be used to supply the latent heat required to vaporize B . We have assumed while using Underwood method that the latent of vaporization is about the same for all components. Under this assumption, the condenser and reboiler duties are directly proportional to their vapor flows. In the current example, it turns out that the vapor flow in the condenser is much higher than the liquid flow in the reboiler. This means that, the reboiler associated with the side-stripper can be replaced with a heat exchanger. Since we withdraw products as liquids, the condenser associated with the side-rectifier was anyway required to condense C . Fortuitously, the vapor duty of the heat integrated configuration (Figure 3.7(b)) matches that of FTC, and hence at *the lowest possible value*! We emphasize that *the maximum benefit* could be attained with just one sloppy-split. For all these reasons, the proposed

configuration may be preferable over the conventional indirect split configuration, even though it introduces the submixture stream DE .

Procedure 2 can also be used to arrive at the same solution. It has been used in our recent work [44] to separate a mixture of alcohols, an example we borrowed from [17]. We had found that this approach also results in configurations that seem attractive from the perspective of both capital and operating costs.

3.2.3 Case Studies from Literature

We have only shown a few ways, out of many, to tailor the MINLP model as per the need of the process designer. Besides heavy crude example, we have tested our model on a few other cases that have been previously reported in the literature. Table 3.4 shows how the problem size varies with the number of components in the process feed. Table 3.5 summarizes the problem data and the results obtained. In all the examples, optimization was performed over the entire search space of regular-column configurations. The computations were done on Dell Optiplex 5040 with 16 GB RAM Intel Core i7-6700 3.4 GHz processor.

The case studies demonstrate that flow of roots constraints, (H28), help expedite the branch-and-bound algorithm. In particular, we bring the attention of the reader to Cases 9 – 14. Flow of roots cuts have an enormous impact on these cases, as is evident from the drastic reduction in the number of iterations and the overall computational time to solve these cases when (H28) was included in the model. Moreover, without these cuts, Cases 10 and 11 could not converge to within 1% optimality gap in the 1200 seconds allotted to the optimization algorithm. We observed that for these cases, the lower bound improved at a rather slow pace.

To further test the effectiveness of our model, we tested it on the 120 four-component cases and 496 five-component cases reported in [35]. For these runs, the relative tolerance for convergence was set to 1% and the time-limit was set at 1200 seconds. The model was solved to ϵ -global optimality by BARON for all the 120

Table 3.4.
Problem size for different systems

	$n = 3$	$n = 4$	$n = 5$	$n = 6$
Binary variables	4	11	21	34
Continuous variables	70	158	292	480
Linear constraints	122	311	617	1073
Nonlinear constraints	10	27	56	100

four-component cases, and for 72.3% of the 496 five-component cases. A few cases that did not converge had a duality gap as high as 82%. Although the model gives insights and worthy configuration alternatives, there is still scope for improvement.

3.3 Concluding Remarks

In this chapter, we introduced a new MINLP formulation to identify energy-efficient distillation configurations. The key contributions are as follows. First, we introduced a new notation that is more natural and easy to work with. Second, we identified that previous MINLP formulations relax the feasible region admitting physically infeasible solutions. We identified the source of the problem to be that Underwood constraints associated with active roots were insufficiently constrained. We showed how to modify these constraints to more accurately model the feasible space. Third, we added valid cuts to model flow of roots, which expedite the branch-and-bound convergence. The effectiveness of the cuts was illustrated with a few examples. Fourth, we proposed a new approach to obtain configurations that reduce operating cost, while being aware that increasing number of sections increases the capital cost, and increasing thermal couplings reduces operability. Finally, we tested our model on a wide-range of cases reported in the literature. While the model could be solved to ϵ -global optimality for a majority of cases, we identified various

Table 3.5.

Case studies from literature. For all the cases, the feed is assumed to be in saturated liquid state ($\Phi = 1$). MINLPs are solved with BARON 18.5.8 [21, 40] on GAMS 25.1 with default values for all options except pDo (pDo is set to -1), and the relative tolerance for convergence (ϵ_r) is set to 1%. Further, the time limit is chosen to be 1200 s for four and five component systems, and 3600 s for six component system. TL corresponds to ‘Time Limit has been reached’, and the number in the parenthesis is the duality gap at the end of 1200 s. I/N/T: Baron Iterations/Maximum number of nodes in memory/Time in seconds. Case 1 is taken from [10], Cases 2, 3 and 4 from [45], Cases 6, 8 and 10 from [9], Cases 7, 9 and 11 from [17]. Relative volatility information for Cases 12 and 13 are taken from [46], and from [47] for Case 14.

Case	n	$\{F_p\}_{p=1}^n$	$\{\alpha_p\}_{p=1}^n$	Without (H28)		With (H28)	
				I/N/T		I/N/T	
1	4	{17.8, 40.4, 27.7, 14.1}	{3.413, 2.146, 1.398, 1}	17/7/4		9/3/2	
2	4	{30, 40, 25, 5}	{29.07, 12.81, 2.35, 1}	15/4/2		13/3/2	
3	4	{5, 25, 40, 30}	{24.19, 4.599, 2.19, 1}	1/1/1		1/1/1	
4	4	{3, 55, 12, 30}	{27.88, 12.28, 4.76, 1}	27/5/3		19/4/2	
5	5	{20, 20, 20, 20, 20}	{39.0625, 15.625, 6.25, 2.5, 1}	29/7/22		29/8/29	
6	5	{10, 10, 40, 30, 10}	{8.9, 5.7, 3.2, 1.55, 1}	35/6/33		24/6/28	
7	5	{10, 20, 30, 20, 20}	{8.9, 5.7, 3.2, 1.55, 1}	29/7/41		29/6/63	
8	5	{10, 10, 40, 30, 10}	{4.1, 3.6, 2.1, 1.42, 1}	7/2/39		9/3/12	
9	5	{20, 30, 20, 20, 10}	{4.1, 3.6, 2.1, 1.42, 1}	189/20/183		61/9/43	
10	5	{15, 15, 40, 20, 10}	{10.5, 4.04, 1.76, 1.31, 1}	TL (8.2%)		245/34/249	
11	5	{4, 6, 50, 35, 5}	{10.5, 4.04, 1.76, 1.31, 1}	TL (8.5%)		367/45/549	
12	5	{5, 15, 25, 20, 35}	{7.98, 3.99, 3, 1.25, 1}	367/34/658		9/3/16	
13	5	{25, 15, 25, 20, 20}	{13.72, 3.92, 3.267, 1.21, 1}	135/16/334		55/9/88	
14	6	{20, 20, 20, 15, 15, 15}	{53.39, 30.51, 14.44, 4.24, 2.16, 1}	455/48/1849		149/36/913	

challenges that guide enhancements reported elsewhere [25]. The proposed method can also be used to identify configurations that are amenable for retrofitting. We have shown that attractive configurations for non-azeotropic mixtures can be systematically determined by solving mathematical models, instead of relying on heuristic procedures. We expect many novel and energy-efficient distillation configurations will result from the application of these techniques to practical distillation problems.

4. IMPROVED FORMULATION AND SOLUTION APPROACH

The computational experiments in the previous chapter show that the state-of-the-art global solvers either fail to converge, or do so very slowly, on several hard instances. Here, we make several enhancements to the formulation and the solution procedure which pave a way towards a reliable solution methodology. The chapter is organized as follows. In §4.1, we define the problem statement and introduce the notation. We formulate the MINLP in §4.2, and outline the overall relaxation and solution procedure in §4.3. We present numerical examples in §4.4 that illustrate the value of our cuts in strengthening the overall relaxation, and report on our computational experiments in §4.5. Finally, we make concluding remarks in §4.6.

4.1 Problem Definition

Figure 4.1 shows all possible streams and heat exchangers in a distillation configuration that separates a four-component mixture into pure components. We represent streams as squares, condensers as filled circles and reboilers as open circles. Each condenser/reboiler is associated with a process stream, that is not the process feed. Throughout the formulation, we denote a stream as $[i, j]$, and heat exchangers as (i, j) , so that condenser (i, j) (*resp.* reboiler (i, j)) represents the heat exchanger through which $[i, j]$ is withdrawn as distillate (*resp.* residue). By Remark 2.1, a configuration cannot contain streams of the form $[i \dots k \ k + l \dots j]$, where $l > 1$.

We denote the set of streams as $\mathcal{T}$, the set of condensers as $\mathcal{C}$, and the set of reboilers as $\mathcal{R}$ (see Table 4.1 for definition). For convenience, we create a set containing streams that are mixtures $\mathcal{P} = \mathcal{T} \setminus \{[1, 1], \dots, [N, N]\}$, and a set containing

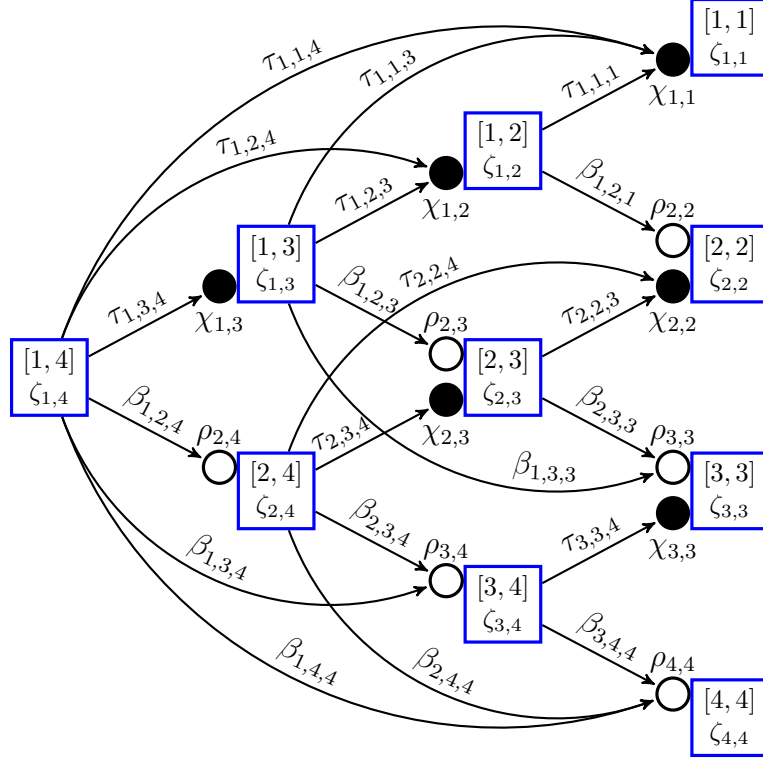


Fig. 4.1. Figure depicting streams ($\zeta_{i,j}$), reboilers ($\rho_{i,j}$) and condensers ($\chi_{i,j}$) present in a four-component system. Section variables $\tau_{i,k,j}$ and $\beta_{i,l,j}$ are defined in (4.1).

submixtures $\mathcal{S} = \mathcal{P} \setminus \{[1, N]\}$. Note that every stream in $\mathcal{P}$ is a mixture, and must undergo a split in order to produce products.

The required input to the problem consists of (1) composition of the process feed $\{F_p\}_{p=1}^N$ either in terms of mole fractions or molar flowrates of the components in the stream, (2) relative volatilities $\{\alpha_p\}_{p=1}^N$ (such that $\alpha_N < \dots < \alpha_1$) of its constituent components; and (3) liquid fraction (fraction of the total flow in liquid phase) of the process feed $\Phi_{1,N}$ and that of the pure components $\{\Phi_{i,i}\}_{i=1}^N$. We write $\{p\}_{p=1}^N$ or $\{p\}_{1 \leq p \leq n}$ to denote the set $\{1, 2, \dots, N\}$, and $\llbracket p \rrbracket_1^N$ to denote $\forall p \in \{1, \dots, N\}$. Given a process feed, the problem is then to identify the best distillation configuration, along with its optimal operating conditions, that requires least vapor duty.

Table 4.1.
Definition of sets.

Set	Symbol	Definition
Streams	$\mathcal{T}$	$\{[i, j] : 1 \leq i \leq j \leq N\}$
Splits	$\mathcal{P}$	$\mathcal{T} \setminus \{[i, i]\}_{i=1}^N$
Submixtures	$\mathcal{S}$	$\mathcal{P} \setminus \{[1, N]\}$
Condensers	$\mathcal{C}$	$\{(i, j) : 1 \leq i \leq j \leq N - 1\}$
Reboilers	$\mathcal{R}$	$\{(i, j) : 2 \leq i \leq j \leq N\}$

4.2 Problem Formulation

We formulate the MINLP in this section. Before proceeding further, we introduce the definition of *parents* and *children* of a stream. By *top* (*resp. bottom*) *parents* of $[i, j]$: we refer to streams $\{[i, n]\}_{n=j+1}^N$ (*resp.* $\{[m, j]\}_{m=1}^{i-1}$) which can produce $[i, j]$ as distillate (*resp.* residue). Analogously, by *top* (*resp. bottom*) *children* of $[i, j]$, we refer to streams $\{[i, k]\}_{k=i}^{j-1}$ (*resp.* $\{[l, j]\}_{l=i+1}^j$) which can be produced as distillate (*resp.* residue) from $[i, j]$. For conciseness, we write $[i, j] \uparrow [i, k]$ (*resp.* $[i, j] \downarrow [l, j]$) to denote stream $[i, k]$ (*resp.* $[l, j]$) is produced as the distillate (*resp.* residue) from $[i, j]$, and $[i, k]/[l, j]$ to denote $[i, k]$ and $[l, j]$ are produced as the distillate and residue from $[i, j]$.

4.2.1 Objective Function

The objective is to determine the configuration(s) which minimizes the total vapor duty:

$$(A): \text{Minimize } \sum_{(i,j) \in \mathcal{R}} FR_{i,j}, \quad (A1)$$

where $FR_{i,j}$ is the vapor flow generated in reboiler (i, j) . The MINLP we develop will be denoted as MINLP (A), and the constraints will be numbered as (A#).

4.2.2 Space of Admissible Configurations

We define *column/stream* binary variables so that $\forall [i, j] \in \mathcal{T}$, $\zeta_{i,j} = 1$ if $[i, j]$ is present and 0 otherwise. Further, we define binary variables associated with the presence/absence of condensers and reboilers so that $\forall (i, j) \in \mathcal{C}$ (*resp.* $\forall (i, j) \in \mathcal{R}$), $\chi_{i,j} = 1$ (*resp.* $\rho_{i,j} = 1$) if condenser (*resp.* reboiler) (i, j) is present and 0 otherwise (See Table 4.1 for set definitions). Although these variables suffice [23], we introduce auxiliary variables to derive a tighter representation.

For every $[i, j] \in \mathcal{P}$, we define *section* variables $\{\tau_{i,k,j}\}_{k=i}^{j-1}$ and $\{\beta_{i,l,j}\}_{l=i+1}^j$, such that $\tau_{i,k,j} = \{1, \text{ if } [i, j] \uparrow [i, k]; 0, \text{ otherwise}\}$ and $\beta_{i,l,j} = \{1, \text{ if } [i, j] \downarrow [l, j]; 0, \text{ otherwise}\}$. In other words, section variables model distillate and residue streams from a mixture. Figure 4.1 shows all the section variables for a four-component mixture. We now relate column and section variables. Consider the split of stream $[i, j]$. In configurations of interest, known as *regular-column configurations*, if $[i, j] \uparrow [i, k]$, for any $i \leq k \leq j-1$, then $[i, j]$ and $[i, k]$ must be present and $\{[i, n]\}_{n=k+1}^{j-1}$ must be absent [9, 10]. Analogously, if $[i, j] \downarrow [l, j]$, for any $i+1 \leq l \leq j$, $[i, j]$ and $[l, j]$ must be present, while $\{[m, j]\}_{m=i+1}^{l-1}$ must be absent. Therefore, section variables are defined as

$$\begin{aligned} \tau_{i,k,j} &= \zeta_{i,j}(1 - \zeta_{i,j-1}) \dots (1 - \zeta_{i,k+1})\zeta_{i,k} \\ &= \prod_{n=k+1}^{j-1} (1 - \zeta_{i,n}) - \prod_{n=k}^{j-1} (1 - \zeta_{i,n}) - \prod_{n=k+1}^j (1 - \zeta_{i,n}) + \prod_{n=k}^j (1 - \zeta_{i,n}), \end{aligned} \quad (4.1)$$

$$\begin{aligned} \beta_{i,l,j} &= \zeta_{i,j}(1 - \zeta_{i+1,j}) \dots (1 - \zeta_{l-1,j})\zeta_{l,j} \\ &= \prod_{m=i+1}^{l-1} (1 - \zeta_{m,j}) - \prod_{m=i}^{l-1} (1 - \zeta_{m,j}) - \prod_{m=i+1}^l (1 - \zeta_{m,j}) + \prod_{m=i}^l (1 - \zeta_{m,j}). \end{aligned}$$

We introduce variables $\{\nu_{i,k,j} : 1 \leq i \leq k \leq j \leq N\}$ and $\{\omega_{i,l,j} : 1 \leq i \leq l \leq j \leq N\}$ to linearize (4.1):

$$\text{for } [i, j] \in \mathcal{P} \quad \begin{cases} \tau_{i,k,j} = \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j}, & \llbracket k \rrbracket_i^{j-1} \\ \beta_{i,l,j} = \omega_{i+1,l-1,j} - \omega_{i,l-1,j} - \omega_{i+1,l,j} + \omega_{i,l,j}, & \llbracket l \rrbracket_{i+1}^j \end{cases} \quad (\text{A2})$$

where $\nu_{i,k,j} = \prod_{n=k}^j (1 - \zeta_{i,n})$ and $\omega_{i,l,j} = \prod_{m=i}^l (1 - \zeta_{m,j})$. Note that $\nu_{i,k+1,j-1}$ (*resp.* $\omega_{i+1,l-1,j}$) are defined as one if $k+1 = j$ (*resp.* $i+1 = l$). Clearly, $\nu_{i,k,j} = \omega_{i,l,j} = 1 - \zeta_{i,j}$

if $k = j$ and $l = i$. Besides this relationship, the introduced variables $\nu_{i,k,j}$ and $\omega_{i,l,j}$ are linearly independent. To see this, note that $\prod_{j \in J} x_j$, where $J \subseteq \{1, \dots, n\}$ are linearly independent and, therefore, so are $\prod_{j \in J} (1 - y_j)$, where $y_j = 1 - x_j$. Since $\nu_{i,k,j}$ and $\omega_{i,l,j}$ are of the latter form, they are linear independent.

We now relax $\nu_{i,k,j}$ and $\omega_{i,l,j}$ variables for $k \neq j$ and $l \neq i$ as follows. Since $\zeta_{i,j}$ is binary, $(1 - \zeta_{i,j})^2 = (1 - \zeta_{i,j})$. We use the definition of $\nu_{i,k,j}$ and $\omega_{i,l,j}$, to derive the following:

$$\text{for } [i, j] \in \mathcal{P} \quad \begin{cases} \nu_{i,k,j} = \nu_{i,k,m} \nu_{i,n,j}, & \llbracket n \rrbracket_{k+1}^{m+1}, \llbracket m \rrbracket_k^{j-1}, \llbracket k \rrbracket_i^{j-1} \\ \omega_{i,l,j} = \omega_{i,m,j} \omega_{n,l,j}, & \llbracket n \rrbracket_{i+1}^{m+1}, \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^j. \end{cases} \quad (4.2)$$

In the above, for $n \leq m+1$, $\nu_{i,n,m}$ (*resp.* $\omega_{n,m,j}$) is a common factor for both $\nu_{i,k,m}$ and $\nu_{i,n,j}$ (*resp.* $\omega_{i,m,j}$ and $\omega_{n,l,j}$). we regard $\nu_{i,n,m}$ and $\omega_{n,m,j}$ as one if $n = m+1$. Thus, $0 \leq \nu_{i,k,m} \leq \nu_{i,n,m}$, $0 \leq \nu_{i,n,j} \leq \nu_{i,n,m}$, $0 \leq \omega_{i,m,j} \leq \omega_{n,m,j}$, and $0 \leq \omega_{n,l,j} \leq \omega_{n,m,j}$. Using these bounds, we relax (4.2) as:

$$\text{for } [i, j] \in \mathcal{P} \quad \begin{cases} \nu_{i,j,j} = \omega_{i,i,j} = 1 - \zeta_{i,j} \\ \max\{0, \nu_{i,k,m} + \nu_{i,n,j} - \nu_{i,n,m}\} \leq \nu_{i,k,j} \leq \min\{\nu_{i,k,m}, \nu_{i,n,j}\}, \\ \llbracket n \rrbracket_{k+1}^{m+1}, \llbracket m \rrbracket_k^{j-1}, \llbracket k \rrbracket_i^{j-1} \\ \max\{0, \omega_{i,m,j} + \omega_{n,l,j} - \omega_{n,m,j}\} \leq \omega_{i,l,j} \leq \min\{\omega_{i,m,j}, \omega_{n,l,j}\}, \\ \llbracket n \rrbracket_{i+1}^{m+1}, \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^j, \end{cases} \quad (A3)$$

where we used $\nu_{i,k,m} = \nu_{i,k,m} \nu_{i,n,m}$, $\nu_{i,n,j} = \nu_{i,n,j} \nu_{i,n,m}$, $\omega_{i,m,j} = \omega_{i,m,j} \omega_{n,m,j}$, and $\omega_{n,l,j} = \omega_{n,l,j} \omega_{n,m,j}$.

Proposition 4.1. *Let $S = \{(x, z) \in [0, 1]^{2n} \mid z_j = \prod_{k=1}^j x_k, \llbracket j \rrbracket_1^n\}$. The convex hull of S , $\text{Conv}(S)$, is the intersection of convex hulls of $z_j = z_{j-1} \cdot x_j$, $\llbracket j \rrbracket_2^n$ over $[0, 1]^2$ (McCormick relaxation).*

Proof. See §B.1 in the appendix. □

We remark that the result in Proposition 4.1 also follows from Theorem 10 in [48]. Our proof is, however, different and elementary. We mention that this proof shows a

previously unobserved connection to the recursive McCormick procedure. Our proof can be used to show that the recursive McCormick procedure, with a few additional linearization variables, yields the convex hull of the multilinear polytopes for γ -acyclic hypergraphs, as obtained in [48].

Remark 4.1. Proposition 4.1 shows that the set of ν (*resp.* ω) variables satisfying (A3) belong to the intersection of simultaneous convex hulls of $(\nu_{i,j,j+1}, \dots, \nu_{i,j,N}, \nu_{i,j,j}, \dots, \nu_{i,N,N})$ for all $[i, j] \in \mathcal{T} \setminus \{[k, N]\}_{k=1}^N$ (*resp.* $(\omega_{1,2,j}, \dots, \omega_{1,i,j}, \omega_{1,1,j}, \dots, \omega_{i,i,j})$ for all $[i, j] \in \mathcal{T} \setminus \{[1, l]\}_{l=1}^N$). $\square$

Remark 4.2. For every $[i, j] \in \mathcal{P}$, $\llbracket k \rrbracket_i^{j-1}$ (*resp.* $\llbracket l \rrbracket_{i+1}^j$), the convex hull of $\tau_{i,k,j}$ (*resp.* $\beta_{i,l,j}$) over $(\zeta_{i,k}, \dots, \zeta_{i,j}) \in [0, 1]^{j-k+1}$ (*resp.* $(\zeta_{i,j}, \dots, \zeta_{l,j}) \in [0, 1]^{l-i+1}$) is implied by (A2) and (A3). (see §B.2 for the proof). $\square$

We now describe the constraints to model the space of admissible distillation configurations.

Presence of process feed and products

Every admissible configuration has the process feed $([1, N])$ and the pure components $(\{[i, i]\}_{i=1}^N)$, *i.e.*,

$$\zeta_{1,N} = \zeta_{1,1} = \dots \zeta_{N,N} = 1. \quad (\text{A4})$$

To restrict the search to a subset of configurations, for example, in order to retrofit an existing design, we may explicitly include (*resp.* eliminate) a specific submixture $[i, j]$ by setting $\zeta_{i,j} = 1$ (*resp.* $\zeta_{i,j} = 0$). We show next that $\zeta_{i,j}$ variables are affinely related to $\tau_{i,k,j}$ and $\beta_{i,l,j}$ variables.

Proposition 4.2. *Let, $x \in [0, 1]^n$, $y_{i,j} = (1-x_i)x_{i+1} \dots x_{j-1}(1-x_j)$ for $1 \leq i < j \leq n$, $z_{i,j} = \prod_{r=i}^j x_r$ for $1 \leq i \leq j \leq n$, and $x_n = 0$, which in turn implies that $z_{i,n} = 0$ for $1 \leq i \leq n$. Then, there is an invertible affine transformation between $\{y_{i,j}\}_{1 \leq i < j \leq n}$ and $\{z_{i,j}\}_{1 \leq i \leq j \leq n}$, given by*

$$y_{i,j} = z_{i+1,j-1} - z_{i+1,j} - z_{i,j-1} + z_{i,j},$$

$$z_{p,q} = 1 - \sum_{r=p}^q \sum_{s=q+1}^n y_{r,s}.$$

Proof. First, we show that $y_{i,j}$ can be written as an affine transformation of $z_{i,j}$. By definition, $y_{i,j} = (1 - x_i)x_{i+1} \dots x_{j-1}(1 - x_j) = \prod_{r=i+1}^{j-1} x_r - \prod_{r=i+1}^j x_r - \prod_{r=i}^{j-1} x_r + \prod_{r=i}^j x_r = \prod_{r=i+1}^{j-1} x_r - z_{i+1,j} - z_{i,j-1} + z_{i,j}$. Substituting the first term in the last equality, $\prod_{r=i+1}^{j-1} x_r$, with 1 if $i+1 = j$, and $z_{i+1,j-1}$ if $i+1 < j$, yields the required affine transformation.

Next, to obtain the inverse affine transformation, we define $w_{k,l} = (1 - x_k)x_{k+1} \dots x_l$ for $1 \leq k \leq l \leq n$. We show the affine transformation between $\{w_{k,l}\}_{1 \leq k \leq l \leq n}$ and $\{y_{i,j}\}_{1 \leq i < j \leq n}$ variables to be

$$w_{k,l} = \sum_{r=l+1}^n y_{k,r}, \quad (4.3)$$

using induction on $n - l$. For $l = n$, (4.3) is trivially satisfied because $w_{k,n} = 0$ as $x_n = 0$. Now, assuming that (4.3) holds for $l+1$, i.e., $w_{k,l+1} = \sum_{r=l+2}^n y_{k,r}$, we show that it holds for $w_{k,l}$ as well: $w_{k,l} = (1 - x_k)x_{k+1} \dots x_l(1 - x_{l+1} + x_{l+1}) = y_{k,l+1} + w_{k,l+1} = y_{k,l+1} + \sum_{r=l+2}^n y_{k,r} = \sum_{r=l+1}^n y_{k,r}$.

In a similar vein, we show for $1 \leq p \leq q \leq n$, the affine transformation between $\{z_{p,q}\}$ and $\{w_{k,l}\}$ variables to be

$$z_{p,q} = 1 - \sum_{r=p}^q w_{r,q}, \quad (4.4)$$

using induction on $q - p$. For $q = p$, (4.4) follows because $z_{q,q} = x_q = 1 - (1 - x_q) = 1 - w_{q,q}$. Next, assuming (4.4) holds for $p+1$ i.e., $z_{p+1,q} = 1 - \sum_{r=p+1}^q w_{r,q}$, we show that it holds for $z_{p,q}$ as well: $z_{p,q} = \prod_{r=p}^q x_r = [1 - (1 - x_p)] \prod_{r=p+1}^q x_r = z_{p+1,q} - w_{p,q} = 1 - \sum_{r=p+1}^q w_{r,q} - w_{p,q} = 1 - \sum_{r=p}^q w_{r,q}$. Finally, substituting (4.3) in (4.4) leads to the required inverse affine transformation given below:

$$z_{p,q} = 1 - \sum_{r=p}^q \sum_{s=q+1}^n y_{r,s}. \quad (4.5)$$

Indeed, the correctness of (4.5) can be checked via direct verification using $y_{r,s} = z_{r+1,s-1} - z_{r+1,s} - z_{r,s-1} + z_{r,s}$, $z_{i,n} = 0$ for $1 \leq i \leq n$, and $z_{i+1,i} = 1$ for $1 \leq i \leq n$. $\square$

We note that Proposition 4.2 shows, by defining $n = N - i + 1$ (*resp.* $n = j$) and $x_r = 1 - \zeta_{i,N-r+1}$ (*resp.* $x_r = 1 - \zeta_{r,j}$), there is an invertible linear transformation between $\{\tau_{i,k,j}\}_{i \leq k < j \leq N}$ and $\{\nu_{i,k,j}\}_{i \leq k \leq j \leq N}$ (*resp.* $\{\beta_{i,l,j}\}_{1 \leq i < l \leq j}$ and $\{\omega_{i,l,j}\}_{1 \leq i < l \leq j}$). We expressed τ (*resp.* β) as an affine function of ν (*resp.* ω) in (A2). The inverse transformation is:

$$\text{for } [i, j] \in \mathcal{T}, \quad \nu_{i,k,j} = \begin{cases} 0, & \text{for } k = i \\ 1 - \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s}, & \text{for } i + 1 \leq k \leq j \end{cases}, \quad (4.6)$$

$$\text{for } [i, j] \in \mathcal{T}, \quad \omega_{i,l,j} = \begin{cases} 1 - \sum_{r=i}^l \sum_{s=l+1}^j \beta_{r,s,j}, & \text{for } i \leq l \leq j - 1 \\ 0, & \text{for } l = j. \end{cases} \quad (4.7)$$

Since $\nu_{i,j,j} = \omega_{i,i,j} = 1 - \zeta_{i,j}$, Corollary 4.1 follows directly from (4.6) and (4.7).

Corollary 4.1. (A2)–(A4) imply that $\sum_{k=i}^{j-1} \tau_{i,k,j} = \sum_{l=i+1}^j \beta_{i,l,j} = \zeta_{i,j}$ for all $[i, j] \in \mathcal{P}$. $\square$

Conservation of components

Corollary 4.1 has the physical interpretation that the stream $[i, j]$, when present, produces exactly one stream as distillate and one stream as residue. However, the distillate and residue streams cannot be chosen arbitrarily. They must be chosen such that, all components are conserved when $[i, j]$ undergoes a split. In other words, for $\llbracket k \rrbracket_i^{j-1}$ (*resp.* $\llbracket l \rrbracket_{i+1}^j$), if $[i, j] \uparrow [i, k]$ (*resp.* $[i, j] \downarrow [l, j]$), then for conservation of components, the residue (*resp.* distillate) from $[i, j]$ must be one of $\{[l, j]\}_{l=i+1}^{k+1}$ (*resp.* $\{[i, k]\}_{k=l-1}^{j-1}$). Consider the digraph shown in Figure 4.2 for stream $[i, j]$.

We partition the nodes into four sets D_1 through D_4 , where $D_1 = \{i\}$ (*resp.* $D_4 = \{j\}$), and $D_2 = \{k\}_{k=i}^{j-1}$ (*resp.* $D_3 = \{l\}_{l=i+1}^j$) contains the heaviest (*resp.* lightest) component in the top (*resp.* bottom) children of $[i, j]$. The edges in $D_1 \times D_2$ (*resp.* $D_3 \times D_4$) correspond to all plausible distillate (*resp.* residue) streams from $[i, j]$. Edges in $D_2 \times D_3$ correspond to feasible *splits* of $[i, j]$, *i.e.*, each node $k \in$

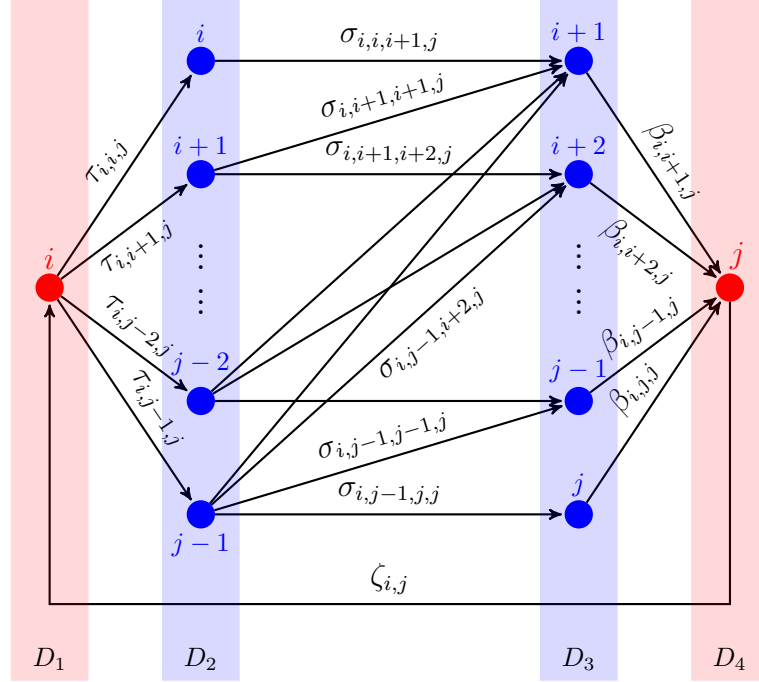


Fig. 4.2. Digraph for deriving conservation of components constraint in §4.2.2

D_2 connects to $\{i+1, \dots, k+1\} \in D_3$. We associate these edges with auxiliary variables $\bigcup_{k=i}^{j-1} \{\sigma_{i,k,l,j}\}_{l=i+1}^{k+1}$, referred as *split* variables hereafter (see Figure 4.2). We let $\sigma_{i,k,l,j} = \{1, \text{ if } [i,k]/[l,j]; 0, \text{ otherwise}\}$, and write mass balances on the network by interpreting stream, section and split variables as material flows along the respective edges of the graph.

$$\text{For } [i,j] \in \mathcal{P} \quad \begin{cases} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} = \tau_{i,k,j}, \llbracket k \rrbracket_i^{j-1}; & \sum_{k=l-1}^{j-1} \sigma_{i,k,l,j} = \beta_{i,l,j}, \llbracket l \rrbracket_{i+1}^j; \\ \sigma_{i,k,l,j} \geq 0, \llbracket l \rrbracket_{i+1}^{k+1}, \llbracket k \rrbracket_i^{j-1}. \end{cases} \quad (\text{A5})$$

Mass balances around the nodes in D_1 and D_4 , and non-negativity constraint on section variables are implied from (A2)–(A4) (see Corollary 4.1 and Remark 4.2), so it is not required to impose them explicitly. We show below that, for any $[i,j] \in \mathcal{P}$, the relaxation (A2)–(A5) is the best possible for the substructure represented by the digraph in Figure 4.2.

Proposition 4.3. *The constraints (A2)–(A5), and $0 \leq \zeta_{i,j} \leq 1$ define a set such that, for any $[i, j] \in \mathcal{P}$, $(\sigma, \tau, \beta, \zeta)$ is contained in the convex hull of*

$$S_{i,j} = \left\{ (\sigma, \tau, \beta, \zeta) \left| \begin{array}{ll} \sigma_{i,k,l,j} = \tau_{i,k,j} \beta_{i,l,j}, & \llbracket l \rrbracket_{i+1}^{k+1}; \quad \llbracket k \rrbracket_i^{j-1} \\ \tau_{i,k,j} \beta_{i,l,j} = 0, & \llbracket l \rrbracket_{k+2}^j; \quad \llbracket k \rrbracket_i^{j-2} \\ \sum_{k=i}^{j-1} \tau_{i,k,j} = \sum_{l=i+1}^j \beta_{i,l,j} = \zeta_{i,j}, & \\ \tau_{i,k,j}, \beta_{i,l,j}, \zeta_{i,j} \in \{0, 1\} & \llbracket l \rrbracket_{i+1}^j; \quad \llbracket k \rrbracket_i^{j-1} \end{array} \right. \right\}. \quad (4.8)$$

Proof. First, note that (A5), equations in Corollary 4.1, $0 \leq \zeta_{i,j} \leq 1$ and non-negativity of section variables together constitute a network flow polytope (see Figure 4.2) in $(\tau, \beta, \sigma, \zeta)$ space. The extreme points of the polytope are integral, and are given by

$$\left. \begin{array}{l} \zeta_{i,j} = \tau_{i,k,j} = \beta_{i,l,j} = \sigma_{i,k,l,j} = 1, \\ \tau_{i,k',j} = \beta_{i,l',j} = 0, \quad \text{for } k' \neq k, l' \neq l \end{array} \right\} \quad \llbracket l \rrbracket_{i+1}^{k+1}; \quad \llbracket k \rrbracket_i^{j-1}, \quad (4.9a)$$

$$\zeta_{i,j} = \tau_{i,k,j} = \beta_{i,l,j} = \sigma_{i,k,l,j} = 0. \quad (4.9b)$$

We show that the only solutions to $S_{i,j}$ are those in (4.9a) and (4.9b). Assume $\zeta_{i,j} = 0$. Then, $\tau_{i,k,j} = 0$ for $\llbracket k \rrbracket_i^{j-1}$, $\beta_{i,l,j} = 0$ for $\llbracket l \rrbracket_{i+1}^j$ and $\sigma_{i,k,l,j} = 0$ for $\llbracket l \rrbracket_{i+1}^{k+1}; \llbracket k \rrbracket_i^{j-1}$. Now, assume $\zeta_{i,j} = 1$. Then, there exists k and l satisfying $i < l \leq k+1 \leq j$ such that $\tau_{i,k,j} = \beta_{i,l,j} = \sigma_{i,k,l,j} = 1$ and for $k' \neq k, l' \neq l$; $\tau_{i,k',j} = \beta_{i,l',j} = \sigma_{i,k',l',j} = 0$. $\square$

Presence of a parent

Stream $[i, j] \in \mathcal{T} \setminus \{[1, N]\}$ is present in a configuration, only if it is produced as a distillate from one of its top parents and/or as a residue from one of its bottom parents. To derive the required constraints, we consider the digraph shown in Figure 4.3.

The graph is inspired from the observation that $\sum_{n=j+1}^{N+1} \tau_{i,j,n} = \zeta_{i,j}$ and $\sum_{m=0}^{i-1} \beta_{m,i,j} = \zeta_{i,j}$, where we define $\tau_{i,j,N+1} = \nu_{i,j+1,N} - \nu_{i,j,N}$ and $\beta_{0,i,j} = \omega_{1,i-1,j} - \omega_{1,i,j}$. From (A3),

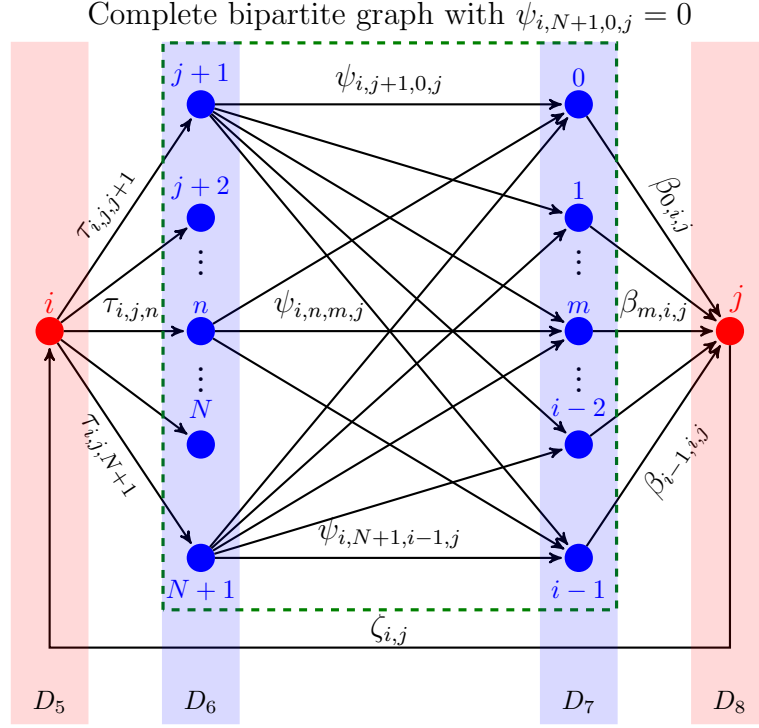


Fig. 4.3. Digraph for deriving presence of parent constraint in §4.2.2

it can be verified that $0 \leq \tau_{i,j,N+1} \leq 1$ and $0 \leq \beta_{0,i,j} \leq 1$. Physically, $\tau_{i,j,N+1} = 1$ (*resp.* $\beta_{0,i,j} = 1$) indicates that $[i, j]$ is not produced as distillate (*resp.* residue), because $\tau_{i,j,N+1} = 1$ (*resp.* $\beta_{0,i,j} = 1$) *iff* $[i, j]$ is present ($\zeta_{i,j} = 1$) and all its top (*resp.* bottom) parents are absent i.e., $\nu_{i,j+1,N} = 1$ (*resp.* $\omega_{1,i-1,j} = 1$).

As in §4.2.2, we partition the nodes into four sets D_5 through D_8 (see Figure 4.3), where $D_5 = \{i\}$ (*resp.* $D_8 = \{j\}$), and $D_6 = \{n\}_{n=j+1}^{N+1}$ (*resp.* $D_7 = \{m\}_{m=0}^{i-1}$) contains the heaviest (*resp.* lightest) component in the top (*resp.* bottom) parents of $[i, j]$. Recall that $m = 0$ and $n = N + 1$ have a special meaning as described in the previous paragraph. The edges in $D_5 \times D_6$ (*resp.* $D_7 \times D_8$) correspond to all plausible ways $[i, j]$ can be produced as distillate (*resp.* residue), and the edges in $D_6 \times D_7$ indicate whether $[i, j]$ is produced only as distillate or only as residue or both. We introduce variables for edges in $D_6 \times D_7$ such that $\psi_{i,n,m,j} = 1$ *iff* $[i, n] \uparrow [i, j]$ and $[m, j] \downarrow [i, j]$.

We require that $\psi_{i,N+1,0,j} = 0$, which, otherwise, would mean that $[i, j]$ can be present even if it is neither produced as distillate nor as residue. Now, we write mass balances on the network.

$$\text{for } [i, j] \in \mathcal{T} \setminus \{[1, N]\} \quad \begin{cases} \sum_{m=0}^{i-1} \psi_{i,n,m,j} = \tau_{i,j,n}, \llbracket n \rrbracket_{j+1}^{N+1}; & \sum_{n=j+1}^{N+1} \psi_{i,n,m,j} = \beta_{m,i,j}, \llbracket m \rrbracket_0^{i-1}; \\ \psi_{i,n,m,j} \geq 0, \llbracket n \rrbracket_{j+1}^{N+1}, \llbracket m \rrbracket_0^{i-1}; & \psi_{i,N+1,0,j} = 0. \end{cases} \quad (4.10)$$

Mass balances around the nodes in D_5 and D_8 , and non-negativity constraint on section variables are implied from (A2) and (A3), so it is not required to impose them explicitly.

Proposition 4.4. *The constraints (A2), (A3), (4.10) and $0 \leq \zeta_{i,j} \leq 1$ define a set such that, for every $[i, j] \in \mathcal{T} \setminus \{[1, N]\}$, $(\tau, \beta, \zeta, \psi)$ is contained in the convex hull of*

$$S_{i,j} = \left\{ (\tau, \beta, \zeta, \psi) \left| \begin{array}{ll} \psi_{i,n,m,j} = \tau_{i,j,n} \beta_{m,i,j}, & \llbracket m \rrbracket_0^{i-1}; \quad \llbracket n \rrbracket_{j+1}^{N+1} \\ \sum_{n=j+1}^{N+1} \tau_{i,j,n} = \sum_{m=0}^{i-1} \beta_{m,i,j} = \zeta_{i,j}, & \psi_{i,N+1,0,j} = 0, \\ \tau_{i,j,n}, \beta_{m,i,j}, \zeta_{i,j} \in \{0, 1\} & \llbracket m \rrbracket_0^{i-1}; \quad \llbracket n \rrbracket_{j+1}^{N+1} \end{array} \right. \right\}. \quad (4.11)$$

Proof. We use a similar argument as the one used to prove Proposition 4.3. We recognize that (4.10), $\sum_{n=j+1}^{N+1} \tau_{i,j,n} = \sum_{m=0}^{i-1} \beta_{m,i,j} = \zeta_{i,j}$, $0 \leq \zeta_{i,j} \leq 1$ and non-negativity requirement on section variables together constitute a network flow polytope, whose extreme points are integral and precisely those in $S_{i,j}$. $\square$

Constraints on Heat Exchanger Variables

For every $(i, j) \in \mathcal{C}$, condenser (i, j) is present only if the stream $[i, j]$ is not produced as residue, *i.e.*, $\beta_{0,i,j} = 1$ [23]. Similarly, for every $(i, j) \in \mathcal{R}$, reboiler (i, j) is present only if the stream $[i, j]$ is not produced as distillate, *i.e.*, $\tau_{i,j,N+1} = 1$.

Further, a condenser (*resp.* reboiler) must be present with a pure component $[i, i]$, if $[i, i]$ is not produced as residue (*resp.* distillate) i.e. $\beta_{0,i,i} = 1$ (*resp.* $\tau_{i,i,N+1} = 1$).

$$\chi_{i,j} \leq \beta_{0,i,j}, \forall (i, j) \in \mathcal{C}; \quad \rho_{i,j} \leq \tau_{i,j,N+1}, \forall (i, j) \in \mathcal{R} \quad (\text{A6})$$

$$\chi_{i,i} \geq \beta_{0,i,i}, \forall (i, i) \in \mathcal{C}; \quad \rho_{i,i} \geq \tau_{i,i,N+1}, \forall (i, i) \in \mathcal{R}. \quad (\text{A7})$$

Proposition 4.5. *The constraints (A2)–(A7), (4.10), $0 \leq \zeta_{i,j} \leq 1$, $\chi_{i,j} \geq 0$ and $\rho_{i,j} \geq 0$ define a set that, for every $[i, j] \in \mathcal{S}$, is contained in the convex hull of solutions that satisfy at least one of the following conditions, where unspecified $\tau_{i,\cdot,j}$, $\beta_{i,\cdot,j}$, $\sigma_{i,\cdot,j}$, $\psi_{i,\cdot,j}$, $\chi_{i,j}$, and $\rho_{i,j}$ variables are zero:*

1. *for some $1 \leq m \leq i-1$, $j+1 \leq n \leq N$, and $i < l \leq k+1 \leq j$, we have*

$$\zeta_{i,j} = \tau_{i,k,j} = \beta_{i,l,j} = \sigma_{i,k,l,j} = \tau_{i,j,n} = \beta_{m,i,j} = \psi_{i,n,m,j} = 1,$$
2. *for some $j+1 \leq n \leq N$, and $i < l \leq k+1 \leq j$, we have $\zeta_{i,j} = \tau_{i,k,j} = \beta_{i,l,j} = \sigma_{i,k,l,j} = \tau_{i,j,n} = \beta_{0,i,j} = \psi_{i,n,0,j} = 1$; $\chi_{i,j} = 1$ or 0 ,*
3. *for some $1 \leq m \leq m-1$, and $i < l \leq k+1 \leq j$, we have $\zeta_{i,j} = \tau_{i,k,j} = \beta_{i,l,j} = \sigma_{i,k,l,j} = \tau_{i,j,N+1} = \beta_{m,i,j} = \psi_{i,N+1,m,j} = 1$; $\rho_{i,j} = 1$ or 0 ,*
4. *all the variables are zero.*

Proof. We modify the graph in Figure 4.3 to accommodate (A6) and (A7), and combine it with the graph in Figure 4.2. The resulting graph is shown in Figure 4.4. Next, observe that (A5), (4.10), $\sum_{k=i}^{j-1} \tau_{i,k,j} = \sum_{l=i+1}^j \beta_{i,l,j} = \sum_{n=j+1}^{N+1} \tau_{i,j,n} = \sum_{m=0}^{i-1} \beta_{m,i,j} = \zeta_{i,j}$, $0 \leq \zeta_{i,j} \leq 1$ (which are implied from (A2)–(A4)), and non-negative constraint on all variables together constitute a network flow polytope. The extreme points this polytope are integral, and are precisely those mentioned in the Proposition. $\square$

Since ψ variables are not used elsewhere, we project (4.10) to the space of section variables (τ, β) .

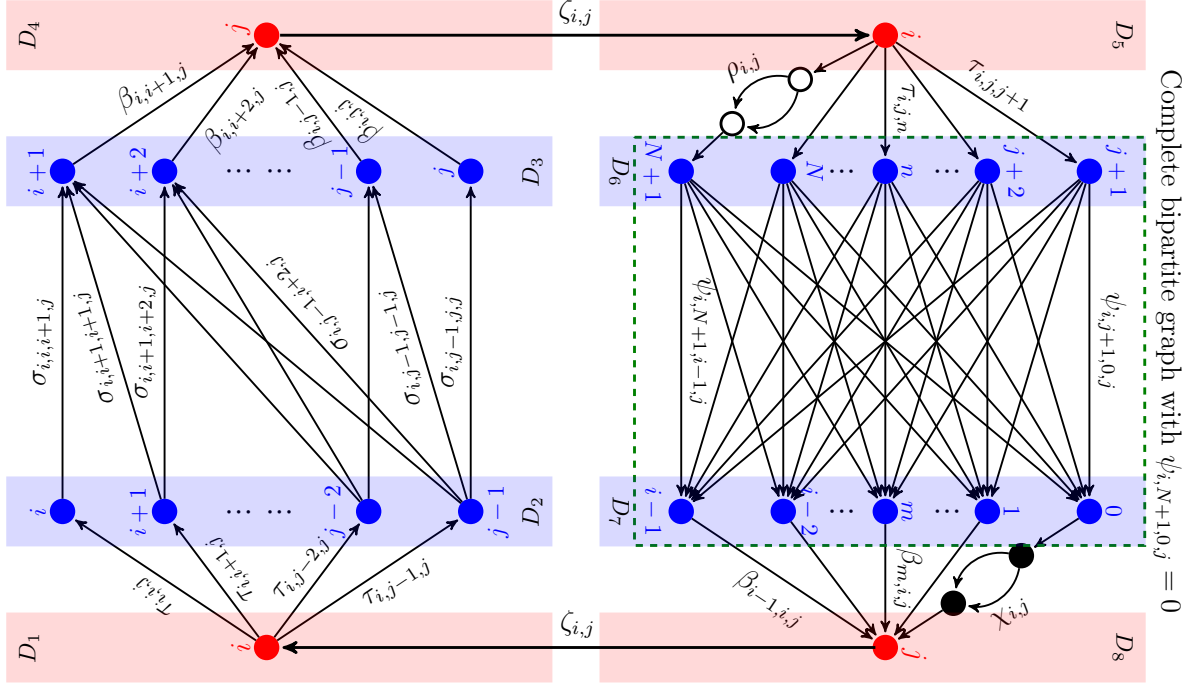


Fig. 4.4. Digraph for the proof of Proposition 4.5

Proposition 4.6. For every $[i, j] \in \mathcal{T} \setminus \{[1, N]\}$, let $S_{i,j} = \{(\tau, \beta, \psi) \mid (4.10); \sum_{m=0}^{i-1} \beta_{m,i,j} = \sum_{n=j+1}^{N+1} \tau_{i,j,n}; \tau_{i,j,n} \geq 0, \llbracket n \rrbracket_{j+1}^{N+1}; \beta_{m,i,j} \geq 0, \llbracket m \rrbracket_0^{i-1}\}$. Then, the projection of $S_{i,j}$ in (τ, β) space is

$$\text{proj}_{(\tau, \beta)}(S_{i,j}) = \left\{ (\tau, \beta) \left| \begin{array}{l} \beta_{0,i,j} \leq \sum_{n=j+1}^N \tau_{i,j,n}; \quad \sum_{m=0}^{i-1} \beta_{m,i,j} = \sum_{n=j+1}^{N+1} \tau_{i,j,n} \\ \tau_{i,j,n} \geq 0, \llbracket n \rrbracket_{j+1}^{N+1}; \quad \beta_{m,i,j} \geq 0, \llbracket m \rrbracket_0^{i-1} \end{array} \right. \right\}. \quad (4.12)$$

Proof. See §B.3 in the Appendix. $\square$

Apart from the following, the remaining constraints in (4.12) follow from (A2) and (A3):

$$\text{for } [i, j] \in \mathcal{T} \setminus \{[1, N]\}, \quad \beta_{0,i,j} \leq \sum_{n=j+1}^N \tau_{i,j,n}. \quad (\text{A8})$$

Remark 4.3. Using (A5), (4.6) and (4.7), τ , β , ν and ω variables can be substituted out. (The resulting model in the lower dimensional space can be found in the chapter).

□

Constraints (A4)–(A8) model the space of admissible configurations. We compare this formulation with CG06, GA10, and TAT19, which refer to the formulations of Caballero and Grossmann [9], Giridhar and Agrawal [10], and Tumbalam Gooty et al. [23], respectively.

Proposition 4.7. *The feasible region defined using constraints (A4)–(A8) is tighter than the set by imposing the constraints in the formulations of CG06, GA10, and TAT19.*

Proof. In addition to the binary variables associated with the presence/absence condensers and reboilers, CG06 has variables for the presence of heat exchanger, which we denote as $\eta_{i,j}$. To our model, we add

$$\eta_{i,j} = \chi_{i,j} + \rho_{i,j}. \quad (4.13)$$

Further, we remark that for $i \leq k \leq j-1$, $[i, j] \in \mathcal{P}$,

$$\sum_{m=i}^k \tau_{i,m,j} \stackrel{(A5)}{=} \sum_{m=i}^k \sum_{l=i+1}^{m+1} \sigma_{i,m,l,j} \stackrel{\text{Fig. 4.2}}{=} \sum_{l=i+1}^{k+1} \sum_{m=l-1}^k \sigma_{i,m,l,j} \stackrel{k \leq j-1}{\leq} \sum_{l=i+1}^{k+1} \sum_{m=l-1}^{j-1} \sigma_{i,m,l,j} \stackrel{(A5)}{=} \sum_{l=i+1}^{k+1} \beta_{i,l,j}. \quad (4.14)$$

In Tables 4.3, 4.2 and 4.4, we prove that the set defined by (A2)–(A8), $\zeta_{i,j} \in [0, 1]$, $\forall [i, j] \in \mathcal{T}$, $\rho_{i,j} \in [0, 1]$, $\forall (i, j) \in \mathcal{R}$ and $\chi_{i,j} \in [0, 1]$, $\forall (i, j) \in \mathcal{C}$ is tighter than CG06, GA10 and TAT19, respectively. We point out that, in GA06, the authors did not consider thermally coupled configurations. Thus, we show the proof only for the constraints they reported.

The fact that our formulation is *strictly tighter* will follow the numerical example below. Consider $N = 4$:

1. When restricted to $\zeta_{1,2} = \zeta_{1,3} = 0$, $\zeta_{1,1} = \zeta_{1,4} = \zeta_{2,2} = \zeta_{3,3} = \zeta_{4,4} = 1$ and $\zeta_{2,3} = \zeta_{2,4} = \zeta_{3,4} = 1/2$, CG06 is feasible, while (A) is infeasible.

2. The point $\tau_{1,1,3} = \tau_{1,2,3} = \tau_{1,1,4} = \tau_{1,3,4} = \tau_{2,2,4} = \beta_{1,2,3} = \beta_{1,3,3} = \beta_{1,3,4} = \beta_{2,4,4} = 0$; $\tau_{1,1,2} = \tau_{1,2,4} = \beta_{1,2,2} = 1$ and $\tau_{2,2,3} = \tau_{2,3,4} = \tau_{3,3,4} = \beta_{1,2,4} = \beta_{1,4,4} = \beta_{2,3,3} = \beta_{2,3,4} = \beta_{3,4,4} = 1/2$ is an extreme point to GA10, and infeasible to (A).
3. When restricted to $\zeta_{3,4} = 0$, $\zeta_{1,1} = \zeta_{1,2} = \zeta_{1,4} = \zeta_{2,2} = \zeta_{3,3} = \zeta_{4,4} = 1$ and $\zeta_{1,3} = \zeta_{2,3} = \zeta_{2,4} = 1/2$, TAT19 is feasible, while (A) is infeasible.

□

Table 4.2.: CG06 for the space of admissible configurations. The first column indicates the constraint number in Table 1 of [9]. ‘Co.’, ‘Re.’ and ‘Pr.’ stand for Corollary, Remark and Proposition, respectively.

#	Proof
[1]	$\sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} \stackrel{(A5), \text{Co.4.1}}{=} \zeta_{i,j} \leq 1$
[2]	$\sum_{n=j+1}^N \sum_{m=i+1}^{j+1} \sigma_{i,j,m,n} \stackrel{(A5)}{=} \sum_{n=j+1}^N \tau_{i,j,n} \stackrel{\text{Pr.4.5}}{\leq} \zeta_{i,j} \leq 1$ $\sum_{m=1}^{i-1} \sum_{n=i-1}^{j-1} \sigma_{m,n,i,j} \stackrel{(A5)}{=} \sum_{m=1}^{i-1} \beta_{m,i,j} \stackrel{\text{Pr.4.5}}{\leq} \zeta_{i,j} \leq 1$
[3]	$\sum_{m=1}^{i-1} \sigma_{m,i-1,i,i} + \sum_{n=i+1}^N \sigma_{i,i,i+1,n} \stackrel{(A5)}{=} \sum_{m=1}^{i-1} \beta_{m,i,i} + \sum_{n=i+1}^N \tau_{i,i,n} \stackrel{(A8)}{\geq} \zeta_{i,i} \stackrel{(A4)}{\geq} 1,$ for $[i, i] \in \mathcal{T}$
[6]	$\left. \begin{aligned} \sigma_{i,j,m,n} &\stackrel{(A5)}{\leq} \tau_{i,j,n} \stackrel{\text{Re.4.2}}{\leq} \zeta_{i,j} \stackrel{\text{Co.4.1}, (A5)}{=} \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} \end{aligned} \right\} \text{ for } \llbracket m \rrbracket_{i+1}^{j+1}; \llbracket n \rrbracket_{j+1}^N$ $\left. \begin{aligned} \sigma_{m,n,i,j} &\stackrel{(A5)}{\leq} \beta_{m,i,j} \stackrel{\text{Re.4.2}}{\leq} \zeta_{i,j} \stackrel{\text{Co.4.1}, (A5)}{=} \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} \end{aligned} \right\} \text{ for } \llbracket n \rrbracket_{i-1}^{j-1}; \llbracket m \rrbracket_1^{i-1}$

continued on next page

Table 4.2.: *continued*

[6]	$\left. \begin{aligned} \sigma_{i,k,l,j} &\stackrel{\text{Pr.4.3}}{\leq} \zeta_{i,j} \stackrel{(\text{A8})}{\leq} \sum_{n=i+1}^N \tau_{i,j,n} + \sum_{m=1}^{i-1} \beta_{m,i,j} \\ &\stackrel{(\text{A5})}{=} \sum_{n=j+1}^N \sum_{m=i+1}^{j+1} \sigma_{i,j,m,n} + \sum_{m=1}^{i-1} \sum_{n=i-1}^{j-1} \sigma_{m,n,i,j} \end{aligned} \right\} \text{ for } \llbracket l \rrbracket_{i+1}^{k+1}; \quad \llbracket k \rrbracket_i^{j-1}$
[12]	$\sigma_{i,k,l,j} \stackrel{\text{Pr.4.3}}{\leq} \zeta_{i,j}, \quad \text{for } \llbracket l \rrbracket_{i+1}^{k+1}; \quad \llbracket k \rrbracket_i^{j-1}$
[13]	$\sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} \stackrel{(\text{A5}), \text{Co.4.1}}{\geq} \zeta_{i,j}$
[4]	$\begin{aligned} 1 - \eta_{i,i} &\stackrel{(4.13)}{\leq}_{\rho_{i,i} \geq 0} 1 - \chi_{i,i} \stackrel{(\text{A7})}{\leq} 1 - \beta_{0,i,i} \\ &\stackrel{(\text{A4}), \text{Pr.4.4}}{=} \sum_{m=1}^{i-1} \beta_{m,i,i} \stackrel{(\text{A5})}{=} \sum_{m=1}^{i-1} \sigma_{m,i-1,i,i}, \quad \text{for } (i,i) \in \mathcal{C} \\ 1 - \eta_{i,i} &\stackrel{(4.13)}{\leq}_{\chi_{i,i} \geq 0} 1 - \rho_{i,i} \stackrel{(\text{A7})}{\leq} 1 - \tau_{i,i,N+1} \\ &\stackrel{(\text{A4}), \text{Pr.4.4}}{=} \sum_{n=i+1}^N \tau_{i,i,n} \stackrel{(\text{A5})}{=} \sum_{n=i+1}^N \sigma_{i,i,i+1,N}, \quad \text{for } (i,i) \in \mathcal{R} \end{aligned}$
[5]	$\begin{aligned} \eta_{i,i} &\stackrel{(4.13)}{=} \chi_{i,j} + \rho_{i,j} \stackrel{(\text{A6}), \text{Pr.4.4}}{\leq} \zeta_{i,j} - \sum_{m=1}^{i-1} \beta_{m,i,j} + \zeta_{i,j} - \sum_{n=j+1}^N \tau_{i,j,n} \\ &\stackrel{\beta, \tau \geq 0}{\leq}_{\zeta_{i,j} \leq 1} (1 - \beta_{m,i,j}) + (1 - \tau_{i,j,n'}) \stackrel{\text{Pr.4.3}}{\leq} (1 - \sigma_{m,n,i,j}) + (1 - \sigma_{i,j,m',n'}) \\ &\quad \text{for } \llbracket m' \rrbracket_{i+1}^{j+1}; \quad \llbracket n' \rrbracket_{j+1}^N; \quad \llbracket m \rrbracket_1^{i-1}; \quad \llbracket n \rrbracket_{i-1}^{j-1} \end{aligned}$
[7]	$\eta_{i,j} \stackrel{(4.13)}{=} \chi_{i,j} + \rho_{i,j} \stackrel{(\text{A6}), \text{Pr.4.4}, (\text{A8})}{\leq} \zeta_{i,j} \stackrel{(\text{A8}), (\text{A5})}{\leq} \sum_{n=j+1}^N \sum_{m=i+1}^{j+1} \sigma_{i,j,m,n} + \sum_{m=1}^{i-1} \sum_{n=i-1}^{j-1} \sigma_{m,n,i,j}$

Table 4.3.

GA10 for space of admissible configurations. The first column indicates the section number in [10]. ‘Co.’, ‘Re.’ and ‘Pr.’ stand for Corollary, Remark and Proposition, respectively.

§	Proof
3.1	$\sum_{l=i+1}^j \beta_{i,l,j} = \sum_{k=i}^{j-1} \tau_{i,k,j} \stackrel{\text{Co.4.1}}{=} \zeta_{i,j} \leq 1$
3.1	$\begin{aligned} \sum_{n=j+1}^N \tau_{i,j,n} &\stackrel{\text{Pr.4.5}}{\leq} \zeta_{i,j} \stackrel{\text{Co.4.1}}{=} \sum_{k=i}^{j-1} \tau_{i,k,j} \\ \sum_{m=1}^{i-1} \beta_{m,i,j} &\stackrel{\text{Pr.4.5}}{\leq} \zeta_{i,j} \stackrel{\text{Co.4.1}}{=} \sum_{k=i}^{j-1} \tau_{i,k,j} \end{aligned}$
3.1	$\sum_{k=i}^{j-1} \tau_{i,k,j} \stackrel{\text{Co.4.1}}{=} \zeta_{i,j} \stackrel{(\text{A8})}{\leq} \sum_{n=j+1}^N \tau_{i,j,n} + \sum_{m=1}^{i-1} \beta_{m,i,j}$
3.2	$\begin{aligned} &\max \left\{ (j-i+1) \sum_{m=1}^{i-1} \beta_{m,i,j}, (j-i+1) \sum_{n=j+1}^N \tau_{i,j,n} \right\} \stackrel{\text{Pr.4.5}}{\leq} (j-i+1) \zeta_{i,j} \stackrel{\text{Co.4.1}}{=} \\ &(j-i+1) \sum_{k=i}^{j-1} \tau_{i,k,j} = \sum_{k=i}^{j-1} (j-k) \tau_{i,k,j} + \sum_{k=i}^{j-1} (k-i+1) \tau_{i,k,j} = \sum_{k=i}^{j-1} \sum_{m=i}^k \tau_{i,m,j} + \\ &\sum_{k=i}^{j-1} (k-i+1) \tau_{i,k,j} \stackrel{(4.14)}{\leq} \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} \beta_{i,l,j} + \sum_{k=i}^{j-1} (k-i+1) \tau_{i,k,j} = \sum_{l=i+1}^j \sum_{k=l-1}^{j-1} \beta_{i,l,j} + \\ &\sum_{k=i}^{j-1} (k-i+1) \tau_{i,k,j} = \sum_{l=i+1}^j (j-l+1) \beta_{i,l,j} + \sum_{k=i}^{j-1} (k-i+1) \tau_{i,k,j} \end{aligned}$
3.3	$\begin{aligned} \sum_{n=j+1}^N \tau_{i,j,n} &\stackrel{\text{Pr.4.5}}{\leq} \zeta_{i,j} \leq 1 \\ \sum_{m=1}^{i-1} \beta_{m,i,j} &\stackrel{\text{Pr.4.5}}{\leq} \zeta_{i,j} \leq 1 \end{aligned}$
3.3	$\sum_{n=i+1}^N \tau_{i,i,n} + \sum_{m=1}^{i-1} \beta_{m,i,i} \stackrel{(\text{A8})}{\geq} \zeta_{i,i} \stackrel{(\text{A4})}{=} 1$

Table 4.4.

TAT19 for the space of admissible configurations. The first column indicates the constraint number in [23]. ‘Re.’ and ‘Pr.’ stand for Remark and Proposition, respectively.

#	Proof
(H2)	$\zeta_{1,N} \stackrel{(A4)}{=} \zeta_{i,i} \stackrel{(A4)}{=} 1$
(H3)	$\zeta_{i,j} \stackrel{(A8)}{\leq} \sum_{n=j+1}^N \tau_{i,j,n} + \sum_{m=1}^{i-1} \beta_{m,i,j} \stackrel{\text{Pr.4.3}}{\leq} \sum_{n=j+1}^N \zeta_{i,n} + \sum_{m=1}^{i-1} \zeta_{m,j}$
(H4)	$\left. \zeta_{i,k} - \sum_{n=k+1}^{j-1} \zeta_{i,n} + \zeta_{i,j} - 1 \stackrel{\text{Re.4.2}}{\leq} \tau_{i,k,j} \stackrel{\text{Pr.4.3}}{\leq} \sum_{l=i+1}^{k+1} \beta_{i,l,j} \stackrel{\text{Re.4.2}}{\leq} \sum_{l=i+1}^{k+1} \zeta_{l,j} \right\} \text{ for } \llbracket k \rrbracket_i^{j-1}$
(H5)	$\left. \zeta_{i,j} - \sum_{m=i+1}^{l-1} \zeta_{m,j} + \zeta_{l,j} - 1 \stackrel{\text{Re.4.2}}{\leq} \beta_{i,l,j} \stackrel{\text{Pr.4.3}}{\leq} \sum_{k=l-1}^{j-1} \tau_{i,k,j} \stackrel{\text{Re.4.2}}{\leq} \sum_{k=l-1}^{j-1} \zeta_{i,k} \right\} \text{ for } \llbracket l \rrbracket_{i+1}^j$
(H6)	$\chi_{i,j} + \rho_{i,j} \stackrel{(A6), \text{Pr. 4.4}}{\leq} \zeta_{i,j} - \sum_{m=1}^{i-1} \beta_{m,i,j} + \zeta_{i,j} - \sum_{n=j+1}^N \tau_{i,j,n} \stackrel{(A8)}{\leq} \zeta_{i,j}$
(H7)	$\begin{aligned} \chi_{i,j} &\stackrel{(A6), \text{Pr. 4.4}}{\leq} \zeta_{i,j} - \sum_{m=1}^{i-1} \beta_{m,i,j} \stackrel{(A8)}{\leq} \sum_{n=j+1}^N \tau_{i,j,n} \stackrel{\text{Re.4.2}}{\leq} \sum_{n=j+1}^N \zeta_{i,n} \\ \rho_{i,j} &\stackrel{(A6), \text{Pr. 4.4}}{\leq} \zeta_{i,j} - \sum_{n=j+1}^N \tau_{i,j,n} \stackrel{(A8)}{\leq} \sum_{m=1}^{i-1} \beta_{m,i,j} \stackrel{\text{Re.4.2}}{\leq} \sum_{m=1}^{i-1} \zeta_{m,j} \end{aligned}$
(H8)	$\chi_{i,j} + \rho_{i,j} \stackrel{(A6)}{\leq} \nu_{i,j+1,N} - \nu_{i,j,N} + \omega_{1,i-1,j} - \omega_{1,i,j} \stackrel{(A3)}{\leq} (1 - \zeta_{i,n}) + (1 - \zeta_{m,j}),$ for $\llbracket m \rrbracket_1^{i-1}; \llbracket n \rrbracket_{j+1}^N$
(H9)	$\chi_{i,i} \stackrel{(A7), \text{Pr. 4.4}}{\geq} \zeta_{i,i} - \sum_{m=1}^{i-1} \beta_{m,i,i} \stackrel{\text{Re.4.2}, (A4)}{\geq} 1 - \sum_{m=1}^{i-1} \zeta_{m,i} \Big\} \text{ for } (i, i) \in \mathcal{C}$
(H10)	$\rho_{i,i} \stackrel{(A7), \text{Pr. 4.4}}{\geq} \zeta_{i,i} - \sum_{n=i+1}^N \tau_{i,i,n} \stackrel{\text{Re.4.2}, (A4)}{\geq} 1 - \sum_{n=i+1}^N \zeta_{i,n} \Big\} \text{ for } (i, i) \in \mathcal{R}$

4.2.3 Mass Balance Constraints

We model the problem as a network flow problem. Figure 4.5 shows the representative nodes and arcs in the network, and variable definitions are in Table 4.5.

Each split $[i, k]/[l, j]$ is performed in a distillation column Q_{iklj} (see Figures 4.5(a) and 4.5(b)). Material flows to and from the column Q_{iklj} only when $\sigma_{iklj} = 1$. The material balances across each column Q_{iklj} are as follows

$$\text{for } [i, j] \in \mathcal{S}, \llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1} :$$

$$\left. \begin{aligned} f_{iklj}^{\text{in}} &= f_{iklj}^{\text{rs}} \delta_{p \leq k} + f_{iklj}^{\text{ss}} \delta_{p \geq l}, \llbracket p \rrbracket_i^j; & U_{iklj}^{\text{rs}} \delta_{j < N} - U_{iklj}^{\text{ss}} \delta_{1 < i} &= V_{iklj}^{\text{rs}} - V_{iklj}^{\text{ss}} \\ K_{iklj}^{\text{ss}} \delta_{1 < i} - K_{iklj}^{\text{rs}} \delta_{j < N} &= L_{iklj}^{\text{ss}} - L_{iklj}^{\text{rs}} \\ 0 \leq (\cdot) &\leq \sigma_{iklj} (\cdot)^{\text{up}}, \forall (\cdot) \in \{\text{All component, liquid and vapor flows}\} \end{aligned} \right\}, \quad (\text{A9})$$

$$\text{for } [i, j] \in \{[1, N]\}, \llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1} :$$

$$\left. \begin{aligned} F_p \sigma_{iklj} &= f_{iklj}^{\text{rs}} \delta_{p \leq k} + f_{iklj}^{\text{ss}} \delta_{p \geq l}, \llbracket p \rrbracket_i^j; & \left(\sum_{p=1}^N F_p \right) (1 - \Phi_{1,N}) \sigma_{iklj} &= V_{iklj}^{\text{rs}} - V_{iklj}^{\text{ss}} \\ \left(\sum_{p=1}^N F_p \right) \Phi_{1,N} \sigma_{iklj} &= L_{iklj}^{\text{ss}} - L_{iklj}^{\text{rs}} \\ 0 \leq (\cdot) &\leq \sigma_{iklj} (\cdot)^{\text{up}}, \forall (\cdot) \in \{\text{All component, liquid and vapor flows}\} \end{aligned} \right\}, \quad (\text{A10})$$

$$\text{for } [i, j] \in \mathcal{P}, \llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1} :$$

$$V_{iklj}^{\text{rs}} - L_{iklj}^{\text{rs}} = \sum_{p=i}^k f_{iklj}^{\text{rs}}; \quad L_{iklj}^{\text{ss}} - V_{iklj}^{\text{ss}} = \sum_{p=l}^j f_{iklj}^{\text{ss}}. \quad (\text{A11})$$

The constraints in (A9) model component, vapor, and liquid mass balances across column Q_{iklj} . In the above $\delta_{(\cdot)}$ is 1 if $(\cdot)$ is true and 0 otherwise. (A10) handles the case where the feed stream is the process feed, $[1, N]$. F_p and $\Phi_{1,N}$ are as defined in §4.1. The last constraint in both (A9) and (A10) suppresses material flows to column Q_{iklj} when $\sigma_{iklj} = 0$. We use $(\cdot)^{\text{up}}$ to denote the upper bound on $(\cdot)$, and discuss how these are obtained later. The first (*resp.* second) constraint in (A11) models that the net distillate (*resp.* residue) flow Q_{iklj} as the difference between the vapor and liquid (*resp.* liquid and vapor) flows in the rectifying (*resp.* stripping) section.

Column Q_{iklj} receives feed from the associated condenser (i, j) and/or reboiler (i, j) (see Figures 4.5(c) and 4.5(d)). Further, condenser (*resp.* reboiler) (i, j) regulates vapor-liquid traffic from all the splits producing $[i, j]$ as distillate (*resp.* residue),

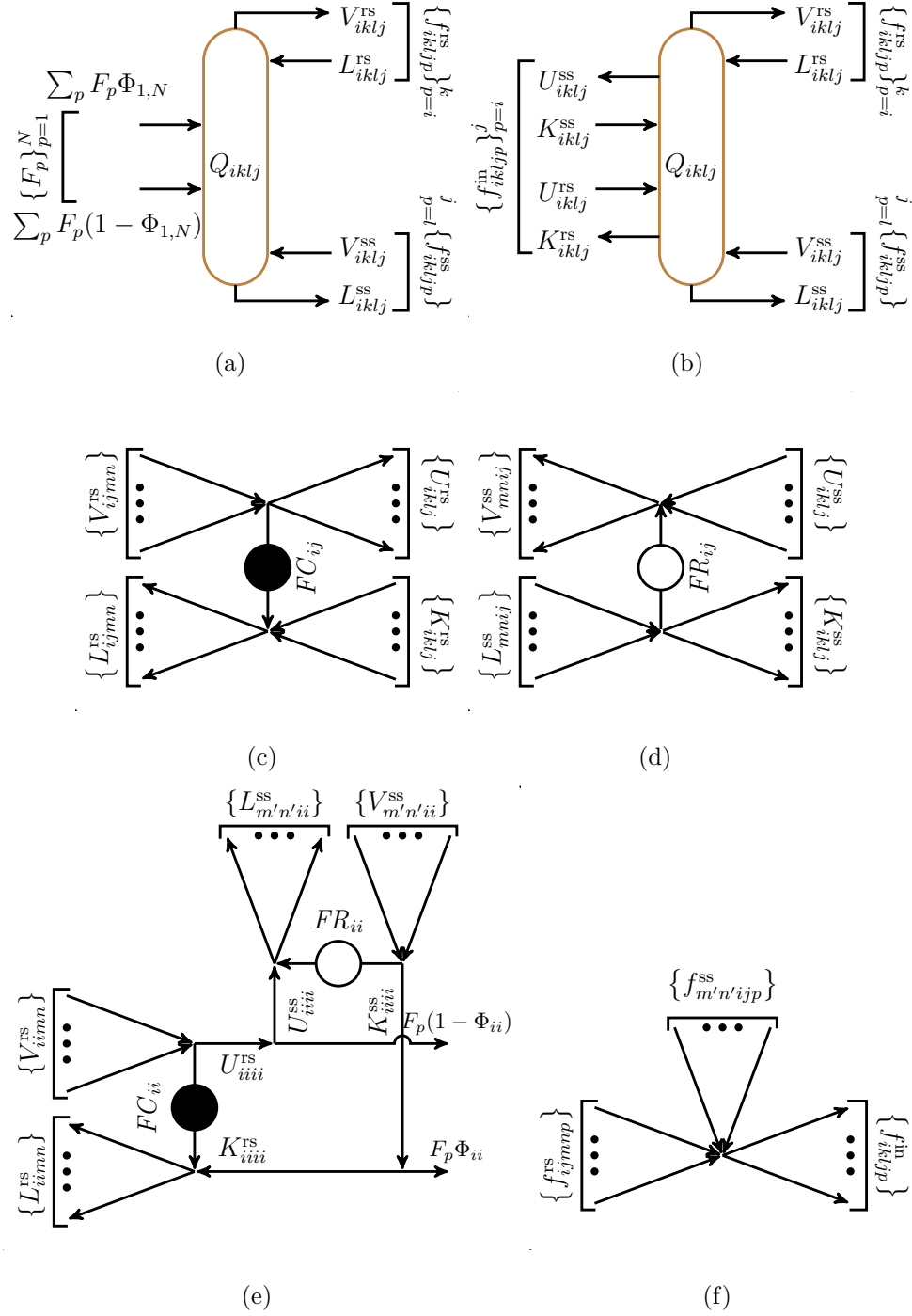


Fig. 4.5. (a) Representative column for splits of process feed i.e., $[i, j] \in \{[1, N]\}$, $\llbracket k \rrbracket_i^{j-1}$, $\llbracket l \rrbracket_{i+1}^{k+1}$ (b) Representative column for the remaining splits $[i, j] \in \mathcal{S}$, $\llbracket k \rrbracket_i^{j-1}$, $\llbracket l \rrbracket_{i+1}^{k+1}$ (c) Representative condenser for $(i, j) \in \mathcal{C} \setminus \{[i, i]\}_{i=1}^{N-1}$ (see (A12) for domain of indices m, n, k, l) (d) Representative reboiler for $(i, j) \in \mathcal{R} \setminus \{[i, i]\}_{i=2}^N$ (see (A13) for domain of indices m, n, k, l) (e) Representative arrangement for pure product withdrawals (see (A14) for domain of indices m, n , and (A15) for domain of indices m', n') (f) Representative arrangement for overall component mass balance for $[i, j] \in \mathcal{S}$ (see (A17) for domain of indices m, n, m', n', k, l)

Table 4.5.
Definition of continuous decision variables.

Variable	Definition
$\{f_{iklj}^{rs}\}_{p=i}^k$	Net molar flow of component p in the rectifying section of Q_{iklj}
$\{f_{iklj}^{ss}\}_{p=l}^j$	Net molar flow of component p in the stripping section of Q_{iklj}
$\{f_{iklj}^{in}\}_{p=i}^j$	Net molar flow of component p in the feed to Q_{iklj}
V_{iklj}^{rs}	Vapor flowrate in the rectifying section of Q_{iklj}
V_{iklj}^{ss}	Vapor flowrate in the stripping section of Q_{iklj}
L_{iklj}^{rs}	Liquid flowrate in the rectifying section of Q_{iklj}
L_{iklj}^{ss}	Liquid flowrate in the stripping section of Q_{iklj}
U_{iklj}^{rs}	Vapor in-flow into Q_{iklj} from condenser (i, j)
U_{iklj}^{ss}	Vapor out-flow from Q_{iklj} to reboiler (i, j)
K_{iklj}^{rs}	Liquid out-flow from Q_{iklj} to condenser (i, j)
K_{iklj}^{ss}	Liquid in-flow into Q_{iklj} from reboiler (i, j)
$\{\theta_{ijq}\}_{q=i}^{j-1}$	Underwood root of Q_{iklj} satisfying $\alpha_{q+1} \leq \theta_{i,j,q} \leq \alpha_q$
Υ_{iklj}^{rs}	Minimum vapor flow required in the rectifying section of Q_{iklj}
Υ_{iklj}^{ss}	Minimum vapor flow required in the stripping section of Q_{iklj}
FC_{ij}	Molar flowrate in condenser (i, j)
FR_{ij}	Molar flowrate in reboiler (i, j)

and distributes flows to all the splits of $[i, j]$. Material balances across these condensers and reboilers are given below:

$$\left. \begin{aligned}
 &\sum_{n=j+1}^N \sum_{m=i+1}^{j+1} V_{ijmn}^{rs} = FC_{ij} + \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} U_{iklj}^{rs}; \\
 &\sum_{n=j+1}^N \sum_{m=i+1}^{j+1} L_{ijmn}^{rs} = FC_{ij} + \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} K_{iklj}^{rs} \\
 &0 \leq FC_{ij} \leq (FC_{ij})^{\text{up}} \chi_{ij}; \\
 &0 \leq K_{iklj}^{rs} \leq (K_{iklj}^{rs})^{\text{up}} (1 - \chi_{ij}), \llbracket l \rrbracket_{i+1}^{k+1}, \llbracket k \rrbracket_i^{j-1}
 \end{aligned} \right\} \quad (\text{A12})$$

$$\left. \begin{aligned}
& \sum_{m=1}^{i-1} \sum_{n=i-1}^{j-1} V_{mni}^{\text{ss}} = FR_{ij} + \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} U_{iklj}^{\text{ss}}; \\
& \sum_{m=1}^{i-1} \sum_{n=i-1}^{j-1} L_{mni}^{\text{ss}} = FR_{ij} + \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} K_{iklj}^{\text{ss}} \\
& 0 \leq FR_{ij} \leq (FR_{ij})^{\text{up}} \rho_{ij}; \\
& 0 \leq U_{iklj}^{\text{ss}} \leq (U_{iklj}^{\text{ss}})^{\text{up}} (1 - \rho_{ij}), \llbracket l \rrbracket_{i+1}^{k+1}, \llbracket k \rrbracket_i^{j-1}.
\end{aligned} \right\} \quad (\text{A13})$$

We are interested in configurations that either have heat exchangers or thermal couplings, but not both. The last two constraints in (A12) and (A13) suppress flows in appropriate arcs if the heat exchangers are absent. The above constraints are written only for heat exchangers associated with mixtures. For heat exchangers associated with pure component products, the vapor and liquid flows are further constrained to produce $\Phi_{i,i} F_i$ and $(1 - \Phi_{i,i}) F_i$ of component i in liquid and vapor phases, respectively (see Figure 4.5(e)). Mass balances around these heat exchangers are given below.

$$\left. \begin{aligned}
& \sum_{n=i+1}^N \sum_{m=i+1}^{i+1} V_{iimn}^{\text{rs}} = FC_{ii} + U_{iii}^{\text{rs}}; \\
& \sum_{n=i+1}^N \sum_{m=i+1}^{i+1} L_{iimn}^{\text{rs}} = FC_{ii} + K_{iii}^{\text{rs}} \\
& 0 \leq U_{iii}^{\text{rs}}; \quad 0 \leq FC_{ii} \leq (FC_{ii})^{\text{up}} \chi_{ii}; \\
& (K_{iii}^{\text{rs}})^{\text{lo}} \leq K_{iii}^{\text{rs}} \leq (K_{iii}^{\text{rs}})^{\text{up}} (1 - \chi_{ii})
\end{aligned} \right\}, \quad (\text{A14})$$

$$\left. \begin{aligned}
& \sum_{m=1}^{i-1} \sum_{n=i-1}^{i-1} V_{mni}^{\text{ss}} = FR_{ii} + U_{iii}^{\text{ss}}; \\
& \sum_{n=i-1}^{i-1} L_{mni}^{\text{ss}} = FR_{ii} + K_{iii}^{\text{ss}} \\
& 0 \leq K_{iii}^{\text{ss}}; \quad 0 \leq FR_{ii} \leq (FR_{ii})^{\text{up}} \rho_{ii}; \\
& (U_{iii}^{\text{ss}})^{\text{lo}} \leq U_{iii}^{\text{ss}} \leq (U_{iii}^{\text{ss}})^{\text{up}} (1 - \rho_{ii})
\end{aligned} \right\}, \quad (\text{A15})$$

$$\text{For } (i, i) \in \mathcal{C} \cap \mathcal{R}: \quad U_{iii}^{\text{rs}} - U_{iii}^{\text{ss}} = F_p (1 - \Phi_{i,i}); \quad K_{iii}^{\text{ss}} - K_{iii}^{\text{rs}} = F_p \Phi_{i,i}. \quad (\text{A16})$$

where $(\cdot)^{\text{lo}}$ denotes the lower bound on $(\cdot)$. From (A16) and (A15) (*resp.* (A14)), $(K_{iii}^{\text{rs}})^{\text{lo}} = -F_p \Phi_{i,i}$ (*resp.* $(U_{iii}^{\text{ss}})^{\text{lo}} = -F_p(1 - \Phi_{i,i})$). For each submixture $[i, j] \in \mathcal{P}$, the net inflow of component p equals the sum of component flows from all the splits that produce $[i, j]$ as distillate or residue. The net inflow is distributed among all splits of $[i, j]$ (see Figure 4.5(f)).

$$\text{For } [i, j] \in \mathcal{S} : \sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} f_{iklj}^{\text{in}} = \sum_{n=j+1}^N \sum_{m=i+1}^{j+1} f_{ijmnp}^{\text{rs}} + \sum_{m=1}^{i-1} \sum_{n=i-1}^{j-1} f_{mnijp}^{\text{ss}}, \quad \llbracket p \rrbracket_i^j. \quad (\text{A17})$$

Finally, modeling the problem in the above manner requires rigorous bounds on all material flows. The net component inflow to and outflow from any column cannot exceed in steady-state the component flow in the process feed. Therefore, the upper bound on all flows of component p is chosen to be F_p i.e., $(f_{iklj}^{\text{in}})^{\text{up}} = (f_{iklj}^{\text{rs}})^{\text{up}} = (f_{iklj}^{\text{ss}})^{\text{up}} = F_p$. However, although required for deriving rigorous relaxations, there is no simple upper bound on vapor and liquid flows in the columns and heat exchangers. For deriving a bound, we use optimality-based bound tightening, where we find feasible flows for an admissible configuration using the technique of [35]. This technique can also be replaced with a local nonlinear programming solver. Let this upper bound be VD^* . Then, we solve the following linear programs (LP) to derive bounds:

$$\max \quad V_{iklj}^{\text{rs}}, \quad \text{s.t.} \quad (\text{A4}) - (\text{A17}), \quad \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \leq \phi VD^* \quad (4.15)$$

We choose $\phi = 1$, if only the optimal solution is desired. Since the model does not capture all operability concerns, such as controllability and suitability w.r.t heat integration with the rest of the plant, and vapor flow predictions are based on shortcut methods rather than rigorous simulations, industrial practitioners are often interested in identifying a ranklist of a few best solutions for this MINLP. Such a ranklist allows them to a posteriori incorporate such considerations. Therefore, to allow construction of such a ranklist, we choose $\phi = 1.5$. With this choice, any configuration that consumes at most 50% more energy than the feasible solution remains in the search space. Our numerical experiments show that each LP can be solved in a fraction of a second using solvers such as Gurobi [49], and the computational time taken to solve all the LPs for a five-component mixture is typically negligible.

4.2.4 Underwood Constraints

As mentioned in §2, for a given split, there is a minimum threshold vapor requirement in each section of a column, below which the products are not produced with the desired purity. A column can, however, carry more vapor than the threshold, and the excess vapor can, if transferred to other columns, be utilized in those columns. The threshold vapor requirement can be computed using Underwood constraints included below:

For $[i, j] \in \mathcal{P}$, $\llbracket k \rrbracket_i^{j-1}$, $\llbracket l \rrbracket_{i+1}^{k+1}$:

$$\sum_{p=i}^j \frac{\alpha_p f_{iklj}^{\text{in}}}{\alpha_p - \theta_{ijq}} = U_{iklj}^{\text{rs}} \delta_{j < N} - U_{iklj}^{\text{ss}} \delta_{1 < i}, \quad \llbracket q \rrbracket_{l-1}^k, \quad (\text{A18})$$

$$\sum_{p=i}^k \frac{\alpha_p f_{iklj}^{\text{rs}}}{\alpha_p - \theta_{ijq}} \leq \Upsilon_{iklj}^{\text{rs}}, \quad \sum_{p=l}^j \frac{\alpha_p f_{iklj}^{\text{ss}}}{\alpha_p - \theta_{ijq}} \geq -\Upsilon_{iklj}^{\text{ss}}, \quad \llbracket q \rrbracket_{l-1}^k, \quad (\text{A19})$$

$$\sum_{p=i}^k \frac{\alpha_p f_{iklj}^{\text{rs}}}{\alpha_p - \theta_{ijq}} \geq \Upsilon_{iklj}^{\text{rs}}, \quad \sum_{p=l}^j \frac{\alpha_p f_{iklj}^{\text{ss}}}{\alpha_p - \theta_{ijq}} \leq -\Upsilon_{iklj}^{\text{ss}}, \quad \llbracket q \rrbracket_l^{k-1}, \quad (\text{A20})$$

$$\alpha_{q+1} \leq \theta_{ijq} \leq \alpha_q, \quad (\text{A21})$$

$$U_{iklj < N}^{\text{rs}} - U_{1 < iklj}^{\text{ss}} = \Upsilon_{iklj}^{\text{rs}} - \Upsilon_{iklj}^{\text{ss}}, \quad (\text{A22})$$

$$0 \leq \Upsilon_{iklj}^{\text{rs}} \leq V_{iklj}^{\text{rs}}, \quad 0 \leq \Upsilon_{iklj}^{\text{ss}} \leq V_{iklj}^{\text{ss}}, \quad (\text{A23})$$

where $\Upsilon_{iklj}^{\text{rs}}$ and $\Upsilon_{iklj}^{\text{ss}}$ denote the threshold vapor flow in rectifying and stripping sections, respectively. Note that, for the process feed $[1, N]$, f_{iklj}^{in} and $U_{iklj}^{\text{rs}} \delta_{j < N} - U_{1 < iklj}^{\text{ss}} \delta_{1 < i}$ in (A18) and (A22) are replaced by $F_p \sigma_{iklj}$ and $\left(\sum_{p=1}^N F_p\right) (1 - \Phi_{1,N}) \sigma_{i,k,l,j}$, respectively. (A18) is commonly known in the literature as the *Underwood feed equation*, and it computes *Underwood roots* $\{\theta_{ijq}\}_{q=l-1}^k$, which satisfy $\alpha_{q+1} \leq \theta_{ijq} \leq \alpha_q$ [31]. (A19) governs the minimum vapor requirement in rectifying and stripping sections as a function of the distillate and residue compositions. (A20) ensures that the minimum vapor constraints are binding for $\{\theta_{ijq}\}_{q=l}^{k-1}$. These constraints are required for the model to have the correct degrees of freedom as described in [23]. (A22) models vapor balance at the feed location in terms of minimum vapor flows. (A23) ensures that the actual vapor in each section is at least as high as the threshold vapor flow.

Remark 4.4. Since the process feed is always present i.e., $\zeta_{1,N} = 1$, and the net component and vapor inflow to columns Q_{1klN} where $1 < l \leq k + 1 \leq N$ are known, we solve the Underwood feed equation (A18) *a priori* to determine the Underwood roots $\{\theta_{1Nq}\}_{q=1}^{N-1}$, and fix these variables to the calculated values. $\square$

Remark 4.5. Recognizing that $f_{iklj p}^{\text{rs}} \geq 0$, $\theta_{ijk} \leq \alpha_k < \alpha_{k-1} < \dots < \alpha_i$, $f_{iklj p}^{\text{ss}} \geq 0$ and $\alpha_j < \alpha_{j-1} < \dots < \alpha_l < \theta_{ijl-1}$, we have

$$\text{for } [i, j] \in \mathcal{P}, \llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1} \quad \left\{ \begin{array}{l} 0 \leq \sum_{p=i}^k \frac{\alpha_p f_{iklj p}^{\text{rs}}}{\alpha_p - \theta_{ijk}}; \quad 0 \leq -\sum_{p=l}^j \frac{\alpha_p f_{iklj p}^{\text{ss}}}{\alpha_p - \theta_{ijl-1}}. \end{array} \right. \quad (4.16)$$

Next, using (4.16), component mass balance $f_{iklj p}^{\text{in}} = f_{iklj p}^{\text{rs}} \delta_{p \leq k} + f_{iklj p}^{\text{ss}} \delta_{p \geq l}$, and (A18), it can be shown that

$$\text{for } [i, j] \in \mathcal{P}, \llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1} \quad \left\{ \begin{array}{l} U_{iklj}^{\text{rs}} \delta_{j < N} - U_{iklj}^{\text{ss}} \delta_{1 < i} \leq \sum_{p=i}^k \frac{\alpha_p f_{iklj p}^{\text{rs}}}{\alpha_p - \theta_{ijl-1}} \\ U_{iklj}^{\text{ss}} \delta_{1 < i} - U_{iklj}^{\text{rs}} \delta_{j < N} \leq -\sum_{p=l}^j \frac{\alpha_p f_{iklj p}^{\text{ss}}}{\alpha_p - \theta_{ijk}}. \end{array} \right. \quad (4.17)$$

Since the vapor flows are bounded, we have finite upper and lower bounds on all nonlinear expressions in (A18)–(A20). $\square$

4.2.5 Exploiting Monotonicity of Underwood Equations

These cuts are inspired from [50] and [34]. Although these relations are implicit in the model, they are not implied in the relaxation, when Underwood constraints are relaxed. We refer to [23] for a derivation.

When $[i, j]$ is produced as distillate from one of its top parent $[i, n]$ where $j + 1 \leq n \leq N$ i.e., $\tau_{i,j,n} = 1$, but not produced as residue from any of its bottom parents i.e., $\beta_{0,i,j} = 1$, and the associated condenser (i, j) is absent, then θ_{inq} lower bounds θ_{ijq} for $\llbracket q \rrbracket_i^{j-1}$. Similarly, when $[i, j]$ is produced as residue from one of its bottom parent $[m, j]$ where $1 \leq m \leq i - 1$ i.e., $\beta_{m,i,j} = 1$, but not produced as distillate from any of

its top parents *i.e.*, $\tau_{i,j,N+1} = 1$, and the associated reboiler (i, j) is absent, then θ_{mq} upper bounds θ_{ijq} for $\llbracket q \rrbracket_i^{j-1}$. These constraints are imposed as follows:

$$\text{for } [i, j] \in \mathcal{S} \quad \begin{cases} \theta_{inq} - \theta_{ijq} \leq M_q [\chi_{i,j} + (1 - \tau_{i,j,n}) + (1 - \beta_{0,i,j})], & \llbracket n \rrbracket_{j+1}^N, \llbracket q \rrbracket_i^{j-1} \\ \theta_{ijq} - \theta_{mq} \leq M_q [\rho_{i,j} + (1 - \beta_{m,i,j}) + (1 - \tau_{i,j,N+1})], & \llbracket m \rrbracket_1^{i-1}, \llbracket q \rrbracket_i^{j-1}, \end{cases} \quad (\text{A24})$$

where $M_q = (\alpha_q - \alpha_{q+1})$ corresponds to the upper bound on the difference of Underwood roots (see (A21)). Numerical examples in [23] illustrate that these cuts help branch & bound converge faster. Given that our formulation has been developed in a lifted space, we use τ and β variables to give a tighter representation of the constraint in (A24). Moreover, if the variables $\psi_{1,m,n,j}$ are not eliminated using Proposition 4.6, they can be used to further tighten the above constraints. For example, in the first constraint, $(1 - \tau_{i,j,n}) + (1 - \beta_{0,i,j})$ can be replaced with $(1 - \psi_{i,n,0,j})$. This concludes the formulation of MINLP (A).

4.3 Relaxation and Solution Procedure

Apart from integrality requirements on stream $(\zeta_{i,j})$ and heat exchanger variables $(\rho_{i,j}$ and $\chi_{i,j})$, the remaining source of nonconvexity in the MINLP is the Underwood constraints. In this section, we describe the construction of a convex relaxation of Underwood constraints ((A18)–(A21)), referred to hereafter as the relaxation, defined using convex constraints that admits all feasible solutions. One of the challenges in constructing a valid relaxation is that the denominator of certain fractions in Underwood constraints can approach arbitrarily close to zero (see (A18)–(A21)). Consequently, off-the-shelf global solvers, such as BARON [21], report an error and are not able to solve the problem. The common strategy used in the literature is to add/subtract ϵ_θ (typically $10^{-2} - 10^{-3}$) from the bounds of θ_{ijq} to prevent it from approaching either α_{q+1} or α_q (see (A21)). However, this ad-hoc strategy has been adopted without a rigorous proof. Our numerical experiments suggest that the choice of this ϵ_θ is not straightforward, and varies from one instance to another. In the follow-

ing, we show that a rigorous relaxation for the fraction can be constructed although the denominator may approach close to zero.

In the following, we drop indices $iklj$. This is because, Underwood equations apply to a column, say Q_{iklj} , and these indices are easily gleaned from the column specification or the associated split $[i, k]/[l, j]$. Moreover, for notational convenience, we describe the relaxation using $\mathcal{U} = \{(f, U, \Upsilon, \theta) \mid (4.18); (f_p^{\text{in}}, f_p^{\text{rs}}, f_p^{\text{ss}}) \in [0, F_p]^3, p = 1, 2; 0 \leq (\cdot) \leq (\cdot)^{\text{up}}, \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}\}$, where

$$\frac{\alpha_1 f_1^{\text{in}}}{\alpha_1 - \theta} - \frac{\alpha_2 f_2^{\text{in}}}{\theta - \alpha_2} = U^{\text{rs}} - U^{\text{ss}}, \quad (4.18a)$$

$$E^{\text{rs}} \leq \frac{\alpha_1 f_1^{\text{rs}}}{\alpha_1 - \theta} - \frac{\alpha_2 f_2^{\text{rs}}}{\theta - \alpha_2} \leq \Upsilon^{\text{rs}}, \quad (4.18b)$$

$$E^{\text{ss}} \leq -\frac{\alpha_1 f_1^{\text{ss}}}{\alpha_1 - \theta} + \frac{\alpha_2 f_2^{\text{ss}}}{\theta - \alpha_2} \leq \Upsilon^{\text{ss}}, \quad (4.18c)$$

$$\alpha_2 \leq \theta^{\text{lo}} \leq \theta \leq \theta^{\text{up}} \leq \alpha_1, \quad (4.18d)$$

$$U^{\text{rs}} - U^{\text{ss}} = \Upsilon^{\text{rs}} - \Upsilon^{\text{ss}}, \quad (4.18e)$$

$$f_p^{\text{in}} = f_p^{\text{rs}} + f_p^{\text{ss}}, \quad p = 1, 2. \quad (4.18f)$$

Here, we assume that column Q_{iklj} performs the split of a binary mixture. Observe that (4.18a), the second inequality in (4.18b) and (4.18c) are simplified versions of (A18) and (A19) for binary mixtures. We ensure that all fractions are non-negative by factoring out a negative sign from the fractions whose denominator is negative (see (4.18)). Next, E^{rs} and E^{ss} denote lower bounds on nonlinear expressions in (4.18b) and (4.18c), respectively. We choose E^{rs} (*resp.* E^{ss}) to be Υ^{rs} (*resp.* Υ^{ss}) if the second inequality in (4.18b) (*resp.* (4.18c)) needs to be binding, as in (A20). Else, we choose the lower bound derived in (4.16) and (4.17). (4.18d), (4.18e), and (4.18f) correspond to (A21), (A22), and (A9), respectively. Lastly, we remark that, in (A19) and (A20), $f_2^{\text{rs}} = f_1^{\text{ss}} = 0$ for a split of a binary mixture. Since our purpose in restricting to the binary case is to illustrate the mathematical structure of relaxations, we do not consider this restriction. In general splits, one or more components may distribute between the distillate and residue.

The first step in standard approaches to relax $\mathcal{U}$ is to linearize Underwood constraints by introducing an auxiliary variable representing the graph of each fraction. Then, the restriction that this variable take the value of the fraction is replaced with the less stringent restriction that the variable lies in a convex set containing the graph of fraction. Instead, we reformulate $\mathcal{U}$ as described in §4.3.1 before linearizing the Underwood constraints.

4.3.1 Reformulation

We adapt classical Reformulation-Linearization Technique (RLT) [26] to fractions, and reformulate $\mathcal{U}$ by appending RLT cuts derived using Underwood constraints. For clarity, we present the derivation of RLT cuts with Underwood minimum vapor constraint in the rectifying section (second inequality in (4.18b)), and describe the entire reformulated set towards the end. We multiply each Underwood constraint with the bound factors of θ , $(\theta - \theta^{\text{lo}})$, and $(\theta^{\text{up}} - \theta)$. A naive approach would then disaggregate the product, leading to

$$\frac{\alpha_1 f_1^{\text{rs}} \theta}{\alpha_1 - \theta} - \frac{\alpha_1 f_1^{\text{rs}} \theta^{\text{lo}}}{\alpha_1 - \theta} - \frac{\alpha_2 f_2^{\text{rs}} \theta}{\theta - \alpha_2} + \frac{\alpha_2 f_2^{\text{rs}} \theta^{\text{lo}}}{\theta - \alpha_2} \leq \Upsilon^{\text{rs}} \cdot \theta - \Upsilon^{\text{rs}} \cdot \theta^{\text{lo}}, \quad (4.19a)$$

$$\frac{\alpha_1 f_1^{\text{rs}} \theta^{\text{up}}}{\alpha_1 - \theta} - \frac{\alpha_1 f_1^{\text{rs}} \theta}{\alpha_1 - \theta} - \frac{\alpha_2 f_2^{\text{rs}} \theta^{\text{up}}}{\theta - \alpha_2} + \frac{\alpha_2 f_2^{\text{rs}} \theta}{\theta - \alpha_2} \leq \Upsilon^{\text{rs}} \cdot \theta^{\text{up}} - \Upsilon^{\text{rs}} \cdot \theta, \quad (4.19b)$$

following which auxiliary variables are introduced to linearize each nonlinear term: $H_p^{\text{rs}} = f_p^{\text{rs}}/|\alpha_p - \theta|$, $\underline{H}\theta_p^{\text{rs}} = f_p^{\text{rs}}\theta/|\alpha_p - \theta|$, for $p = 1, 2$, and $\underline{\Upsilon}\theta^{\text{rs}} = \Upsilon^{\text{rs}} \cdot \theta$. Here, and in the rest of the chapter, the variables introduced to linearize a product will be written by underlining the concatenation of symbols, as in $\underline{\Upsilon}\theta^{\text{rs}} = \Upsilon^{\text{rs}} \cdot \theta$. Instead, we use polynomial long division prior to linearization, which transforms (4.19) to

$$\frac{\alpha_1(\alpha_1 - \theta^{\text{lo}})f_1^{\text{rs}}}{\alpha_1 - \theta} - \alpha_1 f_1^{\text{rs}} + \frac{\alpha_2(\theta^{\text{lo}} - \alpha_2)f_2^{\text{rs}}}{\theta - \alpha_2} - \alpha_2 f_2^{\text{rs}} \leq \Upsilon^{\text{rs}} \cdot \theta - \Upsilon^{\text{rs}} \cdot \theta^{\text{lo}}, \quad (4.20a)$$

$$- \frac{\alpha_1(\alpha_1 - \theta^{\text{up}})f_1^{\text{rs}}}{\alpha_1 - \theta} + \alpha_1 f_1^{\text{rs}} - \frac{\alpha_2(\theta^{\text{up}} - \alpha_2)f_2^{\text{rs}}}{\theta - \alpha_2} + \alpha_2 f_2^{\text{rs}} \leq \Upsilon^{\text{rs}} \cdot \theta^{\text{up}} - \Upsilon^{\text{rs}} \cdot \theta. \quad (4.20b)$$

Next, we introduce auxiliary variables to linearize nonlinear terms: $H_p^{\text{rs}} = f_p^{\text{rs}}/|\alpha_p - \theta|$, for $p = 1, 2$, and $\underline{\Upsilon}\theta^{\text{rs}} = \Upsilon^{\text{rs}} \cdot \theta$. We shall refer to the proposed variant as the

Reformulation-Division-Linearization Technique (RDLT) of fractional terms, in order to easily distinguish and emphasize the use of polynomial division as an intermediate step. Clearly, RDLT cuts require fewer variables than those derived by naive application of RLT as described above. In addition, RDLT cuts lead to a tighter relaxation of $\mathcal{U}$, which we demonstrate below.

Proposition 4.8. *Let $\mathcal{B} = [f_1^{\text{lo}}, f_1^{\text{up}}] \times [f_2^{\text{lo}}, f_2^{\text{up}}] \times [\Upsilon^{\text{lo}}, \Upsilon^{\text{up}}] \times [\theta^{\text{lo}}, \theta^{\text{up}}]$, and $S = \{(f, \Upsilon, \theta) \in \mathcal{B} \mid \frac{\alpha_1 f_1}{\alpha_1 - \theta} - \frac{\alpha_2 f_2}{\theta - \alpha_2} \leq \Upsilon\}$. Let $\underline{\Upsilon}\theta$, H_i , $\underline{H}\theta_i$ be linearizations of $\Upsilon \cdot \theta$, $\frac{f_i}{|\alpha_i - \theta|}$, and $\frac{f_i \theta}{|\alpha_i - \theta|}$ respectively. Define $S_{\text{std}} = \{(f, \Upsilon, \theta, H) \in \mathcal{B} \times \mathbb{R}^2 \mid \alpha_1 H_1 - \alpha_2 H_2 \leq \Upsilon, \check{H}_i \leq H_i \leq \hat{H}_i, i = 1, 2\}$, $S_{\text{RLT}} = \{(f, \Upsilon, \theta, H, \underline{H}\theta, \underline{\Upsilon}\theta) \in C \mid (4.21)\}$, where $C \subseteq \mathcal{B} \times \mathbb{R}^5$ and*

$$\alpha_1(\underline{H}\theta_1 - \theta^{\text{lo}} H_1) - \alpha_2(\underline{H}\theta_2 - \theta^{\text{lo}} H_2) \leq \underline{\Upsilon}\theta - \Upsilon \cdot \theta^{\text{lo}}, \quad (4.21a)$$

$$\alpha_1(\theta^{\text{up}} H_1 - \underline{H}\theta_1) - \alpha_2(\theta^{\text{up}} H_2 - \underline{H}\theta_2) \leq \Upsilon \cdot \theta^{\text{up}} - \underline{\Upsilon}\theta. \quad (4.21b)$$

Let $S_{\text{RDLT}} = \{(f, \Upsilon, \theta, H, \underline{\Upsilon}\theta) \in C' \mid (4.22)\}$, where $C' \subseteq \mathcal{B} \times \mathbb{R}^5$ and

$$\alpha_1(\alpha_1 H_1 - f_1 - \theta^{\text{lo}} H_1) - \alpha_2(\alpha_2 H_2 + f_2 - \theta^{\text{lo}} H_2) \leq \underline{\Upsilon}\theta - \Upsilon \cdot \theta^{\text{lo}}, \quad (4.22a)$$

$$\alpha_1(\theta^{\text{up}} H_1 - \alpha_1 H_1 + f_1) - \alpha_2(\theta^{\text{up}} H_2 - f_2 - \alpha H_2) \leq \Upsilon \cdot \theta^{\text{up}} - \underline{\Upsilon}\theta. \quad (4.22b)$$

Assume that $C \supseteq \{(f, \Upsilon, \theta, H, \underline{H}\theta, \underline{\Upsilon}\theta) \mid (f, \Upsilon, \theta, H, \underline{\Upsilon}\theta) \in C', \underline{H}\theta_1 = \alpha_1 H_1 - f_1, \underline{H}\theta_2 = \alpha_2 H_2 + f_2\}$ and $\text{proj}_{H_1, H_2} C \subseteq [\check{H}_1, \hat{H}_1] \times [\check{H}_2, \hat{H}_2]$. Then, $S_{\text{std}} \supseteq \text{proj}_{(f, \Upsilon, \theta, H)}(S_{\text{RLT}})$ and $S_{\text{RLT}} \supseteq \{(f, \Upsilon, \theta, H, \underline{H}\theta, \underline{\Upsilon}\theta) \in S_{\text{RDLT}} \times \mathbb{R}^2 \mid \underline{H}\theta_1 = \alpha_1 H_1 - f_1, \underline{H}\theta_2 = \alpha_2 H_2 + f_2\}$, where the right hand side is an affine lifting of S_{RDLT} .

Proof. The first part of the statement follows easily because $\alpha_1 H_1 - \alpha_2 H_2 \leq \Upsilon$ is obtained by adding (4.21a) and (4.21b), and the bounds on H_i in S_{std} are implied by our assumption $\text{proj}_{H_1, H_2} C \subseteq [\check{H}_1, \hat{H}_1] \times [\check{H}_2, \hat{H}_2]$. The second part follows similarly because (4.21a) is derived by adding (4.22a) with $\alpha_1(\underline{H}\theta_1 - \alpha_1 H_1 + f_1) = 0$ and $\alpha_2(\underline{H}\theta_2 - \alpha_2 H_2 - f_2) = 0$, and affine lifting of any point in C' that satisfies this equation is assumed to be contained in C . $\square$

The sets C and C' in Proposition 4.8 are typically created by relaxing the nonlinear expressions. We illustrate, via an example, that the relations in Proposition 4.8 can be strict.

Example 4.1. Let, $\alpha_1 = 15$, $\alpha_2 = 9$, $f_1^{\text{lo}} = f_1^{\text{up}} = 0.6$, $f_2^{\text{lo}} = f_2^{\text{up}} = 0.4$, $\Upsilon^{\text{lo}} = -10$, $\Upsilon^{\text{up}} = 10$, $\theta^{\text{lo}} = 9.1$, $\theta^{\text{up}} = 14.9$. The sets C and C' are constructed by under- and over-estimating the nonlinear terms with their respective convex and concave envelopes. Figure 4.6(a) depicts the projection of sets S , S_{std} , S_{RLT} and S_{RDLT} in Υ - θ space. It is clear that $\text{proj}_{(f, \Upsilon, \theta)}(S_{\text{std}}) \supset \text{proj}_{(f, \Upsilon, \theta)}(S_{\text{RLT}}) \supset \text{proj}_{(f, \Upsilon, \theta)}(S_{\text{RDLT}}) \supset S$. Besides improving the quality of relaxation by introducing fewer auxiliary variables, RDLT has another benefit in our context that we describe next.

Even when f_1 and f_2 are fixed, the function $\frac{\alpha_1 f_1}{\alpha_1 - \theta} - \frac{\alpha_2 f_2}{\theta - \alpha_2}$ is nonconvex (see Figure 4.6(b)), because it is a difference of two convex functions. When this function is multiplied by $(\theta - \theta^{\text{lo}})$ (*resp.* $(\theta^{\text{up}} - \theta)$), it becomes convex (*resp.* concave) (see 4.6(b)). In the naive RLT approach, where each fraction is relaxed independently, the product $(f_1/(\alpha_1 - \theta)) \cdot (\theta - \theta^{\text{lo}})$ is disaggregated and relaxed as a difference of the convex envelope of $f_1\theta/(\alpha_1 - \theta)$ with the concave envelope of $f_1/(\alpha_1 - \theta)$. Whereas, the polynomial division step makes the convexity apparent revealing better ways to construct the relaxation. $\square$

We use RDLT to obtain a reformulation of $\mathcal{U}$, denoted as $\mathcal{U}_{\text{ref}}$, in higher dimensional space as $\mathcal{U}_{\text{ref}} = \{(f, U, \Upsilon, \theta, H, \underline{U}\theta, \underline{\Upsilon}\theta) \mid (4.23); (f_p, \theta, H_p) \in \mathcal{F}_p, p = 1, 2; (U, \Upsilon, \theta, \underline{U}\theta, \underline{\Upsilon}\theta) \in \mathcal{V}\}$, where

$$\sum_{p=1}^2 (\alpha_p |\alpha_p - \theta^{\text{lo}}| H_p^{\text{in}} - \alpha_p f_p^{\text{in}}) = (\underline{U}\theta^{\text{rs}} - \theta^{\text{lo}} U^{\text{rs}}) - (\underline{U}\theta^{\text{ss}} - \theta^{\text{lo}} U^{\text{ss}}), \quad (4.23a)$$

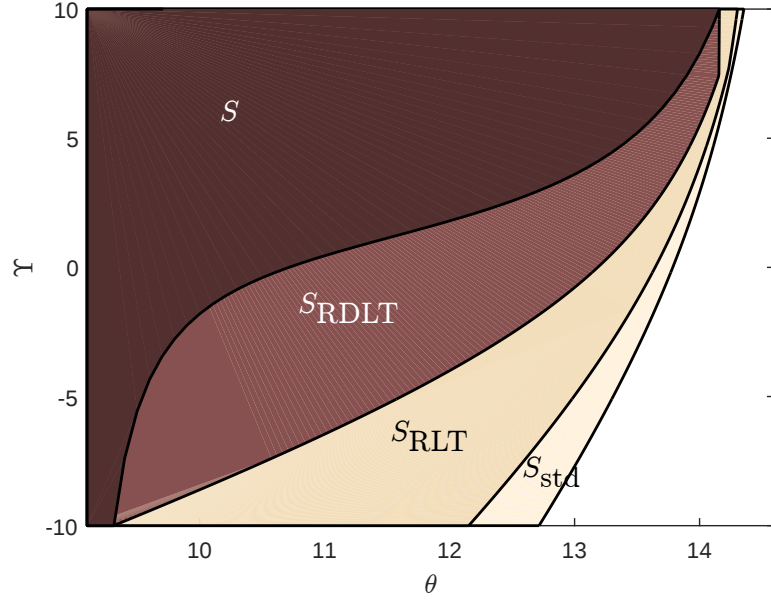
$$\sum_{p=1}^2 (\alpha_p f_p^{\text{in}} - \alpha_p |\alpha_p - \theta^{\text{up}}| H_p^{\text{in}}) = (\theta^{\text{up}} U^{\text{rs}} - \underline{U}\theta^{\text{rs}}) - (\theta^{\text{up}} U^{\text{ss}} - \underline{U}\theta^{\text{ss}}), \quad (4.23b)$$

$$E^{\text{rs}}(\theta - \theta^{\text{lo}}) \leq \sum_{p=1}^2 (\alpha_p |\alpha_p - \theta^{\text{lo}}| H_p^{\text{rs}} - \alpha_p f_p^{\text{rs}}) \leq \underline{\Upsilon}\theta^{\text{rs}} - \theta^{\text{lo}} \Upsilon^{\text{rs}}, \quad (4.23c)$$

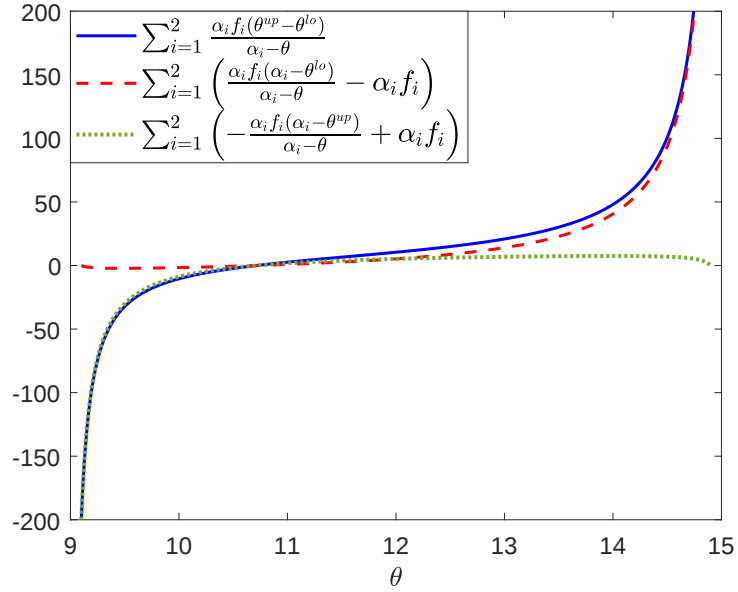
$$E^{\text{rs}}(\theta^{\text{up}} - \theta) \leq \sum_{p=1}^2 (\alpha_p f_p^{\text{rs}} - \alpha_p |\alpha_p - \theta^{\text{up}}| H_p^{\text{rs}}) \leq \theta^{\text{up}} \Upsilon^{\text{rs}} - \underline{\Upsilon}\theta^{\text{rs}}, \quad (4.23d)$$

$$E^{\text{ss}}(\theta - \theta^{\text{lo}}) \leq \sum_{p=1}^2 (\alpha_p f_p^{\text{ss}} - \alpha_p |\alpha_p - \theta^{\text{up}}| H_p^{\text{ss}}) \leq \underline{\Upsilon}\theta^{\text{ss}} - \theta^{\text{lo}} \Upsilon^{\text{ss}}, \quad (4.23e)$$

$$E^{\text{ss}}(\theta^{\text{up}} - \theta) \leq \sum_{p=1}^2 (\alpha_p |\alpha_p - \theta^{\text{lo}}| H_p^{\text{ss}} - \alpha_p f_p^{\text{ss}}) \leq \theta^{\text{up}} \Upsilon^{\text{ss}} - \underline{\Upsilon}\theta^{\text{ss}}. \quad (4.23f)$$



(a)



(b)

Fig. 4.6. (a) Projection of sets S , S_{std} , S_{RLT} and S_{RDLT} in Example 4.1 in $\Upsilon - \theta$ space. (b) Plots of nonlinear expression in Underwood constraint

In the above, $|\cdot|$ denotes absolute value function, and the sets $\mathcal{F}_p$, $p = 1, 2$, and $\mathcal{V}$ are defined as

$$\mathcal{F}_p = \left\{ (f_p, \theta, H_p) \left| \begin{array}{l} H_p^{\text{in}} = f_p^{\text{in}} \cdot T_p(\theta), H_p^{\text{rs}} = f_p^{\text{rs}} \cdot T_p(\theta), H_p^{\text{ss}} = f_p^{\text{ss}} \cdot T_p(\theta) \\ f_p^{\text{in}} = f_p^{\text{rs}} + f_p^{\text{ss}} \\ (f_p^{\text{in}}, f_p^{\text{rs}}, f_p^{\text{ss}}) \in [0, F_p]^3, \theta^{\text{lo}} \leq \theta \leq \theta^{\text{up}} \end{array} \right. \right\}, \quad (4.24)$$

and

$$\mathcal{V} = \left\{ (U, \Upsilon, \theta, \underline{U}\theta, \underline{\Upsilon}\theta) \left| \begin{array}{l} \underline{U}\theta^{\text{rs}} = U^{\text{rs}} \cdot \theta, \underline{U}\theta^{\text{ss}} = U^{\text{ss}} \cdot \theta \\ \underline{\Upsilon}\theta^{\text{rs}} = \Upsilon^{\text{rs}} \cdot \theta, \underline{\Upsilon}\theta^{\text{ss}} = \Upsilon^{\text{ss}} \cdot \theta \\ U^{\text{rs}} - U^{\text{ss}} = \Upsilon^{\text{rs}} - \Upsilon^{\text{ss}} \\ \theta^{\text{lo}} \leq \theta \leq \theta^{\text{up}} \\ 0 \leq U^{\text{rs}} \leq (U^{\text{rs}})^{\text{up}}, 0 \leq U^{\text{ss}} \leq (U^{\text{ss}})^{\text{up}} \\ 0 \leq \Upsilon^{\text{rs}} \leq (\Upsilon^{\text{rs}})^{\text{up}}, 0 \leq \Upsilon^{\text{ss}} \leq (\Upsilon^{\text{ss}})^{\text{up}} \end{array} \right. \right\}, \quad (4.25)$$

where $T_1(\theta) = 1/(\alpha_1 - \theta)$, and $T_2(\theta) = 1/(\theta - \alpha_2)$.

Generalizations

We remark that RDLT can be used for problems with constraints that have the form $\sum_{i=1}^r \frac{x_i g_i(y)}{h_i(y)} \leq x_0$, $\{g_i(y)\}_{i=1}^r$ and $\{h_i(y)\}_{i=1}^r$ are some polynomials of y . We follow the steps below to derive RDLT cuts.

1. We multiply the constraint by some ratio of polynomials of y , $n(y)/d(y)$, such that the sign of the ratio does not change over the domain of y . Here, we assume, w.l.o.g, that $n(y)/d(y) \geq 0$ over the domain of y .
2. We use polynomial long division to express each $\frac{g_i(y) \cdot n(y)}{h_i(y) \cdot d(y)} = m_i(y) + \frac{k_i(y)}{l_i(y)}$ such that $\deg(k_i) < \deg(l_i)$, where $\deg(k_i)$ denotes degree of polynomial $k_i(y)$.
3. We factorize $l_i(y)$ and express it as a product of polynomials $\{q_{ij}(y)\}_{j=1}^{s_i}$ that are non-factorizable over real numbers (e.g., $y + 2$ or $y^2 + y + 1$).

4. We use the general theorem of partial fraction decomposition to express each fraction $k_i(y)/l_i(y)$ as $\sum_{j=1}^{s_i} p_{ij}(y)/q_{ij}(y)$, where $\deg(p_{ij}) < \deg(q_{ij})$. This transforms the constraint to $\sum_{i=1}^r \left(x_i \cdot m_i(y) + \sum_{j=1}^{s_i} x_i \cdot p_{ij}(y)/q_{ij}(y) \right) \leq x_0 \cdot n(y)/d(y)$.
5. We linearize the constraint by introducing auxiliary variables for each nonlinear term.

The reformulation described earlier is a specific case, where we chose to multiply each Underwood constraint by $(\theta - \theta^{\text{lo}})$ and $(\theta^{\text{up}} - \theta)$. By changing the factor used in the reformulation step, we can derive alternative RDLT cuts by following the steps described above. As an illustration, we derive two types of additional RDLT cuts for reformulation of $\mathcal{U}$. While we do not use these cuts for our extensive computational experiments, we demonstrate with numerical examples in §4.4 that they further improve the relaxation for some instances.

RDLT cuts with quadratic polynomials: Here, we choose the product of bound factors of θ , viz. $(\theta - \theta^{\text{lo}})^2$, $(\theta - \theta^{\text{lo}}) \cdot (\theta^{\text{up}} - \theta)$ and $(\theta^{\text{up}} - \theta)^2$, for reformulation. As an illustration, we derive the RDLT cut by multiplying the second inequality in (4.18b) with $(\theta - \theta^{\text{lo}}) \cdot (\theta^{\text{up}} - \theta)$. The remaining RDLT cuts are derived in a similar fashion. Steps 1 and 2 lead to

$$\sum_{p=1}^2 \left(\alpha_p f_p^{\text{rs}} \cdot (\theta + \alpha_p - \theta^{\text{lo}} - \theta^{\text{up}}) - \frac{\alpha_p (\alpha_p - \theta^{\text{lo}}) (\alpha_p - \theta^{\text{up}}) f_p^{\text{rs}}}{\alpha_p - \theta} \right) \leq \Upsilon^{\text{rs}} \cdot (\theta - \theta^{\text{lo}}) \cdot (\theta^{\text{up}} - \theta). \quad (4.26)$$

Since (4.26) is already in the form attained in Step 4, we do not need Steps 3 and 4. Finally, we disaggregate the products of f_p^{rs} and Υ^{rs} with polynomials of θ , and linearize (4.26) by introducing auxiliary variables for $f_p^{\text{rs}}/(\alpha_p - \theta)$, $f_p^{\text{rs}} \cdot \theta$, $\Upsilon^{\text{rs}} \cdot \theta^2$ and $\Upsilon \cdot \theta$.

RDLT cuts with inverse bound factors: Here, we use *inverse bound factors* $\left(\frac{1}{\theta} - \frac{1}{\theta^{\text{up}}}\right)$ and $\left(\frac{1}{\theta^{\text{lo}}} - \frac{1}{\theta}\right)$ for reformulation. Since $\left(\frac{1}{\theta} - \frac{1}{\theta^{\text{up}}}\right) = \frac{\theta^{\text{up}} - \theta}{\theta^{\text{up}} \cdot \theta}$, inverse bound factors are essentially ratios of first-degree polynomial to another first-degree polynomial. As before, for illustration, we derive the RDLT cut obtained by multiplying the

second inequality in (4.18b) with $(\frac{1}{\theta} - \frac{1}{\theta^{\text{up}}})$. The remaining RDLT cuts are obtained in a similar fashion. Step 1 leads to $\sum_{p=1}^2 \frac{\alpha_p f_p^{\text{rs}}}{(\alpha_p - \theta)\theta} - \frac{\alpha_p f_p^{\text{rs}}}{(\alpha_p - \theta)\theta^{\text{up}}} \leq \frac{\Upsilon^{\text{rs}}}{\theta} - \frac{\Upsilon^{\text{rs}}}{\theta^{\text{up}}}$, which is already in the form described in Step 2. Further, the denominator of each fraction is already expressed as product of non-factorizable polynomials. Next, we use partial fraction decomposition (Step 4) to obtain

$$\sum_{p=1}^2 \left(\frac{f_p^{\text{rs}}}{\theta} - \frac{(\alpha_p - \theta^{\text{up}})f_p^{\text{rs}}}{\theta^{\text{up}}(\alpha_p - \theta)} \right) \leq \frac{\Upsilon^{\text{rs}}}{\theta} - \frac{\Upsilon^{\text{rs}}}{\theta^{\text{up}}}. \quad (4.27)$$

Finally, we linearize (4.27) by introducing auxiliary variables for $f_p^{\text{rs}}/|\alpha_p - \theta|$, f_p^{rs}/θ and $\Upsilon^{\text{rs}}/\theta$.

4.3.2 Relaxation for $\alpha_2 < \theta^{\text{lo}}$ and $\theta^{\text{up}} < \alpha_1$

The nonconvexity in $\mathcal{U}_{\text{ref}}$ is due to $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{V}$. We convexify these sets to construct a convex relaxation of $\mathcal{U}_{\text{ref}}$. However, we first assume that $\alpha_2 < \theta^{\text{lo}}$ and $\theta^{\text{up}} < \alpha_1$, and relax this assumption later in §4.3.3. This assumption prevents the denominator of fractions in $\mathcal{F}_1$ and $\mathcal{F}_2$ from becoming zero. This discussion is needed for two reasons: (i) it will guide us in deriving additional valid cuts needed to strengthen the relaxation when $\theta = \alpha_2$ and $\theta = \alpha_1$ are admissible (ii) it is needed to construct a piecewise relaxation in §4.3.4, where we discretize the domain of θ such that every partition excluding the extreme partitions satisfy $\alpha_2 < \theta^{\text{lo}} \leq \theta \leq \theta^{\text{up}} < \alpha_1$.

The standard approach to create a relaxation is to replace each equality $H_p = f_p \cdot T_p(\theta)$ in $\mathcal{F}_p$ (*resp.* $\underline{\Upsilon}\theta = \Upsilon \cdot \theta$ in $\mathcal{V}$) with a less stringent restriction that H_p (*resp.* $\underline{\Upsilon}\theta$) lies in the convex hull of $f_p \cdot T_p(\theta)$ (*resp.* $\Upsilon \cdot \theta$) over a rectangle defined by the ranges of f_p (*resp.* Υ) and θ . However, this approach does not take advantage of the fact that the component (*resp.* vapor) flows are constrained by mass balances (see (4.18e),(4.18f)) and, thus, results in a weaker relaxation. Instead, we use Proposition 4.9, which describes the construction of simultaneous hull of multiple nonlinear terms over a polytope (not necessarily a hyperrectangle), to construct a tighter relaxation of $\mathcal{U}_{\text{ref}}$.

Proposition 4.9. *Let $X = \{x \in \mathbb{R}_+^n \mid Bx \leq b\}$ be a polytope, $g(y)$ be continuous and convex for $y \in [y^{\text{lo}}, y^{\text{up}}] \subset \mathbb{R}$, $\mathcal{D} = X \times [y^{\text{lo}}, y^{\text{up}}] \times \mathbb{R}^{n+n}$, and $S = \{\pi \in \mathcal{D} \mid z_j = x_j \cdot g(y), \underline{xy}_j = x_j \cdot y, \llbracket j \rrbracket_1^n\}$, where $\pi = (x, y, z, \underline{xy})$ denotes an element of S . Then, $\text{Conv}(S) = \text{proj}_\pi\{(\pi, y^1, \dots, y^m, w^1, \dots, w^m, w, \lambda^1, \dots, \lambda^m) \mid (4.28)\}$, where*

$$w^i \geq g^*(\lambda^i, y^i), \quad i = 1, \dots, m \quad (4.28a)$$

$$w^i \leq \lambda^i g(y^{\text{lo}}) + \left(\frac{g(y^{\text{up}}) - g(y^{\text{lo}})}{y^{\text{up}} - y^{\text{lo}}} \right) (y^i - \lambda^i y^{\text{lo}}), \quad i = 1, \dots, m \quad (4.28b)$$

$$\lambda^i y^{\text{lo}} \leq y^i \leq \lambda^i y^{\text{up}}, \quad i = 1, \dots, m \quad (4.28c)$$

$$z = \sum_{i=1}^m v^i w^i, \quad \underline{xy} = \sum_{i=1}^m v^i y^i, \quad w = \sum_{i=1}^m w^i, \quad (4.28d)$$

$$y = \sum_{i=1}^m y^i, \quad x = \sum_{i=1}^m \lambda^i v^i, \quad (\lambda^1, \dots, \lambda^m) \in \Delta^m. \quad (4.28e)$$

Here, $\text{proj}_\pi\{\cdot\}$ represents projection of $\{\cdot\}$ onto the space of $(x, y, z, \underline{xy})$ variables, $\{v^i\}_{i=1}^m$ are the extreme points of X , $\Delta^m = \{(\lambda^1, \dots, \lambda^m) \in \mathbb{R}_+^m \mid \sum_{i=1}^m \lambda^i = 1\}$, and positively homogeneous function $g^*(\lambda^*, y^*)$ related to $g(y) : [y^{\text{lo}}, y^{\text{up}}] \rightarrow \mathbb{R}$ is defined as:

$$g^*(\lambda^*, y^*) = \begin{cases} \lambda^* g((\lambda^*)^{-1} y^*), & \text{if } (\lambda^*)^{-1} y^* \in [y^{\text{lo}}, y^{\text{up}}], \lambda^* > 0 \\ 0, & \text{if } \lambda^* = 0, y^* = 0. \end{cases} \quad (4.29)$$

Proof. Since S is compact, its convex hull is compact and, by Krein-Milman theorem, is the convex hull of its extreme points. Therefore, we determine the extreme points of S , and take their convex hull to obtain $\text{Conv}(S)$. When y is restricted to $\bar{y} \in [y^{\text{lo}}, y^{\text{up}}]$, the set $S = \{(x, y, z, \underline{xy}) \mid z = g(\bar{y}) x, \underline{xy} = \bar{y} x, x \in X, y = \bar{y}\}$ can be expressed as an affine transform of X . Thus, the extreme points of S project to the set of extreme points of X and we may restrict attention to these points in order to construct $\text{Conv}(S)$. Let S^i , for $i = 1, \dots, m$, denote the set S where x is restricted to v^i i.e., $S^i = \{(x, y, z, \underline{xy}) \mid z = v^i g(y), \underline{xy} = v^i y, x = v^i, y \in [y^{\text{lo}}, y^{\text{up}}]\}$. Then, $\text{Conv}(S)$ is given as the convex hull of disjunctive union of S^i , $i = 1, \dots, m$, i.e., $\text{Conv}(S) = \text{Conv}(S^1 \cup \dots \cup S^m) = \text{Conv}(\text{Conv}(S^1) \cup \dots \cup \text{Conv}(S^m))$.

To determine $\text{Conv}(S^i)$, we reformulate each S^i as $S^i = \{(x, y, z, \underline{xy}, w) \mid z = v^i w, \underline{xy} = v^i y, w = g(y), x = v^i, y^{\text{lo}} \leq y \leq y^{\text{up}}\}$, which is an affine transform of

the set $\{(y, w) \in [y^{\text{lo}}, y^{\text{up}}] \times \mathbb{R} \mid w = g(y)\}$. This implies that it suffices to convexify the latter set to obtain $\text{Conv}(S^i) = \text{proj}_\pi\{(\pi, w) \mid (4.30)\}$, where

$$w \geq g(y), \quad (4.30a)$$

$$w \leq g(y^{\text{lo}}) + \left(\frac{g(y^{\text{up}}) - g(y^{\text{lo}})}{y^{\text{up}} - y^{\text{lo}}} \right) (y - y^{\text{lo}}), \quad (4.30b)$$

$$y^{\text{lo}} \leq y \leq y^{\text{up}}, \quad (4.30c)$$

$$z = v^i w, \quad \underline{xy} = v^i y, \quad x = v^i. \quad (4.30d)$$

The disjunctive union of $\text{Conv}(S^i)$, $i = 1, \dots, m$, leads to (4.28), where w^i and y^i are to be regarded as linearization of $\lambda^i w$ and $\lambda^i y$, respectively. $\square$

Remark 4.6. In Proposition 4.9, if $\text{Conv}(S^i)$ (see proof for definition) is bounded, closed and cone-quadratic representable (CQR), for $i = 1, \dots, m$, then $\text{Conv}(S)$ is CQR (see Proposition 3.3.5 in [51]). This result also applies to other conic representations. Let $P_3^{\delta, 1-\delta} := \{x \in \mathbb{R}^3 \mid x_1^\delta \cdot x_2^{1-\delta} \geq |x_3|\}$ where $0 < \delta < 1$ is the *power-cone*, and $K_{\text{exp}} = \{x_1 \geq x_2 \cdot \exp(x_3/x_2), x_2 > 0\} \cup \{(x_1, 0, x_3) \mid x_1 \geq 0, x_3 \leq 0\}$ is the *exponential-cone*. It is known that various elementary functions have cone representations [52]. For example, let $g(y) = |y|^\delta$ where $\delta > 1$ (*resp.* $g(y) = y^\delta$ where $\delta < 0$). Then, $w^i \geq g^*(\lambda^i, y^i)$ in Proposition 4.9 can be replaced with $(w^i, \lambda^i, y^i) \in P_3^{1/\delta, 1-1/\delta}$ (*resp.* $(w^i, y^i, \lambda^i) \in P_3^{1/(1-\delta), -\delta/(1-\delta)}$). For this work, we are interested in $\delta = -1$ and $\delta = 2$ (for reformulation with quadratic polynomials described in §4.3.1). Next, let $g(y) = -\ln(y)$, $y > 0$ (*resp.* $g(y) = \exp(y)$), which arises in formulations for identifying thermodynamically efficient distillation configurations (see [53]). Here, we replace $w^i \geq g^*(\lambda^i, y^i)$ in Proposition 4.9 with $(y^i, \lambda^i, -w^i) \in K_{\text{exp}}$ (*resp.* $(w^i, \lambda^i, y^i) \in K_{\text{exp}}$). $\square$

Remark 4.7. In Proposition 4.9, when $g(y)$ is nonlinear, the convex hull description has nonlinear constraints (see (4.28a)). To capitalize on LP solvers, we derive a polyhedral outer-approximation of $\text{Conv}(S)$ by outer-approximating the convex hull of each S^i before taking their disjunctive union. Let $\bar{y}^r \in [y^{\text{lo}}, y^{\text{up}}]$ for $r = 1, \dots, R$. Then, an outer-approximation of the convex hull of S^i is given

by $\text{Conv}_{OA}(S^i) = \text{proj}_\pi\{(\pi, w) \mid w \geq \max\{g(\bar{y}^r) + g'(\bar{y}^r)(y - \bar{y}^r)\}_{r=1}^R; (4.30b) - (4.30d)\}$, where $g'(y)$ denotes the first derivative of $g(y)$. The disjunctive union of $\text{Conv}_{OA}(S^i)$, $i = 1, \dots, m$, yields an outer-approximation of the convex hull of S , given by $\text{Conv}_{OA}(S) = \text{proj}_\pi\{(\pi, y^1, \dots, y^m, w^1, \dots, w^m, w, \lambda^1, \dots, \lambda^m) \mid w^i \geq \max\{g(\bar{y}^r)\lambda^i + g'(\bar{y}^r)(y^i - \bar{y}^r\lambda^i)\}_{r=1}^R; (4.28b) - (4.28e)\}$. $\square$

Now, consider the set $\mathcal{F}_p$. We lift $\mathcal{F}_p$ to a higher dimensional space by appending bilinear terms of the form $f_p \cdot \theta$ i.e., $\mathcal{F}_p = \{(f, \theta, H, \underline{f}\theta) \mid (4.24), \underline{f}\theta_p^{\text{in}} = f_p^{\text{in}} \cdot \theta, \underline{f}\theta_p^{\text{rs}} = f_p^{\text{rs}} \cdot \theta, \underline{f}\theta_p^{\text{ss}} = f_p^{\text{ss}} \cdot \theta\}$. Observe that the fractions and bilinear terms in $\mathcal{F}_p$ are defined over the polytope obtained by the intersection of hyperplane $f_p^{\text{in}} = f_p^{\text{rs}} + f_p^{\text{ss}}$ with the hypercube $[0, F_p]^3$ (see (4.24)). We now use Proposition 4.9 to obtain $\text{Conv}(\mathcal{F}_p) = \{(f_p, \theta, H_p, \underline{f}\theta_p) \mid (4.31)\}$ (see §B.4 for a detailed derivation), where

$$H_p^{\text{rs}} \geq F_p T_p^* \left(\frac{f_p^{\text{rs}}}{F_p}, \frac{f\theta_p^{\text{rs}}}{F_p} \right), \quad H_p^{\text{ss}} \geq F_p T_p^* \left(\frac{f_p^{\text{ss}}}{F_p}, \frac{f\theta_p^{\text{ss}}}{F_p} \right), \quad (4.31a)$$

$$H_p^{\text{rs}} \leq f_p^{\text{rs}} T_p(\theta^{\text{lo}}) + \left[\frac{T_p(\theta^{\text{up}}) - T_p(\theta^{\text{lo}})}{\theta^{\text{up}} - \theta^{\text{lo}}} \right] (\underline{f}\theta_p^{\text{rs}} - f_p^{\text{rs}} \theta^{\text{lo}}), \quad (4.31b)$$

$$H_p^{\text{ss}} \leq f_p^{\text{ss}} T_p(\theta^{\text{lo}}) + \left[\frac{T_p(\theta^{\text{up}}) - T_p(\theta^{\text{lo}})}{\theta^{\text{up}} - \theta^{\text{lo}}} \right] (\underline{f}\theta_p^{\text{ss}} - f_p^{\text{ss}} \theta^{\text{lo}}), \quad (4.31c)$$

$$(F_p - f_p^{\text{in}}) \theta^{\text{lo}} \leq F_p \theta - \underline{f}\theta_p^{\text{in}} \leq (F_p - f_p^{\text{in}}) \theta^{\text{up}}, \quad (4.31d)$$

$$f_p^{\text{rs}} \theta^{\text{lo}} \leq \underline{f}\theta_p^{\text{rs}} \leq f_p^{\text{rs}} \theta^{\text{up}}, \quad f_p^{\text{ss}} \theta^{\text{lo}} \leq \underline{f}\theta_p^{\text{ss}} \leq f_p^{\text{ss}} \theta^{\text{up}}, \quad (4.31e)$$

$$H_p^{\text{in}} = H_p^{\text{rs}} + H_p^{\text{ss}}, \quad \underline{f}\theta_p^{\text{in}} = \underline{f}\theta_p^{\text{rs}} + \underline{f}\theta_p^{\text{ss}}, \quad f_p^{\text{in}} = f_p^{\text{rs}} + f_p^{\text{ss}}. \quad (4.31f)$$

and the positively homogeneous function $T_p^*(\lambda, \theta)$ is defined as in (4.29) from $T_p(\theta)$. Note that the convex hull description does not require introduction of auxiliary variables. This is in contrast to the typical application of disjunctive programming, where new variables are introduced to derive the convex hull in a lifted space. We remark that the above yields a tighter relaxation of $\mathcal{F}_p$ compared to the one obtained by relaxing each fraction and bilinear term separately over the bounds of f_p^{in} , f_p^{rs} , f_p^{ss} , and θ . This is because the first two equations in (4.31f) are not implied in the latter set. Although, these relations can be obtained using RLT, appending these constraints does not result in (4.31). This is because, the set described in (4.31) is the simultane-

ous convex hull of the fraction and bilinear terms. It is known that the simultaneous hull of these functions is strictly contained in the intersection of their individual hulls (see Example 3.8 in [54]). In particular, (4.31b) and (4.31c), which are linearizations of $-\frac{f_p^{\text{rs}}}{|\alpha_p - \theta|} \cdot (\theta^{\text{up}} - \theta) \cdot (\theta - \theta^{\text{lo}}) \leq 0$ and $-\frac{f_p^{\text{ss}}}{|\alpha_p - \theta|} \cdot (\theta^{\text{up}} - \theta) \cdot (\theta - \theta^{\text{lo}}) \leq 0$ respectively, are not implied in the intersection of individual convex hulls.

The convex hull description in (4.31) is cone-quadratic representable (see Remark 4.6), since the constraints in (4.31a) can be expressed as second-order cones. For example, $H_1^{\text{rs}} \geq F_1 T_1^*(f_1^{\text{rs}}/F_1, \underline{f}\theta_1^{\text{rs}}/F_1) = (f_1^{\text{rs}})^2/(\alpha_1 f_1^{\text{rs}} - \underline{f}\theta_1^{\text{rs}})$, or $\sqrt{H_1} \cdot \sqrt{(\alpha_1 f_1^{\text{rs}} - \underline{f}\theta_1^{\text{rs}})} \geq |f_1^{\text{rs}}|$ (Note that $0 \leq \theta^{\text{lo}} f_1^{\text{rs}} - \underline{f}\theta_1^{\text{rs}} < \alpha_1 f_1^{\text{rs}} - \underline{f}\theta_1^{\text{rs}}$). However, we use the cone-quadratic representation only in §4.4. For our computational experiments in §4.5, we use its outer-approximation given by $\text{Conv}_{\text{OA}}(\mathcal{F}_p) = \{(f_p, \theta, H_p, \underline{f}\theta_p) \mid H_p^{\text{rs}} \geq \max\{f_p^{\text{rs}} T_p(\bar{\theta}^r) + T_p'(\bar{\theta}^r)(\underline{f}\theta_p^{\text{rs}} - \bar{\theta}^r f_p^{\text{rs}})\}_{r=1}^R, H_p^{\text{ss}} \geq \max\{f_p^{\text{ss}} T_p(\bar{\theta}^r) + T_p'(\bar{\theta}^r)(\underline{f}\theta_p^{\text{ss}} - \bar{\theta}^r f_p^{\text{ss}})\}_{r=1}^R, (4.31b) - (4.31f)\}$ for some $\bar{\theta}^r \in [\theta^{\text{lo}}, \theta^{\text{up}}]$, $r = 1, \dots, R$, where $T_p'(\bar{\theta}^r)$ denotes the derivative of $T_p(\theta)$ at $\bar{\theta}^r$; see Remark 4.7.

Next, consider the set $\mathcal{V}$, which contains bilinear terms defined over a polytope obtained by the intersection of a hyperrectangle in the positive orthant with the hyperplane $U^{\text{rs}} - U^{\text{ss}} = \Upsilon^{\text{rs}} - \Upsilon^{\text{ss}}$. Clearly, Proposition 4.9 can be used to construct the convex hull of $\mathcal{V}$ (only (4.28c), the equation with $\underline{xy}$ as the left-hand-side in (4.28d), and (4.28e) are needed to construct the hull). However, Proposition 4.9 requires enumeration of the extreme points of X . Instead, in this context, it is more convenient to directly use Proposition 4.10, which is a special case of Proposition 2.2 in [55], to obtain $\text{Conv}(\mathcal{V}) = \{(U, \Upsilon, \theta, \underline{U}\theta, \underline{\Upsilon}\theta) \mid (4.32)\}$, where

$$(\underline{U}\theta^{\text{rs}} - \theta^{\text{lo}} U^{\text{rs}}) - (\underline{U}\theta^{\text{ss}} - \theta^{\text{lo}} U^{\text{ss}}) = (\underline{\Upsilon}\theta^{\text{rs}} - \theta^{\text{lo}} \Upsilon^{\text{rs}}) - (\underline{\Upsilon}\theta^{\text{ss}} - \theta^{\text{lo}} \Upsilon^{\text{ss}}), \quad (4.32a)$$

$$(\theta^{\text{up}} U^{\text{rs}} - \underline{U}\theta^{\text{rs}}) - (\theta^{\text{up}} U^{\text{ss}} - \underline{U}\theta^{\text{ss}}) = (\theta^{\text{up}} \Upsilon^{\text{rs}} - \underline{\Upsilon}\theta^{\text{rs}}) - (\theta^{\text{up}} \Upsilon^{\text{ss}} - \underline{\Upsilon}\theta^{\text{ss}}), \quad (4.32b)$$

$$0 \leq \underline{(\cdot)}\theta - \theta^{\text{lo}} (\cdot) \leq (\cdot)^{\text{up}} (\theta - \theta^{\text{lo}}), \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}, \quad (4.32c)$$

$$0 \leq \theta^{\text{up}} (\cdot) - \underline{(\cdot)}\theta \leq (\cdot)^{\text{up}} (\theta^{\text{up}} - \theta), \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}. \quad (4.32d)$$

Finally, we construct the convex relaxation of $\mathcal{U}$ as, $\mathcal{U}_{\text{Relax}} = \{(f, U, \Upsilon, \theta, H, \underline{U}\theta, \underline{\Upsilon}\theta, \underline{f}\theta) \mid (4.23), (f_p, H_p, \theta, \underline{f}\theta_p) \in \text{Conv}(\mathcal{F}_p), p = 1, 2, (U, \Upsilon, \theta, \underline{U}\theta, \underline{\Upsilon}\theta) \in \text{Conv}(\mathcal{V})\}$.

Proposition 4.10 ([55]). Let $X = \{x \in \mathbb{R}^n \mid Bx \leq b\}$ be a polytope, $\mathcal{D} = X \times [y^{\text{lo}}, y^{\text{up}}] \times \mathbb{R}^n$, and $S = \{(x, y, z) \in \mathcal{D} \mid \underline{xy}_j = x_j \cdot y, j = 1, \dots, n\}$. Then, $\text{Conv}(S) = \{(x, y, \underline{xy}) \mid y^{\text{lo}} \leq y \leq y^{\text{up}}, B(\underline{xy} - y^{\text{lo}}x) \leq b(y - y^{\text{lo}}), B(y^{\text{up}}x - \underline{xy}) \leq b(y^{\text{up}} - y)\}$.

Proof. When x is restricted to $\bar{x} \in X$, the set $S = \{(x, y, \underline{xy}) \in \mathcal{D} \mid \underline{xy} = \bar{x} \cdot y, x = \bar{x}\}$ can be expressed as an affine transformation of $y^{\text{lo}} \leq y \leq y^{\text{up}}$, whose extreme points are $y \in \{y^{\text{lo}}, y^{\text{up}}\}$. Therefore, the extreme points of convex hull of S are contained in the set of points where $y \in \{y^{\text{lo}}, y^{\text{up}}\}$. Let $S^1 = \{(x, y, \underline{xy}) \mid \underline{xy} = y^{\text{lo}}x, Bx \leq b, y = y^{\text{lo}}\}$ and $S^2 = \{(x, y, \underline{xy}) \mid \underline{xy} = y^{\text{up}}x, Bx \leq b, y = y^{\text{up}}\}$. Then, by Krein-Milman theorem, convex hull of S is obtained by taking the disjunctive union of S^1 and S^2 , i.e., $\text{Conv}(S) = \text{proj}_{(x, y, \underline{xy})} \{(x, y, \underline{xy}, x^1, x^2, \lambda^1, \lambda^2) \mid Bx^i \leq b\lambda^i, i = 1, 2, (4.33), \lambda^1 \geq 0, \lambda^2 \geq 0\}$, where

$$\underline{xy} = x^1 y^{\text{lo}} + x^2 y^{\text{up}}, \quad x = x^1 + x^2, \quad (4.33a)$$

$$y = y^{\text{lo}}\lambda^1 + y^{\text{up}}\lambda^2, \quad \lambda^1 + \lambda^2 = 1, \quad (4.33b)$$

Solving the above equations leads to

$$x^1 = \frac{y^{\text{up}}x - \underline{xy}}{y^{\text{up}} - y^{\text{lo}}}, \quad x^2 = \frac{\underline{xy} - y^{\text{lo}}x}{y^{\text{up}} - y^{\text{lo}}}, \quad \lambda^1 = \frac{y^{\text{up}} - y}{y^{\text{up}} - y^{\text{lo}}}, \quad \lambda^2 = \frac{y - y^{\text{lo}}}{y^{\text{up}} - y^{\text{lo}}}. \quad (4.34)$$

Using the above relations, we substitute out x^1, x^2, λ^1 and λ^2 to obtain the convex hull description in the proposition. □

Remark 4.8. We remark that $\text{Conv}(\mathcal{F}_p)$ and $\text{Conv}(\mathcal{V})$ in (4.31) and (4.32) imply the convex envelope of $\sum_{p=1}^2 \left[-\frac{\alpha_p(\alpha_p - \theta^{\text{up}})f_p^{\text{rs}}}{\alpha_p - \theta} + \alpha_p f_p^{\text{rs}} \right] + \Upsilon^{\text{rs}} \cdot \theta - \Upsilon^{\text{rs}} \cdot \theta^{\text{up}}$ over bound constraints on f_p^{rs}, θ , and Υ^{rs} (see (4.20b)). This is because, when all f_p^{rs} and Υ^{rs} are fixed, the function is concave in θ . Then, by Theorem 1.4 in [56], it follows that the convex envelope is obtained by replacing $-\frac{\alpha_p(\alpha_p - \theta^{\text{up}})f_p^{\text{rs}}}{\alpha_p - \theta}$ for all p and $\Upsilon^{\text{rs}} \cdot \theta$ by their convex envelopes. □

We comment on the construction of convex relaxations of $\mathcal{U}$ when additional RDLT cuts described in §4.3.1 are appended to $\mathcal{U}_{\text{ref}}$. Reformulation of Underwood

constraints using quadratic polynomials of θ introduces nonconvex terms of the form $f_p \cdot \theta$, $\Upsilon \cdot \theta^2$ (see (4.26)), in addition to the existing $f_p \cdot T_p(\theta)$ and $\Upsilon \cdot \theta$ terms in $\mathcal{U}_{\text{ref}}$. We relax $f_p \cdot T_p(\theta)$ and $f_p \cdot \theta$ using the simultaneous hull description in (4.31). Although Proposition 4.9 yields the simultaneous hull of $\Upsilon \cdot \theta^2$ and $\Upsilon \cdot \theta$ terms over the polytope in $\mathcal{V}$, we do not implement this relaxation. This is because the hull description does not project onto the space of problem variables in a straightforward manner. Instead, we convexify each pair of $\Upsilon \cdot \theta^2$ and $\Upsilon \cdot \theta$ terms over a box using Corollary 4.2, and append the RLT cuts $U^{\text{rs}} \cdot \theta^2 - U^{\text{ss}} \cdot \theta^2 = \Upsilon^{\text{rs}} \cdot \theta^2 - \Upsilon^{\text{ss}} \cdot \theta^2$ and $U^{\text{rs}} \cdot \theta - U^{\text{ss}} \cdot \theta = \Upsilon^{\text{rs}} \cdot \theta - \Upsilon^{\text{ss}} \cdot \theta$.

On the other hand, reformulation of Underwood constraints using inverse bound factors introduces nonconvex terms of the form $f_p \cdot \theta^{-1}$ and $\Upsilon \cdot \theta^{-1}$ (see (4.27)), in addition to the existing $f_p \cdot T_p(\theta)$ and $\underline{\Upsilon}\theta$ terms in $\mathcal{U}_{\text{ref}}$. We relax $f_p \cdot T_p(\theta)$ and $f_p \cdot \theta$ using the simultaneous hull description in (4.31). We use a similar hull description, obtained using Proposition 4.9, to relax $f_p \cdot \theta^{-1}$ and $f_p \cdot \theta$. Finally, for the same reason mentioned above, we convexify each pair of $\Upsilon \cdot \theta^{-1}$ and $\Upsilon \cdot \theta$ terms using Corollary 4.2, and append RLT cuts $U^{\text{rs}} \cdot \theta^{-1} - U^{\text{ss}} \cdot \theta^{-1} = \Upsilon^{\text{rs}} \cdot \theta^{-1} - \Upsilon^{\text{ss}} \cdot \theta^{-1}$ and $U^{\text{rs}} \cdot \theta - U^{\text{ss}} \cdot \theta = \Upsilon^{\text{rs}} \cdot \theta - \Upsilon^{\text{ss}} \cdot \theta$.

Corollary 4.2. Let $\mathcal{B} = [x^{\text{lo}}, x^{\text{up}}] \times [y^{\text{lo}}, y^{\text{up}}] \times \mathbb{R}^2$, where we assume $0 \leq x^{\text{lo}}$, $g(y) : [y^{\text{lo}}, y^{\text{up}}] \rightarrow \mathbb{R}$ is convex, and $S = \{(x, y, z, \underline{xy}) \in \mathcal{B} \mid \underline{xy} = x \cdot y, z = x \cdot g(y)\}$. Then, $\text{Conv}(S) = \{(x, y, z, \underline{xy}) \mid (4.35)\}$, where

$$z \geq x^{\text{lo}} g^* \left(\frac{x^{\text{up}} - x}{x^{\text{up}} - x^{\text{lo}}}, \frac{x^{\text{up}} y - \underline{xy}}{x^{\text{up}} - x^{\text{lo}}} \right) + x^{\text{up}} g^* \left(\frac{x - x^{\text{lo}}}{x^{\text{up}} - x^{\text{lo}}}, \frac{\underline{xy} - x^{\text{lo}} y}{x^{\text{up}} - x^{\text{lo}}} \right), \quad (4.35a)$$

$$z \leq g(y^{\text{lo}}) \cdot x + \left[\frac{g(y^{\text{up}}) - g(y^{\text{lo}})}{y^{\text{up}} - y^{\text{lo}}} \right] (\underline{xy} - y^{\text{lo}} x), \quad (4.35b)$$

$$(x^{\text{up}} - x) y^{\text{lo}} \leq x^{\text{up}} y - \underline{xy} \leq (x^{\text{up}} - x) y^{\text{up}}, \quad (x - x^{\text{lo}}) y^{\text{lo}} \leq \underline{xy} - x^{\text{lo}} y \leq (x - x^{\text{lo}}) y^{\text{up}}, \quad (4.35c)$$

and $g^*(\lambda^*, y^*)$ is defined as in (4.29). Further, the outer-approximation of the convex hull is $\text{Conv}_{OA}(S) = \{(x, y, z, \underline{xy}) \mid (4.36), (4.35b) - (4.35c)\}$, where

$$z \geq \frac{x^{\text{lo}}}{x^{\text{up}} - x^{\text{lo}}} \max \{ g(\bar{y}^r) (x^{\text{up}} - x) + g'(\bar{y}^r) (x^{\text{up}} y - \underline{xy} - (x^{\text{up}} - x) \bar{y}^r) \}_{r=1}^R +$$

$$\frac{x^{\text{up}}}{x^{\text{up}} - x^{\text{lo}}} \max\{g(\bar{y}^r)(x - x^{\text{lo}}) + g'(\bar{y}^r)(\underline{xy} - x^{\text{lo}}y - (x - x^{\text{lo}})\bar{y}^r)\}_{r=1}^R, \quad (4.36)$$

for some $\bar{y}^r \in [y^{\text{lo}}, y^{\text{up}}]$, $r = 1, \dots, R$, and $g'(y)$ denotes the first derivative of $g(y)$ w.r.t y .

Proof. See §B.5 in the Appendix. $\square$

4.3.3 Valid Relaxation for $\theta^{\text{lo}} = \alpha_2$ and/or $\theta^{\text{up}} = \alpha_1$

In the previous subsection, we have assumed that $\alpha_2 < \theta^{\text{lo}}$ and $\theta^{\text{up}} < \alpha_1$. Instead, if α_1 and/or α_2 is an admissible value of θ , we cannot directly use (4.31) to convexify $\mathcal{F}_p$, because $T_1(\alpha_1)$ and $T_2(\alpha_2)$ are not well-defined. To construct a valid relaxation, we first restrict the admissible values of θ to a subset of the interval $[\alpha_2, \alpha_1]$ by recognizing that each fraction in $\mathcal{U}$ is bounded.

Proposition 4.11. (i) Valid upper bounds on $\frac{f_1^{\text{rs}}}{\alpha_1 - \theta}$, $(H_1^{\text{rs}})^{\text{up}}$, and on $\frac{f_2^{\text{rs}}}{\theta - \alpha_2}$, $(H_2^{\text{rs}})^{\text{up}}$, are given by

$$(H_1^{\text{rs}})^{\text{up}} = \frac{(\Upsilon^{\text{rs}})^{\text{up}}(\alpha_1 - \alpha_2) + \alpha_1 F_1 + \alpha_2 F_2}{\alpha_1(\alpha_1 - \alpha_2)}, \quad (4.37a)$$

$$(H_2^{\text{rs}})^{\text{up}} = \frac{\alpha_1 F_1 + \alpha_2 F_2 - (E^{\text{rs}} \cdot (\alpha_1 - \theta))^{\text{lo}}}{\alpha_2(\alpha_1 - \alpha_2)}. \quad (4.37b)$$

(ii) the admissible region of θ in the interval $[\alpha_2, \alpha_1]$ is given by

$$\alpha_2 + \frac{f_2^{\text{rs}}}{(H_2^{\text{rs}})^{\text{up}}} \leq \theta \leq \alpha_1 - \frac{f_1^{\text{rs}}}{(H_1^{\text{rs}})^{\text{up}}}. \quad (4.38)$$

Proof. (i) Consider the second inequality in (4.23c). Since this inequality holds for any θ^{lo} less than θ , if we substitute θ^{lo} with α_2 , the inequality remains valid. Then, we obtain (4.37a) from $\alpha_1(\alpha_1 - \alpha_2)H_1^{\text{rs}} \leq (\underline{\Upsilon}\theta^{\text{rs}} - \alpha_2\Upsilon^{\text{rs}}) + \alpha_1f_1^{\text{rs}} + \alpha_2f_2^{\text{rs}} \leq (\Upsilon^{\text{rs}})^{\text{up}}(\alpha_1 - \alpha_2) + \alpha_1F_1 + \alpha_2F_2$, where the last inequality is because $f_1^{\text{rs}} \leq F_1$, $f_2^{\text{rs}} \leq F_2$, and $\underline{\Upsilon}\theta^{\text{rs}} - \alpha_2\Upsilon^{\text{rs}} \leq (\Upsilon^{\text{rs}})^{\text{up}}(\alpha_1 - \alpha_2)$. Similarly, we substitute $\theta^{\text{up}} = \alpha_1$ in the first inequality in (4.23d), and rearrange to get $\alpha_2(\alpha_1 - \alpha_2)H_2^{\text{rs}} \leq -E^{\text{rs}} \cdot (\alpha_1 - \theta) + \alpha_1f_1^{\text{rs}} + \alpha f_2^{\text{rs}}$. We maximize the right hand side by substituting $f_1^{\text{rs}} = F_1$, $f_2^{\text{rs}} = F_2$, and $(E^{\text{rs}} \cdot (\alpha_1 - \theta))$

by its lower bound which is computed using the bounds on E^{rs} and θ . This leads to the bound in (4.37b).

(ii) Every point feasible to $\mathcal{U}$ satisfies $f_1^{\text{rs}}/(\alpha_1 - \theta) \leq (H_1^{\text{rs}})^{\text{up}}$ and $f_2^{\text{rs}}/(\theta - \alpha_2) \leq (H_2^{\text{rs}})^{\text{up}}$. Rearranging the inequalities yields (4.38). $\square$

We remark that the bounds on H_p^{in} and H_p^{ss} for $p = 1, 2$ can be computed in the same manner as in the proof of (i) in Proposition 4.11. Even when additional fractions are present in the Underwood constraints, each fraction can be bounded, since the remaining fractions are strictly bounded in the interval of θ . We revisit the argument on bounds of θ in light of Proposition 4.11. As mentioned before, the common approach used in the literature to overcome the singularity arising due to θ approaching one of the adjoining relative volatilities has been to restrict θ to belong to $[\alpha_2 + \epsilon_\theta, \alpha_1 - \epsilon_\theta]$. However, observe that our bounds in (4.38) depend on f_1^{rs} and f_2^{rs} . This explains the difficulty we encountered in choosing a value for ϵ_θ in our computations with prior formulations. We have found that there are instances when θ is fairly close to one of the relative volatilities, particularly when the corresponding flow is small. We will provide a rigorous approach to addressing this singularity using (4.38). Our approach will be to construct a relaxation of $\mathcal{F}_1$ as the intersection of simultaneous convex hulls of $f_1 \cdot T_1(\theta)$ and $f_1 \cdot \theta$. For brevity, we only discuss the relaxation for $\mathcal{F}_1$ in detail, and remark that a similar result is easily derived for $\mathcal{F}_2$.

Proposition 4.12. *Let $\mathcal{H}_1 = \{(f_1, \theta, H_1, \underline{f}\theta_1) \mid 0 \leq f_1 \leq F_1, \theta^{\text{lo}} \leq \theta \leq \alpha_1 - f_1/H_1^{\text{up}}, H_1 = f_1 \cdot T_1(\theta), \text{ if } \theta < \alpha_1; H_1 \in [0, H_1^{\text{up}}] \text{ if } \theta = \alpha_1, \underline{f}\theta_1 = f_1 \cdot \theta\}$, where $\alpha_2 \leq \theta^{\text{lo}}$. Then, $\text{Conv}(\mathcal{H}_1) = \text{proj}_{(f_1, \theta, H_1, \underline{f}\theta_1)}\{(f_1, \theta, H_1, \underline{f}\theta_1, \theta^a, \theta^b, \theta^c, \lambda^a, \lambda^b, \lambda^c) \mid (4.39)\}$, where*

$$\left. \begin{aligned} H_1 &\geq H_1^{\text{up}}\lambda^b + F_1 T_1^*(\lambda^c, \theta^c) \\ H_1 &\leq H_1^{\text{up}} \left(\frac{\theta^a - \theta^{\text{lo}}\lambda^a}{\alpha_1 - \theta^{\text{lo}}} \right) + H_1^{\text{up}}\lambda^b + \frac{F_1\lambda^c}{\alpha_1 - \theta^{\text{lo}}} + H_1^{\text{up}} \left(\frac{\theta^c - \theta^{\text{lo}}\lambda^c}{\alpha_1 - \theta^{\text{lo}}} \right) \end{aligned} \right\} \quad (4.39a)$$

$$\left. \begin{aligned} \underline{f}\theta_1 &\geq H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) (\alpha_1\lambda^b - \theta^b) + F_1\theta^c \\ \underline{f}\theta_1 &\leq \underline{f}\theta_1 \leq H_1^{\text{up}} \left(\alpha_1\theta^b - \frac{(\theta^b)^2}{\lambda^b} \right) + F_1\theta^c \end{aligned} \right\} \quad (4.39b)$$

$$\left. \begin{aligned} \theta^{\text{lo}} \lambda^a &\leq \theta^a \leq \alpha_1 \lambda^a \\ \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \lambda^b &\leq \theta^b \leq \alpha_1 \lambda^b \\ \theta^{\text{lo}} \lambda^c &\leq \theta^c \leq \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \lambda^c \end{aligned} \right\} \quad (4.39c)$$

$$\left. \begin{aligned} f_1 &= H_1^{\text{up}}(\alpha_1 \lambda^b - \theta^b) + F_1 \lambda^c \\ \theta &= \theta^a + \theta^b + \theta^c \\ \lambda^a + \lambda^b + \lambda^c &= 1, \quad \lambda^a, \lambda^b, \lambda^c \geq 0. \end{aligned} \right\} \quad (4.39d)$$

Proof. See §B.6 in the Appendix. $\square$

The convex hull in Proposition 4.12 requires several additional variables. To avoid the introduction of these additional variables, we use its relaxation, $\mathcal{H}_{1,\text{Relax}}$, derived in §B.7 and shown below:

$$\max \left\{ f_1 T_1(\bar{\theta}^r) + T_1'(\bar{\theta}^r)(\underline{f}\theta_1 - \bar{\theta}^r f_1) \right\}_{r=1}^R \leq H_1 \leq \frac{f_1}{\alpha_1 - \theta^{\text{lo}}} + H_1^{\text{up}} \left(\frac{\theta - \theta^{\text{lo}}}{\alpha_1 - \theta^{\text{lo}}} \right), \quad (4.40a)$$

$$\max \{ \theta^{\text{lo}} f_1, F_1 \theta + \alpha_1 f_1 - \alpha_1 F_1 \} \leq \min \{ \alpha_1 f_1, F_1 \theta + \theta^{\text{lo}} f_1 - \theta^{\text{lo}} F_1 \}, \quad (4.40b)$$

$$\theta^{\text{lo}} \leq \theta \leq \alpha_1 - \frac{f_1}{H_1^{\text{up}}}, \quad (4.40c)$$

where $\bar{\theta}^r \in [\theta^{\text{lo}}, \alpha_1)$, $r = 1, \dots, R$. Here, we argue from first principles that (4.40) is a valid relaxation. To derive the first inequality in (4.40a), observe that $H_1 \geq f_1 \cdot T_1(\theta) \geq f_1 \cdot \max \{ T_1(\bar{\theta}^r) + T_1'(\bar{\theta}^r)(\theta - \bar{\theta}^r) \}_{r=1}^R$. Disaggregating the product and linearizing the bilinear term yields (4.40a). To derive the second inequality in (4.40a), we begin with $H_1 \cdot (\alpha_1 - \theta) \leq f_1$, and replace the bilinear term on the left hand side with its convex envelope. (4.40b) is the convex hull of $\underline{f}\theta_1 = f_1 \cdot \theta$ over $[0, F_1] \times [\theta^{\text{lo}}, \alpha_1]$, and (4.40c) is the same as (4.38). Using (4.40), we obtain a valid relaxation of $\mathcal{F}_1$ given by $\mathcal{F}_{1,\text{Relax}} = \{ (f_1, \theta, H_1, \underline{f}\theta_1) \mid (f_1^{\text{in}}, \theta, H_1^{\text{in}}, \underline{f}\theta_1^{\text{in}}) \in \mathcal{H}_{1,\text{Relax}}^{\text{in}}, (f_1^{\text{rs}}, \theta, H_1^{\text{rs}}, \underline{f}\theta_1^{\text{rs}}) \in \mathcal{H}_{1,\text{Relax}}^{\text{rs}}, (f_1^{\text{rs}}, \theta, H_1^{\text{rs}}, \underline{f}\theta_1^{\text{rs}}) \in \mathcal{H}_{1,\text{Relax}}^{\text{ss}}, H_1^{\text{in}} = H_1^{\text{rs}} + H_1^{\text{ss}}, \underline{f}\theta_1^{\text{in}} = \underline{f}\theta_1^{\text{rs}} + \underline{f}\theta_1^{\text{ss}} \}$. Inspired from (4.31), the last two equations in the relaxation are derived by multiplying the component mass balance, (4.18f), with $T_1(\theta)$ and θ , respectively.

4.3.4 Discretization and Solution Procedure

In this work, instead of using convex relaxations of $\mathcal{U}$ in a spatial branch-and-bound framework to solve the MINLP, we construct a piecewise relaxation (see Definition 4.1) that is iteratively improved until we prove ϵ_r -optimality. This approach capitalizes on state-of-the-art MIP solvers, such as Gurobi.

Definition 4.1 (Piecewise Relaxation). *Let $x = (x_1, \dots, x_n)$, $\mathcal{B} = [x^{\text{lo}}, x^{\text{up}}] \times [y^{\text{lo}}, y^{\text{up}}] \subset \mathbb{R}^{n+1}$, $S = \{(x, y) \in \mathcal{B} \mid g_i(x, y) \leq 0, i = 1, \dots, m\}$, and $S_{\text{Relax}} = \{(x, y) \in \mathcal{B} \mid \check{g}_i(x, y) \leq 0, i = 1, \dots, m\}$ be its convex relaxation, where $\{\check{g}_i\}_{i=1}^m$ denote convex underestimators of $\{g_i\}_{i=1}^m$ over $\mathcal{B}$. Let, the domain of y be partitioned as $\mathcal{I} = \{[Y^0, Y^1], \dots, [Y^{|\mathcal{I}|-1}, Y^{|\mathcal{I}|}]\}$ with $Y^0 = y^{\text{lo}}$, $Y^{|\mathcal{I}|} = y^{\text{up}}$ and $Y^0 \leq Y^1 \leq \dots \leq Y^{|\mathcal{I}|}$. By piecewise relaxation of S , we refer to $\bigcup_{t=1}^{|\mathcal{I}|} S_{t,\text{Relax}}$, where $S_{t,\text{Relax}} = \{(x, y, z) \in \mathcal{B}_t \mid \check{g}_{i,t}(x, y) \leq 0, i = 1, \dots, m\}$, $\mathcal{B}_t = [x^{\text{lo}}, x^{\text{up}}] \times [Y^{t-1}, Y^t]$, and $\check{g}_{i,t}$ is the convex under-estimator of g_i over $\mathcal{B}_t$. $\square$*

Piecewise relaxation of $\mathcal{U}$ can be constructed by partitioning the domain of Underwood root as $\mathcal{I} = \{[\Theta^0, \Theta^1], \dots, [\Theta^{|\mathcal{I}|-1}, \Theta^{|\mathcal{I}|}]\}$, where $\Theta^0 = \alpha_2$, $\Theta^{|\mathcal{I}|} = \alpha_1$, and $\Theta^0 \leq \Theta^1 \leq \dots \leq \Theta^{|\mathcal{I}|}$, and taking the union of sets $\bigcup_{t=1}^{|\mathcal{I}|} \mathcal{U}_{t,\text{Relax}}$, where $\mathcal{U}_{t,\text{Relax}}$ denotes the convex relaxation of $\mathcal{U}$ restricted to $\theta \in [\Theta^{t-1}, \Theta^t]$. The set $\mathcal{U}_{t,\text{Relax}}$ is constructed as outlined in §4.3.2 and §4.3.3. Next, using standard disjunctive programming techniques, the piecewise relaxation can be expressed as a Mixed Integer Program (MIP). While this approach leads to a *locally ideal* formulation, it leads to a bigger problem size, because of which the computational time required is higher. Thus, in favor of smaller problem size, we do the following.

Instead of reformulating $\mathcal{U}$ in each partition using the local bound factors of θ , we reformulate with the overall bound factors of θ : $(\theta - \alpha_2)$ and $(\alpha_1 - \theta)$. Next, we require that $(f_p, \theta, H_p, \underline{f}\theta_p)$, $p = 1, 2$, and $(U, \Upsilon, \theta, \underline{U}\theta, \underline{\Upsilon}\theta)$ lie in piecewise relaxations of $\mathcal{F}_p$ and $\mathcal{V}$, respectively. We choose piecewise relaxation of $\mathcal{F}_1$ to be $\bigcup_{t=1}^{|\mathcal{I}|-1} \text{Conv}_{\text{OA}}(\mathcal{F}_{1,t}) \cup \mathcal{F}_{1,|\mathcal{I}|,\text{Relax}}$, piecewise relaxation of $\mathcal{F}_2$ to be $\mathcal{F}_{2,1,\text{Relax}} \cup \bigcup_{t=2}^{|\mathcal{I}|} \text{Conv}_{\text{OA}}(\mathcal{F}_{2,t})$, and piecewise relaxation of $\mathcal{V}$ to be $\bigcup_{t=1}^{|\mathcal{I}|} \text{Conv}(\mathcal{V}_t)$. Here, the additional subscript t denotes

that the set is restricted to $\theta \in [\Theta^{t-1}, \Theta^t]$. Observe that if zero is not an admissible value to the denominators of the fractions, we use outer-approximation of convex hulls derived in §4.3.2 to relax $\mathcal{F}_p$. Otherwise, we use a relaxation of the convex hull description, such as the one derived in §4.3.3. We use disjunctive programming to express the piecewise relaxations as the following mixed-integer sets.

$$[t]_1^{|I|}, \quad H_1^{\text{rs}} \geq f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}), \quad (4.41\text{a})$$

$$[t]_1^{|I|}, \quad H_1^{\text{rs}} \geq f_{1,t}^{\text{rs}} T_1(\Theta^t) + T_1'(\Theta^t)(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^t f_{1,t}^{\text{rs}}), \quad (4.41\text{b})$$

$$[t]_1^{|I|}, \quad H_1^{\text{ss}} \geq f_{1,t}^{\text{ss}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}), \quad (4.41\text{c})$$

$$[t]_1^{|I|}, \quad H_1^{\text{ss}} \geq f_{1,t}^{\text{ss}} T_1(\Theta^t) + T_1'(\Theta^t)(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^t f_{1,t}^{\text{ss}}), \quad (4.41\text{d})$$

$$\begin{aligned} H_1^{\text{rs}} \leq & \sum_{t=1}^{|I|-1} f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + \left[\frac{T_1(\Theta^{t-1}) - T_1(\Theta^t)}{\Theta^{t-1} - \Theta^t} \right] (\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}) \\ & + \frac{f_{1,|I|}^{\text{rs}}}{\alpha_1 - \Theta^{|I|-1}} + (H_1^{\text{rs}})^{\text{up}} \left[\frac{\theta_{|I|} - \Theta^{|I|-1} \mu_t}{\Theta^{|I|} - \Theta^{|I|-1}} \right] \end{aligned} \quad (4.41\text{e})$$

$$\begin{aligned} H_1^{\text{ss}} \leq & \sum_{t=1}^{|I|-1} f_{1,t}^{\text{ss}} T_1(\Theta^{t-1}) + \left[\frac{T_1(\Theta^{t-1}) - T_1(\Theta^t)}{\Theta^{t-1} - \Theta^t} \right] (\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}) \\ & + \frac{f_{1,|I|}^{\text{ss}}}{\alpha_1 - \Theta^{|I|-1}} + (H_1^{\text{ss}})^{\text{up}} \left[\frac{\theta_{|I|} - \Theta^{|I|-1} \mu_t}{\Theta^{|I|} - \Theta^{|I|-1}} \right] \end{aligned} \quad (4.41\text{f})$$

$$[t]_1^{|I|}, \quad (F_1 \mu_t - f_{1,t}^{\text{rs}} - f_{1,t}^{\text{ss}}) \Theta^{t-1} \leq (F_1 \theta_t - \underline{f}\theta_{1,t}^{\text{rs}} - \underline{f}\theta_{1,t}^{\text{ss}}) \leq (F_1 \mu_t - f_{1,t}^{\text{rs}} - f_{1,t}^{\text{ss}}) \Theta^t, \quad (4.41\text{g})$$

$$[t]_1^{|I|}, \quad f_{1,t}^{\text{rs}} \Theta^{t-1} \leq \underline{f}\theta_{1,t}^{\text{rs}} \leq f_{1,t}^{\text{rs}} \Theta^t, \quad f_{1,t}^{\text{ss}} \Theta^{t-1} \leq \underline{f}\theta_{1,t}^{\text{ss}} \leq f_{1,t}^{\text{ss}} \Theta^t, \quad (4.41\text{h})$$

$$H_1^{\text{in}} = H_1^{\text{rs}} + H_1^{\text{ss}}, \quad \underline{f}\theta_1^{\text{in}} = \underline{f}\theta_1^{\text{rs}} + \underline{f}\theta_1^{\text{ss}} \quad (4.41\text{i})$$

$$f_1^{\text{rs}} = \sum_{t=1}^{|I|} f_{1,t}^{\text{rs}}, \quad f_1^{\text{ss}} = \sum_{t=1}^{|I|} f_{1,t}^{\text{ss}}, \quad \theta = \sum_{t=1}^{|I|} \theta_t \quad (4.41\text{j})$$

$$\sum_{t=1}^{|I|} \mu_t = 1, \quad \mu_t \in \{0, 1\}, \quad [t]_1^{|I|} \quad (4.41\text{k})$$

and

$$\underline{U}\theta^{\text{rs}} - \underline{U}\theta^{\text{ss}} = \underline{\Upsilon}\theta^{\text{rs}} - \underline{\Upsilon}\theta^{\text{ss}} \quad (4.42\text{a})$$

$$[t]_1^{|I|}, \quad U_t^{\text{rs}} - U_t^{\text{ss}} = \Upsilon_t^{\text{rs}} - \Upsilon_t^{\text{ss}}, \quad (4.42\text{b})$$

$$\llbracket t \rrbracket_1^{|I|}, \quad 0 \leq \underline{(\cdot)\theta} - \sum_{t=1}^{|I|} \Theta^{t-1}(\cdot)_t \leq (\cdot)^{\text{up}}\theta - (\cdot)^{\text{up}} \sum_{t=1}^{|I|} \Theta^{t-1}\mu_t, \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}, \quad (4.42\text{c})$$

$$\llbracket t \rrbracket_1^{|I|}, \quad 0 \leq \sum_{t=1}^{|I|} \Theta^t(\cdot)_t - \underline{(\cdot)\theta} \leq (\cdot)^{\text{up}} \sum_{t=1}^{|I|} \Theta^t\mu_t - (\cdot)^{\text{up}}\theta, \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}, \quad (4.42\text{d})$$

$$(\cdot) = \sum_{t=1}^{|I|} (\cdot)_t, \quad 0 \leq (\cdot) \leq (\cdot)^{\text{up}}\mu_t, \quad ; \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\} \quad (4.42\text{e})$$

$$\sum_{t=1}^{|I|} \mu_t = 1, \quad \mu_t \in \{0, 1\}, \quad \llbracket t \rrbracket_1^{|I|}. \quad (4.42\text{f})$$

The piecewise relaxation of $\mathcal{F}_2$ can be expressed as a mixed-integer set in a similar manner. The derivation of these sets is provided in §B.8 and §B.9.

In §4.4, we illustrate through numerical examples the impact of various aspects described in this section in strengthening the overall relaxation of MINLP (A). Finally, Algorithm 2 outlines our approach to solve the MINLP. We start with a coarse discretization and use an adaptive partitioning scheme to iteratively refine the partitions until ϵ_r -optimality is achieved. To avoid numerical issues, we maintain that each partition, $(\Theta_{ijq}^t - \Theta_{ijq}^{t-1})$, is at least `MinPrtSize` in length.

4.4 Effect of Individual Cuts on Relaxation

This section illustrates, through numerical examples, the impact of various aspects described in §4.3 in strengthening the overall relaxation of MINLP (A). We highlight the individual effect of RDLT cuts derived from Underwood constraints, simultaneous hulls derived in §4.3.2, and discretization on the overall relaxation. In all the scenarios below, stream and heat exchanger variables are considered to be binary.

Scenario 1 : (BARON's root node relaxation) Here, we use BARON 18.5.8, on GAMS 25.1, to construct and solve the relaxation of MINLP (A). This is achieved by specifying BARON option `MaxIter` = 1, which terminates the branch-and-cut

Input : $N, \alpha = (\alpha_1, \dots, \alpha_N), F = (F_1, \dots, F_N), \Phi = (\Phi_{1,N}, \Phi_{1,1}, \dots, \Phi_{N,N})$

Output : Vectors $\mathbf{y}$ and $\mathbf{x}$ containing optimal values of discrete (streams and heat exchangers present in the configuration), and continuous (material flows in columns and heat exchangers) variables, respectively

Parameters: Relative tolerance for convergence $\epsilon_r = 0.01$, Minimum length of each partition $\text{MinPrtSize} = 10^{-3}$.

```

1 Function  $[\mathbf{y}, \mathbf{x}] = \text{VAPORDUTY}(N, \alpha, F, \Phi)$ 
2   Initialization: For every  $[i, j] \in \mathcal{S}$ ,  $\llbracket q \rrbracket_i^{j-1}, \mathcal{I}_{ijq} \leftarrow \left\{ \left[ \Theta_{ijq}^0, \Theta_{ijq}^1 \right], \dots, \left[ \Theta_{ijq}^{|\mathcal{I}_{ijq}|-1}, \Theta_{ijq}^{|\mathcal{I}_{ijq}|} \right] \right\}$ , where
    $\Theta_{ijq}^0 = \alpha_{q+1}, \Theta_{ijq}^{|\mathcal{I}_{ijq}|} = \alpha_q, \Theta_{ijq}^0 < \Theta_{ijq}^1 < \dots < \Theta_{ijq}^{|\mathcal{I}_{ijq}|}$ .
3    $[VD^{\text{lo}}, \mathbf{y}^{\text{rlx}}, \mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}}] \leftarrow \text{RELAXATION}(N, \alpha, F, \Phi, \mathcal{I}_{ijq})$ , (see function for definitions)
4    $[VD^{\text{up}}, \mathbf{x}^{\text{fsp}}] \leftarrow \text{LOCALSOLUTION}(N, \alpha, F, \Phi, \mathbf{y}^{\text{rlx}}, \mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}})$ , (see function for definitions)
5   if  $\left( \frac{VD^{\text{up}} - VD^{\text{lo}}}{VD^{\text{up}}} \right) < \epsilon_r$  then
6     The relative tolerance  $\epsilon_r$  is achieved.
7      $\mathbf{y} = \mathbf{y}^{\text{rlx}}$  and  $\mathbf{x} = \mathbf{x}^{\text{fsp}}$ 
8   else
9     For every  $[i, j] \in \mathcal{S}$ ,  $\llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1}$ , if split  $[i, k]/[l, j]$  is absent in  $\mathbf{y}^{\text{rlx}}$ , then  $\mathcal{I}_{ijq} \leftarrow \mathcal{I}_{ijq}$ .
     Otherwise,  $\llbracket q \rrbracket_{l-1}^k, \mathcal{I}_{ijq} \leftarrow \text{REFINEDISCRETIZATION}(\mathcal{I}_{ijq}, \boldsymbol{\theta}_{ijq}^{\text{rlx}})$ 
10    Go to Line 3
11  end
12 end
13 Function  $[VD^{\text{lo}}, \mathbf{y}^{\text{rlx}}, \mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}}] = \text{RELAXATION}(N, \alpha, F, \Phi, \mathcal{I}_{ijq})$ 
14  Construct relaxation (A)rlx: Formulate MINLP (A) described in §4. For  $[i, j] \in \mathcal{S}$ ,  $\llbracket k \rrbracket_i^{j-1}, \llbracket l \rrbracket_{i+1}^{k+1}, \llbracket q \rrbracket_{l-1}^k$ , reformulate Underwood constraints as described in §5.1, and construct
   piecewise relaxations of sets  $\mathcal{F}_{ikljqp}$ ,  $p = i, \dots, j$ , and  $\mathcal{V}_{ikljq}$  as described in §5.4.
15  Solve the resulting MI(L/SOC)P
16   $VD^{\text{lo}} \leftarrow$  Optimum objective function value
17   $[\mathbf{y}^{\text{rlx}}, \mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}}] \leftarrow$  Optimal values of discrete ( $\mathbf{y}^{\text{rlx}}$ ) and continuous ( $\mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}}$ ) decision
   variables. Vectors  $\mathbf{x}^{\text{rlx}}$  and  $\boldsymbol{\theta}^{\text{rlx}}$  contain optimal values of material flows and Underwood
   roots, respectively.
18 end
19 Function  $[VD^{\text{up}}, \mathbf{x}^{\text{fsp}}] = \text{LOCALSOLUTION}(N, \alpha, F, \Phi, \mathbf{y}^{\text{rlx}}, \mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}})$ 
20  Formulate MINLP (A) described in §4, and fix discrete decisions  $\mathbf{y} = \mathbf{y}^{\text{rlx}}$ .
21  Using  $(\mathbf{x}^{\text{rlx}}, \boldsymbol{\theta}^{\text{rlx}})$  as initial point, solve the resulting NLP using local solvers
22   $VD^{\text{up}} \leftarrow$  Optimum objective function value
23   $\mathbf{x}^{\text{fsp}} \leftarrow$  Optimal values of material flows
24 end
25 Function  $\text{REFINEDISCRETIZATION}(\mathcal{I}_{ijq}, \boldsymbol{\theta}_{ijq}^{\text{rlx}})$ 
26  Identify  $1 \leq t \leq |\mathcal{I}_{ijq}|$  such that  $\boldsymbol{\theta}_{ijq}^{\text{rlx}} \in [\boldsymbol{\theta}_{ijq}^{t-1}, \boldsymbol{\theta}_{ijq}^t]$ 
27  if  $(\boldsymbol{\theta}_{ijq}^{\text{rlx}} - \boldsymbol{\theta}_{ijq}^{t-1}) < \text{MinPrtSize}$  or  $(\boldsymbol{\theta}_{ijq}^t - \boldsymbol{\theta}_{ijq}^{\text{rlx}}) < \text{MinPrtSize}$  then
28     $\mathcal{I}_{ijq} \leftarrow \mathcal{I}_{ijq}$ 
29  else
30     $\mathcal{I}_{ijq} \leftarrow \{\mathcal{I}_{ijq} \setminus \{[\boldsymbol{\theta}_{ijq}^{t-1}, \boldsymbol{\theta}_{ijq}^t]\}\} \cup \{[\boldsymbol{\theta}_{ijq}^{t-1}, \boldsymbol{\theta}_{ijq}^{\text{rlx}}], [\boldsymbol{\theta}_{ijq}^{\text{rlx}}, \boldsymbol{\theta}_{ijq}^t]\}$ 
31  end
32  return  $\mathcal{I}_{ijq}$ 
33 end

```

Algorithm 2: Adaptive partitioning scheme to solve MINLP (A)

algorithm after processing the root node. We let $\theta_{ijq} \in [\alpha_{q+1} + \epsilon_\theta, \alpha_q - \epsilon_\theta]$, with $\epsilon_\theta = 10^{-7}$, for every $\llbracket q \rrbracket_i^{j-1}$, $[i, j] \in \mathcal{S}$ to avoid a possible division by zero. We use BARON's root node relaxation as a reference for comparison. We remark that BARON solves MIP relaxations as needed [40]. We also verified that the bound obtained is close to solving a factorable MIP relaxation.

Scenario 2 : (Simultaneous hull of fractional terms) This scenario illustrates the improvement in relaxation due to the use of simultaneous convexification techniques. We linearize all Underwood constraints in the MINLP by introducing auxiliary variables for each fraction. To relax fractional terms, we use (4.31), or (4.40) if zero is an admissible value for the range of the denominator of fractions. The nonlinear constraints in (4.31) are expressed as second-order cones, and the resulting Mixed Integer Second-order Cone Program (MISOCP) is solved with Gurobi 8.0 using Gurobi/MATLAB interface.

Scenario 3 : (RDLT with linear polynomials of θ) This scenario illustrates the improvement in relaxation due to reformulation of Underwood constraints using RDLT. We reformulate Underwood constraints as in (4.23), convexify fractional terms using (4.31) or (4.40), and convexify bilinear terms of the form $\Upsilon\theta = \Upsilon \cdot \theta$ using (4.32).

Scenario 4 : (RDLT with quadratic polynomials of θ) To the relaxation in Scenario 3, we add cuts derived by reformulating Underwood constraints with quadratic polynomials of θ (see (4.26)), as described in §4.3.1. This introduces additional nonlinear terms of the form $\Upsilon \cdot \theta^2$, which we relax in the manner described towards the end of §4.3.2.

Scenario 5 : (RDLT with inverse bound factors of θ) To the relaxation in Scenario 3, we add cuts derived by reformulating Underwood constraints with inverse bound factors (see (4.27)). This introduces additional nonlinear terms of the form f_1/θ and V/θ , which we relax in the manner described towards the end of §4.3.2.

Scenario 6 : (Discretization) Finally, to illustrate the potential of discretization, we construct piecewise relaxation of Scenario 3. We discretize the domain of each Underwood root into two partitions, and choose the roots of columns performing the split of the process feed, $\{\theta_{1Nq}\}_{q=1}^{N-1}$, as the partition points. In other words, we let $\mathcal{I}_{ijq} = \{[\alpha_{q+1}, \theta_{1Nq}], [\theta_{1Nq}, \alpha_q]\}$ for $i \leq q < j$ and $[i, j] \in \mathcal{S}$. As pointed out in Remark 4.4, these roots can be computed prior to solving the optimization problem. We construct the piecewise relaxation of MINLP (A) as outlined in §4.3.4.

Table 4.6 reports the percentage gap value, defined as

$$\% \text{ Gap} = 100 \times \left(1 - \frac{\text{Optimal value of relaxation}}{\text{Optimal value of (A)}} \right) \quad (4.43)$$

on a set of cases evaluated for all the Scenarios. To compare against BARON, we also report % gap closed (numbers in parenthesis in Table 4.6), defined as

$$\% \text{ Gap Closed} = 100 \times \left(1 - \frac{\text{Optimal value of (A)} - \text{Optimal value of relaxation}}{\text{Optimal value of (A)} - \text{Optimal value in Scenario 1}} \right) \quad (4.44)$$

We refer to a particular combination of parameter settings: N , $\{F_p\}_{p=1}^N$, $\{\alpha_p\}_{p=1}^N$, $\Phi_{1,N}$ and $\{\Phi_{p,p}\}_{p=1}^N$, as a *case*. The parameter settings for the cases considered in Table 4.6 are listed in the caption. It is worth noting that **Case-A** [17], **Case-B** and **Case-C** [46] correspond to physical mixtures: mixture of alcohols, mixture of light paraffins and mixture of light olefins and paraffins. The remaining cases do not directly correspond to physical mixtures, but are representative of specific classes of separations (see [39] for more details). Under Scenario 2, we report % Gap value, and % Gap closed for all cases when simultaneous hulls are used to convexify fractions. It can be observed that, this approach closes on an average 45.8% of the gap. In particular, in **Case-E**, implementation of simultaneous hull completely closes the gap at root node. Next, under Scenario 3, we report the combined effect of our RDLT approach and simultaneous hulls. This approach closes on an average 74.1% of the gap. Under Scenarios 4 and 5, we report further improvement in relaxation due to

addition RDLT cuts discussed in §4.3.1 to the relaxation in Scenario 3. RDLT cuts with quadratic polynomials of Underwood roots closes the gap completely in **Case-B**. Finally, the gap can be completely closed for all the cases considered in Table 4.6 by discretizing the domain of Underwood root into two partitions, as described in Scenario 6.

4.5 Computational Results

We conducted computational experiments on a *test set* of 496 cases, taken from [35, 39], which is a representative of a majority of separations. Parameter settings for the test set are listed in §A in e-companion. In this section, we demonstrate that our proposed approach is able to solve MINLP (A) within a relative tolerance of 1%. We also compare the performance of our approach with prior approaches in the literature [17, 22, 23]. Since the prior approaches develop an (MI)NLP model, we use BARON 18.5.8 via GAMS 25.1 to solve these (MI)NLPs, where all BARON options are set at their default values. For the adaptive partitioning scheme described in Algorithm 1, we use Gurobi 8.0 [49] to solve the resulting MIPs, and use IPOPT [57] as a local solver. The model is loaded into Gurobi using the MATLAB/Gurobi interface, while IPOPT is used via MATLAB/GAMS interface and GAMS 25.1. We used single CPU thread to solve the MIPs so as to keep the comparison with BARON fair. Besides the setting of number of threads, the remaining options for Gurobi and IPOPT were left at their defaults. All computations were done on a Dell Optiplex 5040 with Intel Core i7-6700 3.4 GHz processor and 16 GB RAM, and is running 64-bit Windows 7.

Table 4.6.

Variation of duality gap across the scenarios described in §4.4. Here, a gap value less than $10^{-4}\%$ is marked as 0%. For all the cases, $N = 5$, $\Phi_{1,N} = \Phi_{1,1} = \dots = \Phi_{N,N} = 1$. In **Case-A**, $F = \{20, 30, 20, 20, 10\}$ and $\alpha = \{4.1, 3.6, 2.1, 1.42, 1\}$; In **Case-B**, $F = \{5, 15, 25, 20, 35\}$ and $\alpha = \{7.98, 3.99, 3, 1.25, 1\}$; In **Case-C**, $F = \{25, 10, 25, 20, 20\}$ and $\alpha = \{13.72, 3.92, 3.267, 1.21, 1\}$; In **Case-D**, $F = \{42.5, 42.5, 5, 5\}$ and $\alpha = \{3.3275, 3.025, 1.21, 1.1, 1\}$; In **Case-E**, $F = \{30, 30, 5, 5, 30\}$ and $\alpha = \{1.4641, 1.331, 1.21, 1.1, 1\}$; In **Case-F**, $F = \{5, 5, 5, 42.5, 42.5\}$ and $\alpha = \{3.3275, 1.331, 1.21, 1.1, 1\}$; In **Case-G**, $F = \{5, 5, 5, 42.5, 42.5\}$ and $\alpha = \{17.1875, 6.875, 2.75, 1.1, 1\}$; In **Case-H**, $F = \{20, 20, 20, 20, 20\}$ and $\alpha = \{7.5625, 3.025, 1.21, 1.1, 1\}$. Data for **Case-A** is taken from [17], **Case-B** and **Case-C** from [46], and **Case-D** through **Case-H** from [39].

% Gap as defined in (4.43) (% Reduced Gap as defined in (4.44))						
	Optimum	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Case-A	402.7	31.2%	27% (13.5%)	19.6% (37.2%)	13.9% (55.4)	15.5% (50.3%)
Case-B	272.5	38.3%	13.7% (64.2%)	3.3% (91.4%)	0% (100%)	0.1% (99.7%)
Case-C	260	25.7%	17.7% (31.1%)	6% (76.7%)	1% (96.1%)	1.1% (95.7%)
Case-D	896.4	45%	20.4% (54.7%)	14.9% (66.9%)	7.8% (82.7%)	8.4% (81.3%)
Case-E	695.6	27.7%	0% (100%)	0% (100%)	0% (100%)	0% (100%)
Case-F	929.1	32.5%	21.4% (34.2%)	4.4% (86.5%)	2.4% (92.6%)	4% (87.7%)
Case-G	902.7	45.8%	22.7% (50.4%)	7.4% (83.8%)	1.2% (97.4%)	3.2% (93%)
Case-H	542	27.8%	22.7% (18.3%)	13.8% (50.4%)	3.8% (86.3%)	4.4% (84.2%)
Average Gap Closed			45.8%	74.1%	88.8%	86.5%
						100%

4.5.1 Comparison with Prior Approaches

Here, we compare the performance of three approaches, namely those of [9, 23], and the one proposed here. For all the computations, we set the relative tolerance for convergence (ϵ_r), defined as

$$\epsilon_r = \left(1 - \frac{\text{BestLB}}{\text{BestUB}}\right) \quad (4.45)$$

where BestLB and BestUB are the best-known relaxation bound and feasible solution, to 1% i.e., $\epsilon_r = 0.01$. We impose a CPU time limit of five hours as the termination criterion.

Approach 1 : We solve MINLP (A) using the adaptive partitioning approach described in Algorithm 2. We begin with four partitions for each Underwood root *i.e.*, $\mathcal{I}_{ijq} = \{[\alpha_{q+1}, (\alpha_{q+1} + \theta_{1Nq})/2], [(\alpha_{q+1} + \theta_{1Nq})/2, \theta_{1Nq}], [\theta_{1Nq}, (\alpha_q + \theta_{1Nq})/2], [(\theta_{1Nq} + \alpha_q)/2, \alpha_q]\}$ for every $\llbracket q \rrbracket_i^{j-1}$, $[i, j] \in \mathcal{S}$. We compute the Underwood roots for the splits of the process feed $\{\theta_{1Nq}\}_{q=1}^{N-1}$ prior to solving the MINLP (see Remark 4.4). For all but 4 cases, we set `MinPrtSize` = 10^{-3} . For the remaining cases, we reduced `MinPrtSize` to 10^{-4} in order to achieve the relative tolerance of 1%. Finally, we point out that the upper bounds on material flows are computed by solving (4.15), where we choose

$$VD^* = \max_{q \in \{1, \dots, N-1\}} \sum_{p=1}^q \frac{\alpha_p F_p}{\alpha_p - \theta_{1Nq}} \quad (4.46)$$

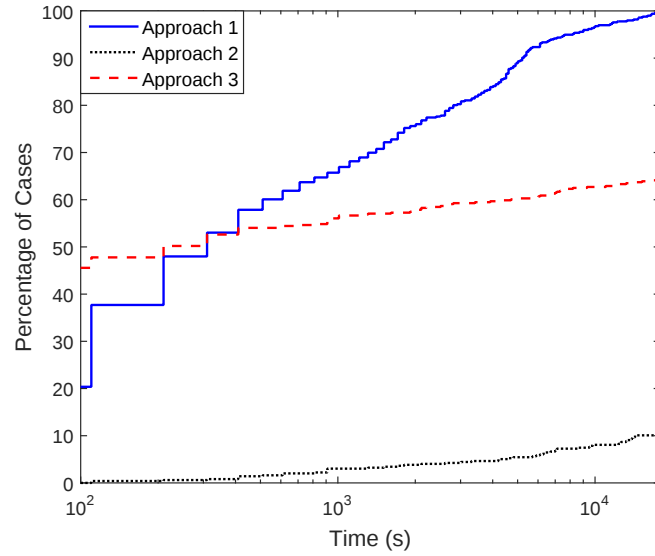
and $\phi = 1.5$. We note that (4.46) is the objective function value corresponding to a feasible point of one of the admissible configurations, commonly known in literature as Fully Thermally Coupled or Petlyuk configuration (see [14, 15]).

Approach 2 : We obtained the GAMS code of the model proposed in [9] from the MINLP library [24]. There, the authors were interested in identifying the configuration minimizing the total annual cost. For our computations, we modify their code in the following manner. First, as mentioned in [23], the model of [24] admits solutions that are physically infeasible. This is because the constraints

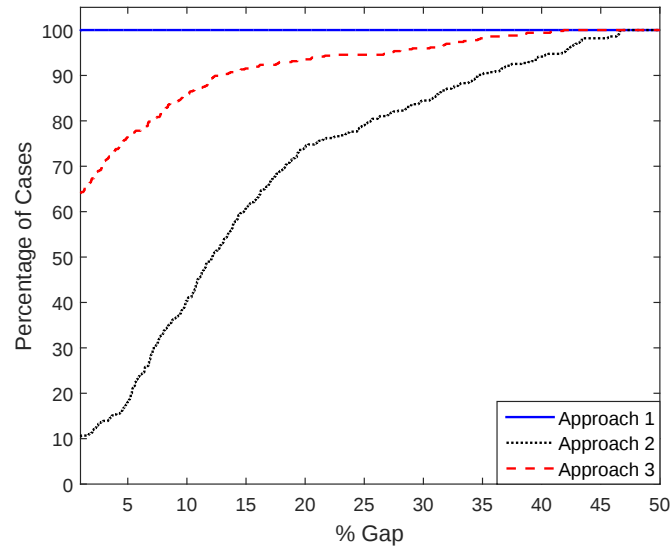
corresponding to (A19) in their model should be tight for certain Underwood roots, and their model does not impose this requirement. We have added these missing constraints to their GAMS code. Second, the authors employed the BigM approach in order to transform certain disjunctions into a set of inequalities. Unfortunately, the BigM value used for vapor and liquid bypass in their GAMS code made a few test cases infeasible. Therefore, we specified $2.5 VD^*$ as the BigM value for the vapor and liquid bypasses. This number was found by choosing the smallest BigM value for which we found a feasible solution. Third, the authors use a parameter ϵ_θ and restrict $\theta_{ijq} \in [\alpha_{q+1} + \epsilon_\theta, \alpha_q - \epsilon_\theta]$ for $[[q]]_i^{j-1}$, $[i, j] \in \mathcal{P}$ in order to avoid the singularity associated with θ_{ijq} approaching α_q or α_{q+1} . Their choice of ϵ_θ , in some cases, made the optimal solution infeasible. Empirically, we found that $\epsilon_\theta = 10^{-4}$ does not cut off the optimal solution, so we set $\epsilon_\theta = 10^{-4}$. Fourth, the cost equations required for the evaluation of the objective function were removed from the model, and the objective function was modified to compute the total vapor duty instead. The resulting MINLP is then solved with BARON.

Approach 3 : Here, we consider the MINLP proposed in [23]. For a consistent comparison, we set the upper bound on all vapor flows to be $1.5 VD^*$. Further, we restrict $\theta_{ijq} \in [\alpha_{q+1} + \epsilon_\theta, \alpha_q - \epsilon_\theta]$, where $\epsilon_\theta = 10^{-4}$, for $[[q]]_i^{j-1}$, $[i, j] \in \mathcal{S}$ in order to avoid the singularity associated with θ_{ijq} approaching α_q or α_{q+1} . The resulting MINLP is then solved using BARON.

Figure 4.7(a) shows the percentage of cases solved to 1%-optimality against time, with Approach 1 (solid blue curve), Approach 2 (dotted black curve), and Approach 3 (dashed red curve). Observe that Approach 2 solves about 10% of cases to 1%-optimality within five hours. This is not surprising because [9, 17] also reported difficulties in convergence. To overcome the challenges, the authors architected an algorithm by modifying logic-based outer-approximation. While the method resulted



(a)



(b)

Fig. 4.7. (a) Plot showing percentage of cases solved to 1%-optimality against time. Here, Approach 1 corresponds to the current work, Approach 2 corresponds to the model proposed in [9] solved with BARON, after making the changes described in §4.5, and Approach 3 corresponds to the model proposed in [23] solved with BARON (b) Plot showing the remaining duality Gap at the end of five hours for all the three approaches.

in good solutions, optimality was not guaranteed. Approach 3 solves 64% of the cases in the test set.

We remark that [23] introduced a new search-space formulation, derived cuts that exploit monotonicity of Underwood constraints, and modeled the absence/presence of a column using disjunctions. Nevertheless, this approach fails to solve the problem to 1%-optimality for 36% of the cases. The progress of lower bound for a majority of these cases is either stagnant or very slow. Figure 4.7(b) depicts the cumulative percentage of cases as a function of the remaining duality gap at the end of five hours. In contrast, our approach, for the first time, solves all 496 cases from this test set within an optimality tolerance of 1%.

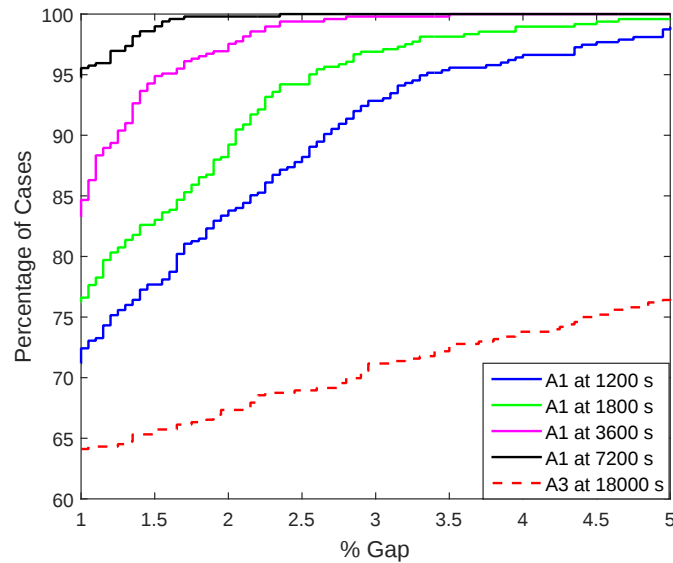


Fig. 4.8. Profiles showing remaining % Gap at the end of specific time instances for Approach 1 (A1) and Approach 3 (A3).

Figure 4.8 depicts cumulative percentage of cases as a function of the remaining duality gap at specific time instances for Approach 1. This graph demonstrates that our solution approach, with a CPU time of twenty minutes, already outperforms the best prior MINLP based approach allowed to run for a CPU time of five hours. Further, within 1800 s (green curve), 3600s (magenta curve) and 7200s (black curve),

the proposed approach solves all 496 cases to less than 5.5%, 3.5% and 2.5% gap, respectively. Since (A) is primarily designed as a screening tool for an otherwise highly cumbersome search of optimal distillation configuration, practicing engineers can use Approach 1 to quickly identify near optimal solutions that are worthy of further exploration. Although we do not provide specific configurations found using our procedure, the potential benefits are documented in [5, 23] for a crude distillation case study.

4.5.2 Comparison with Nallasivam et al. [22]

Recently, [22] proposed an alternative technique that relies on explicit enumeration for identifying distillation configuration requiring the least vapor duty. After enumerating all the configurations, an NLP is formulated for each configuration and solved to 1%–optimality with BARON. We refer to this as **Approach 4**. We compare the performance of Approach 4, with Approaches 1 and 3 by fixing the discrete decisions to a specific configuration. We choose Fully Thermally Coupled (FTC) configuration, characterized by $\zeta_{i,j} = 1 \ \forall [i, j] \in \mathcal{T}$, $\chi_{i,j} = 0 \ \forall (i, j) \in \mathcal{C} \setminus \{(1, 1)\}$, $\chi_{1,1} = 1$, $\rho_{i,j} = 0 \ \forall (i, j) \in \mathcal{R} \setminus \{(N, N)\}$, and $\rho_{N,N} = 1$, for comparison. This comparison ignores the advances in the search space formulation discussed in §4.2.2 and other advances that relate Underwood constraints with stream variables, since we fix the binary variables a priori. We set the time limit as one hour and a relative gap of 1% ($\epsilon_r = 0.01$) as termination criteria.

Figure 4.9 depicts the percentage of cases solved as a function of computational time for the three approaches. Clearly, BARON solves more number of cases to 1%–optimality with Approach 3 than with Approach 4. Despite the improvement, only 82% of the cases are solved to 1%–optimality using Approach 3. In contrast, our approach solves all cases in this test set within 100 s.

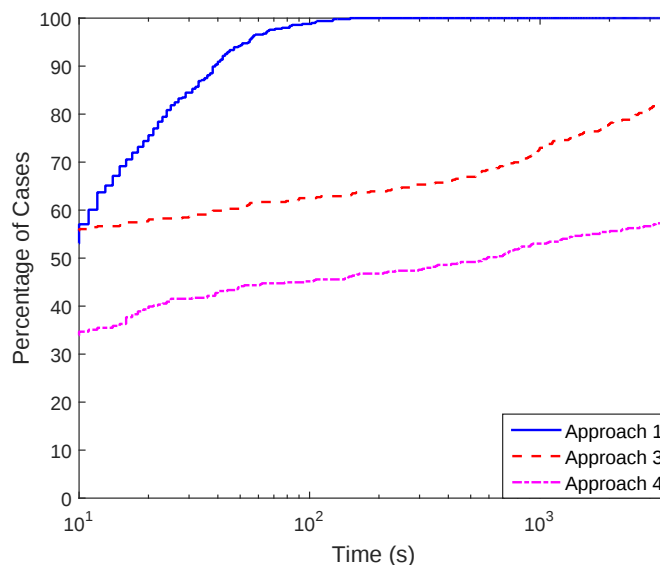


Fig. 4.9. Plot showing percentage of cases solved to 1%–optimum against time, when discrete variables are fixed to fully thermally coupled configuration (see §4.5.2). Approach 4 corresponds to the model proposed in [22] solved with BARON.

4.6 Concluding Remarks

This work addressed the optimal design of distillation configurations, which are widely used in all chemical and petrochemical industries, and are significant consumers of energy in the world economy. We proposed a novel MINLP that identifies energy-efficient configurations for a given application. Given the complexity from combinatorial explosion of the choice set and nonconvex Underwood constraints, this problem has resisted solution approaches. In this paper, we report on the first successful approach and solve this problem to global optimality for five-component mixtures. The key contributions that make this possible are (i) new formulation for discrete choices that is strictly tighter than the previous formulations, (ii) new valid cuts to the problem using RDLT, and various other convexification results for special structures, and (iii) discretization techniques and an adaptive partitioning scheme to solve the MINLP to ϵ –optimality. On a test set that is a representative of a majority of five-component separations, we demonstrated that our approach solves all

the instances in a reasonable amount of time, which was not possible using existing approaches. In summary, this paper describes the first solution approach that can reliably and quickly screen several thousands of alternative distillation configurations and identify solutions that consume less energy and, thereby, lead to less greenhouse gas emissions. This approach has the potential to reduce the carbon footprint and energy usage of thermal separation processes.

5. MINIMIZATION OF EXERGY LOSS

Process designers may use several criteria to filter configurations from the *search space*. In the last two chapters, we have shown how to identify the distillation configuration that requires least vapor vapor duty. In this chapter, we extend the procedure for identifying the configuration maximizing the thermodynamic efficiency (η), defined as

$$\eta = \frac{\text{Minimum work of separation}}{\text{Total work of separation}} \quad (5.1)$$

Since all configurations in the search space perform the same task of separating a multicomponent mixture into its constituent components, the minimum work of separation is the same for all configurations. Thus, configurations with the highest thermodynamic efficiency require the least total or net work for separation. Thermodynamic analysis is crucial for work-driven distillations like cryogenic/sub-ambient separations and above-ambient distillations employing heat pumps (vapor recompression cycle, for example). [28] elucidate how thermodynamic analysis provides valuable insights for synthesis of separation systems. For example, in cases where economic models optimizing either operating cost (minimization of heat duty as in [22, 23]) or total annualized cost (annualized CAPEX+OPEX as in [9, 58]) yield distinct configurations having comparable objective function values, thermodynamic analysis may be supplemented to further discriminate such configurations. The other instance where thermodynamic analysis may be useful is when extensive heat integration is desired, both within the separation unit and with the rest of the plant. While it is possible to formulate a model with extensive heat integration, it may prove computationally challenging to solve the model. Since the method in this article is intended as a screening step, thermodynamic analysis considered here may be faster and it yields a model that is relatively easier to solve.

Now, consider the definition of thermodynamic efficiency in (5.1). The total work required for separation is obtained as

$$\text{Total work of separation} = \text{Minimum work of separation} + \text{Exergy loss.} \quad (5.2)$$

Exergy of a stream is defined as its *work potential i.e.*, the maximum work that can be extracted when a stream is brought from its current state to thermal, mechanical and chemical equilibrium with the environment via a reversible path. *Exergy loss* of a system corresponds to the loss of work potential due to irreversibilities in the system. We refer readers to [59] for a lucid description of exergy concepts and applications. From (5.2), a configuration with higher irreversibilities has higher exergy losses; thereby requiring higher total work for separation. This makes the overall separation process less efficient, thermodynamically. In contrast, a configuration with lower irreversibilities requires lower total work, making it thermodynamically more efficient. Thus, the problem of identifying thermodynamically efficient configurations is equivalent to the problem of identifying configurations with least exergy losses. Since both problems are equivalent, we choose the latter problem as it has a simpler objective function.

While the formulation in Chapter 3 identifies configurations requiring least heat duty (*first law savings*), it does not take into account the temperature levels of streams. As a result, it may lead to a solution that is attractive in terms of energy consumption, but may require utilities at extreme temperatures (*second law penalty*). Such solutions are not thermodynamically efficient. Therefore, in this work, we build upon the formulation introduced in Chapter 3 to account for temperature levels of streams in order to identify exergetically-efficient configurations.

Here, we review the studies which used thermodynamic analysis to obtain insights on improving the efficiency of the separation process. Use of thermodynamic analysis in the context of multicomponent distillation may be traced back to as early as 1960s. [60] introduced a novel configuration, which is now commonly known as Fully Thermally Coupled (FTC) configuration, and used thermodynamic arguments to support their design. Subsequent works include [28, 29, 61–67]. In particular, we

highlight that [28, 63, 65] independently derived a relation relating temperature as a function of liquid mole fraction and relative volatilities of constituent components, for a multicomponent mixture using the Clausius-Claypeyron equation. Using this relation, all the aforementioned authors independently derived an expression for net work/thermodynamic efficiency/exergy loss that does not depend on temperature explicitly. Such results hold for both sub-ambient and above-ambient operation, provided the mixture has the same composition and relative volatilities for constituent components. However, [28] focused only on sharp-split configurations. Though their equations also hold for sloppy split configurations, a global optimization framework, such as the one described here, is needed [35]. On the other hand, [29, 62–65] considered only binary and ternary mixtures. To the best of our knowledge, a framework that screens through all regular-column configurations, and identifies exergetically-efficient configurations for a general N -component zeotropic mixture was not unavailable until recently. [53] developed an explicit enumeration approach that formulates a Nonlinear Program (NLP), with minimization of exergy loss as the objective function, for each configuration in the search space. NLPs are then solved to ϵ -global optimality using BARON [21, 40]. The configurations are then ranklisted in ascending order of the minimum exergy loss, to identify a handful of configurations attractive for the given application. The contributions in this chapter are closely related to [53], and they differ from the latter in the following aspects. Explicit enumeration, like the one proposed by [53], can be computationally intensive for mixtures containing five or more components. Whereas, the current approach is computationally more efficient as it enumerates implicitly and avoids solving every configuration in the search space. Further, this work proposes a novel reformulation of equations for exergy loss calculation. This reformulated model requires fewer number of nonlinear nonconvex equations compared to the model used in [53].

The rest of the chapter is organized as follows. In §5.1, we formulate the MINLP. In §5.2, we consider a case study concerning the recovery of Natural Gas Liquids (NGLs) from shale gas. In addition, we also consider several examples from literature

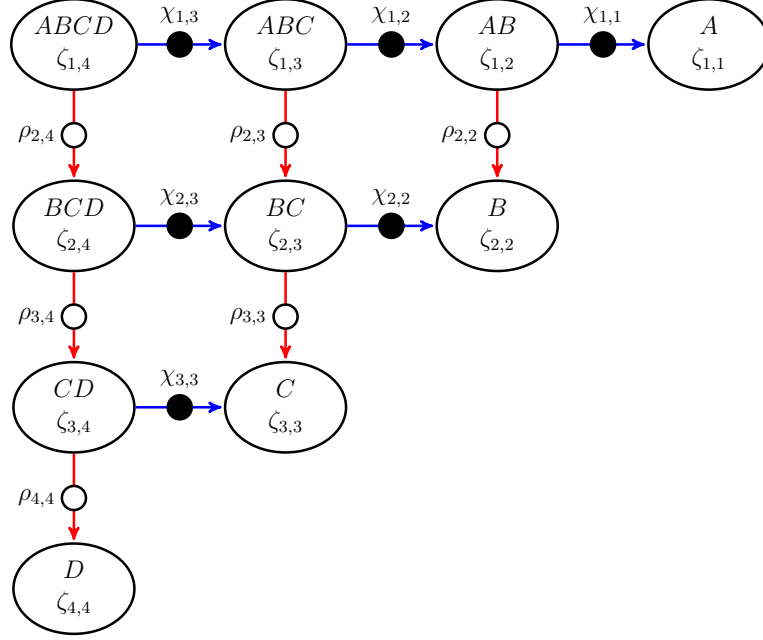


Fig. 5.1. A four-component system

to investigate the solution performance to changes in process parameters. Finally, we conclude the chapter in §5.3.

5.1 Problem Formulation

Figure 5.1 shows a four-component system ($N = 4$). We use the same notation introduced in Chapter 4.

The problem definition can be briefly stated as follows. Given an N -component mixture along with the composition of the *process feed* $\{F_p\}_{p=1}^N$ and component relative volatilities $\{\alpha_p\}_{p=1}^N$ (measured w.r.t the least volatile component *i.e.*, $\alpha_1 > \dots > \alpha_n$), identify the optimal distillation configuration and its operating conditions minimizing the total exergy loss ($\Delta\mathcal{E}_{\text{loss}}$) for separating the given mixture into its constituent components. Here, and in the rest of the chapter, $\{p\}_{p=1}^n$ is a shorthand notation for the set $\{1, \dots, n\}$.

For better readability, we begin the formulation (W) with the problem constraints, and defer the description of the objective function to the end. Table 5.1 shows the definition of relevant sets required for the problem formulation.

Table 5.1.
Definition of sets.

Set	Symbol	Definition
Streams	$\mathcal{T}$	$\{[i, j] : 1 \leq i \leq j \leq N\}$
Splits	$\mathcal{P}$	$\mathcal{T} - \{[i, i] : 1 \leq i \leq N\}$
Submixtures	$\mathcal{S}$	$\mathcal{P} - \{[1, N]\}$
Condensers	$\mathcal{C}$	$\{(i, j) : 1 \leq i \leq j \leq N - 1\}$
Reboilers	$\mathcal{R}$	$\{(i, j) : 2 \leq i \leq j \leq N\}$
Quadrature points	$\mathcal{G}$	$\{1, 2\}$

5.1.1 Space of Admissible Distillation Configurations

We first describe the constraints to model the space of all regular-column configurations (both basic and thermally coupled configurations). We define binary variables $\zeta_{i,j} = \{1, \text{ if } [i, j] \text{ is present; } 0, \text{ otherwise}\}$ for $[i, j] \in \mathcal{T}$, $\chi_{i,j} = \{1, \text{ if condenser } (i, j) \text{ is present; } 0, \text{ otherwise}\}$ for $(i, j) \in \mathcal{C}$, and $\rho_{i,j} = \{1, \text{ if reboiler } (i, j) \text{ is present; } 0, \text{ otherwise}\}$ for $(i, j) \in \mathcal{R}$. In addition, we define split variables $\sigma_{i,k,l,j}$, such that $\sigma_{i,k,l,j} = 1$ if $[i, k]$ and $[l, j]$ are produced as distillate and residue from $[i, j]$, and $\sigma_{i,k,l,j} = 0$ otherwise. Hereafter, for conciseness, we write $[i, k]/[l, j]$ to represent that split of stream $[i, j]$ produces $[i, k]$ and $[l, j]$ as distillate and residue. We note that, for a given distillate stream $[i, k]$ from $[i, j]$, $i \leq k \leq j - 1$, only the splits in $\{[i, k]/[l, j]\}_{l=i+1}^{k+1}$ are feasible. Whereas, the remaining splits $\{[i, k]/[l, j]\}_{l=k+2}^j$ are infeasible, because one or more components are not conserved in these splits. Similarly, for a given residue stream $[l, j]$ from $[i, j]$, $i + 1 \leq l \leq j$, only the splits in $\{[i, k]/[l, j]\}_{k=l-1}^{j-1}$ are feasible. We

define split variables only for feasible splits. Thus, the domain of indices for split variables is given by $\bigcup_{k=i}^{j-1} \{\sigma_{i,k,l,j}\}_{l=i+1}^{k+1}$ or $\bigcup_{l=i+1}^j \{\sigma_{i,k,l,j}\}_{k=l-1}^{j-1}$ for $[i, j] \in \mathcal{P}$.

We borrow constraints (A2) through (A8) in Chapter 4 for modeling the space of admissible configurations, and substitute out the auxiliary variables $(\tau, \beta, \nu, \omega)$ to obtain the formulation in the space of $(\zeta, \sigma, \rho, \chi)$ variables. The variable elimination process is described in the supplementary information. The resulting constraints are described below.

$$\zeta_{1,N} = \zeta_{1,1} = \dots = \zeta_{N,N} = 1, \quad (\text{W1})$$

$$\sum_{k=i}^{j-1} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} = \zeta_{i,j} \quad \forall \quad [i, j] \in \mathcal{P}, \quad (\text{W2})$$

$$\zeta_{i,j} \leq \sum_{l=j+1}^N \sum_{k=i+1}^{j+1} \sigma_{i,j,k,l} + \sum_{k=1}^{i-1} \sum_{l=i-1}^{j-1} \sigma_{k,l,i,j} \quad \forall \quad [i, j] \in \mathcal{T} \setminus \{[1, N]\}, \quad (\text{W3})$$

$$\left. \begin{aligned} \chi_{i,j} &\leq \zeta_{i,j} - \sum_{k=1}^{i-1} \sum_{l=i-1}^{j-1} \sigma_{k,l,i,j}, \quad \forall \quad (i, j) \in \mathcal{C} \\ \rho_{i,j} &\leq \zeta_{i,j} - \sum_{l=j+1}^N \sum_{k=i+1}^{j+1} \sigma_{i,j,k,l}, \quad \forall \quad (i, j) \in \mathcal{R} \end{aligned} \right\}, \quad (\text{W4})$$

$$\left. \begin{aligned} \chi_{i,i} &\geq \zeta_{i,i} - \sum_{k=1}^{i-1} \sigma_{k,i-1,i,i} \quad \forall \quad (i, i) \in \mathcal{C} \\ \rho_{i,i} &\geq \zeta_{i,i} - \sum_{l=j+1}^N \sigma_{i,i,i+1,l} \quad \forall \quad (i, i) \in \mathcal{R} \end{aligned} \right\}, \quad (\text{W5})$$

$$\left. \begin{aligned} \sum_{s=n}^N \sum_{r=i}^{n-1} \sum_{m=i+1}^{r+1} \sigma_{i,r,m,s} &\leq \sum_{s=k}^N \sum_{r=i}^{k-1} \sum_{m=i+1}^{r+1} \sigma_{i,r,m,s} \leq 1, \quad \llbracket n \rrbracket_{k+1}^N, \llbracket k \rrbracket_{i+1}^{N-1} \\ \sum_{r=1}^m \sum_{s=m+1}^j \sum_{n=s-1}^{j-1} \sigma_{i,n,s,j} &\leq \sum_{r=1}^l \sum_{s=l+1}^j \sum_{n=s-1}^{j-1} \sigma_{i,n,s,j} \leq 1, \quad \llbracket m \rrbracket_1^{l-1}, \llbracket l \rrbracket_2^{j-1} \end{aligned} \right\}. \quad (\text{W6})$$

Here, $\llbracket n \rrbracket_{k+1}^N$ is a shorthand notation for $\forall n \in \{k+1, \dots, N\}$. (W1) ensures that the process feed and the pure components are present in a configuration. (W2) ensures that if a submixture $[i, j]$ is produced in a configuration, it undergoes a split. (W3) ensures that every stream except the process feed stream is produced from at least

one split. (W4) and (W5) model the presence/absence of heat exchangers. (W6) ensures that only regular column configurations are present in the search space.

Proposition 5.1. *Let $S = \{(\zeta, \sigma, \chi, \rho, \tau, \beta) \in \text{Unit hypercube} \mid (\text{W1}) - (\text{W6}), \tau_{i,k,j} = \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j}, \beta_{i,l,j} = \sum_{k=l-1}^{j-1} \sigma_{i,k,l,j}, \forall [i,j] \in \mathcal{P}\}$ be an affine lift of (W1) – (W6). Then, S is tighter than the intersection of the search space formulations in (i) [9], (ii) [10], and (iii) [23].*

Proof. See proof of Proposition 7 in Chapter 4. □

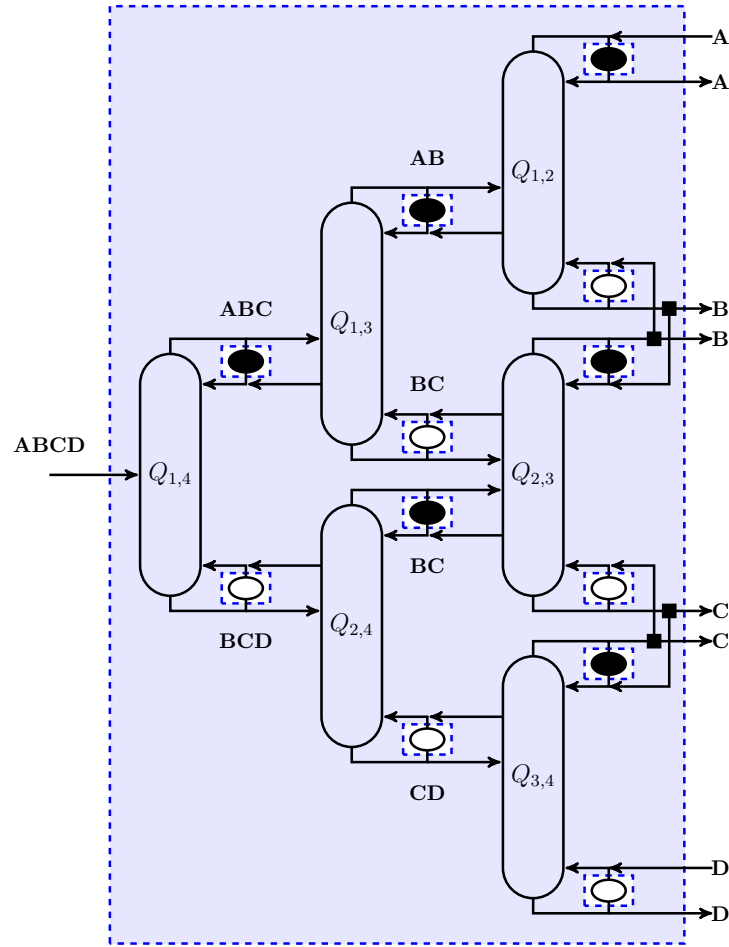
5.1.2 Mass Balance Constraints

Figures 5.2(a) and 5.2(b) show the superstructure for a four-component system and schematic of a representative column. The mass balances are the same as those described in Chapter 3:

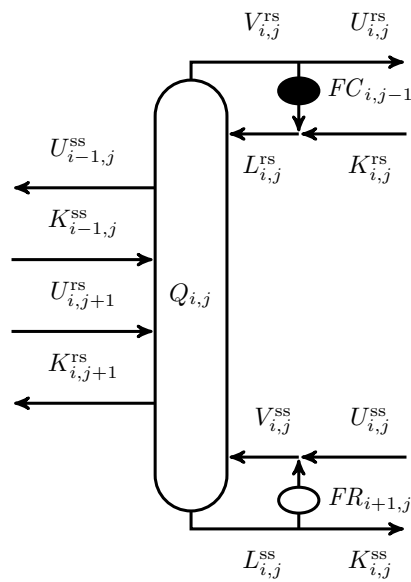
$$(\text{H12}) - (\text{H17}) \tag{W7}$$

The superstructure enables modeling of single/two phase heat exchangers, thermal coupling and a combination of thermal coupling and heat exchanger. In this chapter, we restrict that either a two-phase heat exchanger or a thermal coupling can be present, but not both. Further, we model two-phase heat exchangers as shown in Figure 5.3.

For a condenser, we split the inlet vapor stream into two; such that one stream passes through the condenser to supply the required liquid reflux and the liquid portion of the feed to the subsequent column, while the other stream supplies the vapor portion of the feed to the subsequent column. In a similar manner, for a reboiler, we split the inlet liquid stream into two; such that one stream passes through the reboiler to supply the required vapor reflux and the vapor portion of the feed to the subsequent column, while the other stream supplies liquid portion of the feed to the subsequent column. Note that the vapor and liquid streams will not be in equilibrium



(a)



(b)

Fig. 5.2. (a) Four-component superstructure and (b) A representative column in the superstructure. Reproduced from [23].

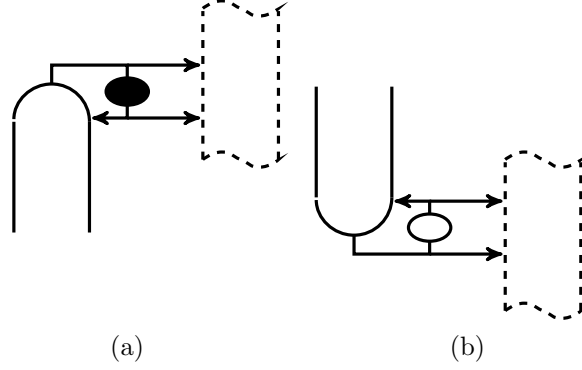


Fig. 5.3. (a) Two-phase condenser (b) Two-phase reboiler

unlike in conventional two-phase heat exchangers. Though this alternate approach of modeling two-phase heat exchangers does not affect minimum vapor requirement calculations, it affects exergy calculations. This is because, the temperature variation in each condenser (*resp.* reboiler) goes from dew point to all the way to bubble point (*resp.* bubble point to dew point). Whereas in a conventional two-phase condenser (*resp.* reboiler), the temperature is fixed at the vapor-liquid equilibrium temperature which is higher (*resp.* lower) than the bubble point (*resp.* dew point). Despite this limitation, we choose to model two-phase exchangers in this manner to reduce the complexity (number of nonlinear nonconvex equations) of the model. We believe that this simplification does not change the ranklist of configurations significantly, and thus, it can be used for first-step screening. The required constraints are modeled as

$$\left. \begin{aligned} 0 &\leq FC_{i,j} \leq (FC_{i,j})^{\text{up}} \chi_{i,j} \\ (K_{i,j+1}^{\text{rs}})^{\text{lo}} \chi_{i,j} &\leq K_{i,j+1}^{\text{rs}} \leq (K_{i,j+1}^{\text{rs}})^{\text{up}} (1 - \chi_{i,j}) \end{aligned} \right\} \quad \forall \quad (i,j) \in \mathcal{C}, \quad (\text{W8})$$

$$\left. \begin{aligned} 0 &\leq FR_{i,j} \leq (FR_{i,j})^{\text{up}} \rho_{i,j} \\ (U_{i-1,j}^{\text{ss}})^{\text{lo}} \rho_{i,j} &\leq U_{i-1,j}^{\text{ss}} \leq (U_{i-1,j}^{\text{ss}})^{\text{up}} (1 - \rho_{i,j}) \end{aligned} \right\} \quad \forall \quad (i,j) \in \mathcal{R}. \quad (\text{W9})$$

Here, $(\cdot)^{\text{up}}$ and $(\cdot)^{\text{lo}}$ represent upper and lower bounds on $(\cdot)$, respectively. Observe that, when condenser (*resp.* reboiler) is absent *i.e.*, $\chi_{i,j} = 0$ (*resp.* $\rho_{i,j} = 0$), material flow into the condenser $FC_{i,j}$ (*resp.* reboiler $FR_{i,j}$) is forced to zero and $K_{i,j+1}^{\text{rs}} \geq 0$ (*resp.* $U_{i-1,j}^{\text{ss}} \geq 0$). This scenario models thermal coupling. On the other hand, the

scenario where $\chi_{i,j} = 1$ (*resp.* $\rho_{i,j} = 1$) models a two-phase condenser (*resp.* reboiler) by enforcing $K_{i,j+1}^{\text{rs}} \leq 0$ (*resp.* $U_{i-1,j}^{\text{ss}} \leq 0$). When condenser (i, j) (*resp.* reboiler (i, j)) is present, the maximum liquid (*resp.* vapor) flowrate to the pseudocolumn $Q_{i,j}$ from $Q_{i,j+1}$ (*resp.* $Q_{i,j-1}$) is $\sum_{p=i}^j F_p$. Therefore, we choose $(K_{i,j+1}^{\text{rs}})^{\text{lo}} = -\sum_{p=i}^j F_p$ and $(U_{i-1,j}^{\text{ss}})^{\text{lo}} = -\sum_{p=i}^j F_p$. The choice of upper bounds is described in Chapter 3.

5.1.3 Underwood Constraints

Consider a pseudocolumn $Q_{i,j}$ in the superstructure. It performs the separation of stream $[i, j]$, when $[i, j]$ is present. The resulting component distribution, vapor and liquid flows are governed by Underwood constraints. On the other hand, when $[i, j]$ is absent, pseudocolumn $Q_{i,j}$ bypasses material flows from the rectifying (*resp.* stripping) section of the parent pseudocolumn $Q_{i,j+1}$ (*resp.* $Q_{i,j-1}$) to the rectifying (*resp.* stripping) section of $Q_{i,j}$. The relevant constraints are expressed as a disjunction, which is transformed to a set of equalities and inequalities using disjunctive programming techniques. We omit the details for brevity, and refer the reader to Chapter 3. For every $[i, j] \in \mathcal{P}$, we impose the following constraints:

$$\left\{ \begin{array}{l} \sum_{p=i}^j \frac{\alpha_p (f_{i,j+1,p}^{\text{rsI}} + f_{i-1,j,p}^{\text{ssI}})}{\alpha_p - \theta_{i,j,q}} = U_{i,j+1}^{\text{rsz}} - U_{i-1,j}^{\text{ssz}} \\ \sum_{p=i}^{j-1} \frac{\alpha_p f_{i,j,p}^{\text{rsO}}}{\alpha_p - \theta_{i,j,q}} \leq \Upsilon_{i,j}^{\text{rsz}} \\ - \sum_{p=i+1}^j \frac{\alpha_p f_{i,j,p}^{\text{ssO}}}{\alpha_p - \theta_{i,j,q}} \leq \Upsilon_{i,j}^{\text{ssz}} \\ - \sum_{p=i}^{j-1} \frac{\alpha_p f_{i,j,p}^{\text{rsO}}}{\alpha_p - \theta_{i,j,q}} + \Upsilon_{i,j}^{\text{rsz}} \leq M_{ij1} \left(\zeta_{i,j} - \sum_{k=q+1}^{j-1} \sum_{l=i+1}^q \sigma_{i,k,l,j} \right) \\ \sum_{p=i+1}^j \frac{\alpha_p f_{i,j,p}^{\text{ssO}}}{\alpha_p - \theta_{i,j,q}} + \Upsilon_{i,j}^{\text{ssz}} \leq M_{ij2} \left(\zeta_{i,j} - \sum_{k=q+1}^{j-1} \sum_{l=i+1}^q \sigma_{i,k,l,j} \right) \end{array} \right\}_{q=i}^{j-1} \quad (\text{W10})$$

$$\Upsilon_{i,j}^{\text{rsz}} \leq V_{i,j}^{\text{rsz}}, \quad \Upsilon_{i,j}^{\text{ssz}} \leq V_{i,j}^{\text{ssz}} \quad (\text{W11})$$

$$U_{i,j+1}^{\text{rsz}} - U_{i-1,j}^{\text{ssz}} = \Upsilon_{i,j}^{\text{rsz}} - \Upsilon_{i,j}^{\text{ssz}} \quad (\text{W12})$$

$$\left\{ \begin{array}{l} \{f_{i,j+1,p}^{\text{rs}} - f_{i,j+1,p}^{\text{rsI}} - f_{i,j,p}^{\text{rs}} + f_{i,j,p}^{\text{rsO}} = 0\}_{p=i}^{j-1} \\ \{f_{i-1,j,p}^{\text{ss}} - f_{i-1,j,p}^{\text{ssI}} - f_{i,j,p}^{\text{ss}} + f_{i,j,p}^{\text{ssO}} = 0\}_{p=i+1}^j \\ U_{i,j+1}^{\text{rs}} - U_{i,j+1}^{\text{rsz}} - V_{i,j}^{\text{rs}} + V_{i,j}^{\text{rsz}} = 0 \\ U_{i-1,j}^{\text{ss}} - U_{i-1,j}^{\text{ssz}} - V_{i,j}^{\text{ss}} + V_{i,j}^{\text{ssz}} = 0 \end{array} \right. \quad (\text{W13})$$

The auxiliary variables, defined as $a_{ijp} = \left[f_{i,j+1,p}^{\text{rsI}}, f_{i-1,j,p}^{\text{ssI}}, f_{i,j,p}^{\text{rsO}}, f_{i,j,p}^{\text{ssO}}, U_{i,j+1}^{\text{rsz}}, U_{i-1,j}^{\text{ssz}}, V_{i,j}^{\text{rsz}}, V_{i,j}^{\text{ssz}} \right] = b_{ijp} \cdot \zeta_{i,j}$, where $b_{ijp} = [f_{i,j+1,p}^{\text{rs}}, f_{i-1,j,p}^{\text{ss}}, f_{i,j,p}^{\text{rs}}, f_{i,j,p}^{\text{ss}}, U_{i,j+1}^{\text{rs}}, U_{i-1,j}^{\text{ss}}, V_{i,j}^{\text{rs}}, V_{i,j}^{\text{ss}}]$, are introduced for linearization. The bilinear equalities are relaxed with McCormick envelopes [37]. The relaxation is exact, because $\zeta_{i,j}$ is always at its bound.

$$\left. \begin{array}{l} a_{ijp} \leq \min\{b_{ijp}, \zeta_{i,j} b_{ijp}^{\text{up}}\} \\ a_{ijp} \geq \max\{0, b_{ijp} + b_{ijp}^{\text{up}}(\zeta_{i,j} - 1)\} \end{array} \right\} \quad \forall \quad p = \{i, \dots, j\}, [i, j] \in \mathcal{P} \quad (\text{W14})$$

The first constraint in (W10), commonly referred as Underwood feed equation, governs Underwood roots $(\theta_{i,j,q})$ satisfying $\alpha_{q+1} \leq \theta_{i,j,q} \leq \alpha_q$. The second and third constraints in (W10) govern minimum vapor requirement in rectifying and stripping sections. The fourth and fifth constraints in (W10) enforce that the minimum vapor constraints are binding in the presence of certain splits [23]. They are imposed in BigM form with $M_{ij1} = (\Upsilon_{i,j}^{\text{rsz}})^{\text{up}} - (U_{i,j}^{\text{in}})^{\text{lo}}$ and $M_{ij2} = (\Upsilon_{i,j}^{\text{ssz}})^{\text{up}} + (U_{i,j}^{\text{in}})^{\text{up}}$, where $U_{i,j}^{\text{in}} = U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}}$ denotes the net vapor inflow into pseudocolumn $Q_{i,j}$ (see Proposition 5.2). Next, (W11) ensures that the actual vapor flow in both rectifying and stripping sections is at least as high the threshold vapor flow. (W12) implies that the difference in minimum vapor flows between the rectifying and stripping sections equals the net vapor inflow into the pseudocolumn. Finally, (W13) models component and vapor bypass, when pseudocolumn $Q_{i,j}$ is absent. The liquid bypass constraints are implied from (W13), and hence are not included in the model.

Proposition 5.2. *A valid lower bound on the nonlinear expression in (i) the second constraint in (W10) is $\min\{0, (V^{\text{in}})^{\text{lo}}\}$, and (ii) the third constraint in (W10) is $-\max\{0, (V^{\text{in}})^{\text{up}}\}$, where V^{in} is the net vapor inflow.*

Proof. (i). We denote the net component and vapor inflow into column $Q_{i,j}$ by $f_{i,j,p}^{\text{in}} = f_{i,j+1,p}^{\text{rs}} + f_{i-1,j,p}^{\text{ss}}$ and $V_{i,j}^{\text{in}} = U_{i,j+1}^{\text{rs}} - U_{i-1,j}^{\text{ss}}$. For brevity, we drop indices i and j . Let, the recovery of component p in distillate be $\gamma_p = f_p^{\text{rs}}/f_p^{\text{in}} \implies f_p^{\text{rs}} = \gamma_p f_p^{\text{in}}$. The following optimization problem is formulated:

$$\begin{aligned}
& \text{Minimize} && \sum_{p=i}^j \frac{\alpha_p \gamma_p f_p^{\text{in}}}{\alpha_p - \theta_q} \\
& \text{s.t.} && \sum_{p=i}^j \frac{\alpha_p f_p^{\text{in}}}{\alpha_p - \theta_q} = V^{\text{in}} \\
& && 1 = \gamma_i \geq \gamma_{i+1} \geq \dots \geq \gamma_j = 0 \\
& && \alpha_{q+1} \leq \theta_q \leq \alpha_q
\end{aligned} \tag{5.3}$$

The first constraint is the Underwood feed equation, and the second constraint is due to the fact that the recovery a more volatile component in the distillate is always higher than that for a less volatile component Remark. By inspection, the terms in the objective function are obtained by multiplying each fraction of the feed equation with respective recovery variable. Hence, we add the following redundant constraints to the optimization problem.

$$\frac{\alpha_p f_p^{\text{in}}}{\alpha_p - \theta_q} (\gamma_p - \gamma_q) \geq 0 \text{ for } p \in \{i, \dots, j\} \tag{5.4}$$

The constraints are valid, because for $p \leq q$ (*resp.* $p > q$) both the fraction and the recovery difference is positive (*resp.* negative). Linear combination of all the redundant constraints leads to

$$\sum_{p=i}^j \frac{\alpha_p \gamma_p f_p^{\text{in}}}{\alpha_p - \theta_q} \geq \gamma_q \sum_{p=i}^j \frac{\alpha_p f_p^{\text{in}}}{\alpha_p - \theta_q} = V^{\text{in}} \gamma_q \tag{5.5}$$

The equality is due to feed equation in (5.3). Therefore, the optimum of (5.3) is $V^{\text{in}} \gamma_q$, which is bilinear. As it is not used in the rest of the formulation, we use its lower bound instead. Since $0 \leq \gamma_q \leq 1$, the lower bound on $V^{\text{in}} \gamma_q$ is $\min\{0, (V^{\text{in}})^{\text{lo}}\}$; 0 when $(V^{\text{in}})^{\text{lo}} > 0$ and $(V^{\text{in}})^{\text{lo}}$ when $(V^{\text{in}})^{\text{lo}} < 0$.

Although similar arguments can be used to prove (ii), we describe a simpler alternative. We start with the feed equation

$$\begin{aligned}
 V^{\text{in}} &= \sum_{p=i}^j \frac{\alpha_p f_p^{\text{in}}}{\alpha_p - \theta_q} = \sum_{p=i}^j \frac{\alpha_p \gamma_p f_p^{\text{in}}}{\alpha_p - \theta_q} + \frac{\alpha_p (1 - \gamma_p) f_p^{\text{in}}}{\alpha_p - \theta_q} \\
 - \sum_{p=i}^j \frac{\alpha_p (1 - \gamma_p) f_p^{\text{in}}}{\alpha_p - \theta_q} &= \sum_{p=i}^j \frac{\alpha_p \gamma_p f_p^{\text{in}}}{\alpha_p - \theta_q} - V^{\text{in}} \geq -V^{\text{in}}(1 - \gamma_p)
 \end{aligned} \tag{5.6}$$

By recognizing that $(1 - \gamma_p) f_p^{\text{in}} = f_p^{\text{ss}}$ and $0 \leq (1 - \gamma_p) \leq 1$, the lower bound on the nonlinear expression in the third constraint in (W10) can be shown to be $-\max\{0, (V^{\text{in}})^{\text{up}}\}$. $\square$

Before proceeding further, we make a few remarks on the Underwood constraints for pseudocolumn $Q_{1,N}$. Since the composition of the feed and net vapor inflow to $Q_{1,N}$ are known, its Underwood roots $(\{\theta_{1,N,q}\}_{q=1}^{N-1})$ can be calculated a priori by solving the feed equation. Therefore, to reduce the number of nonlinear nonconvex equations in the model, we fix $\{\theta_{1,N,q}\}_{q=1}^{N-1}$ to the calculated value, and discard the feed equation for $Q_{1,N}$. Further, we substitute $\{\theta_{1,N,q}\}_{q=1}^{N-1}$ in Underwood minimum vapor constraints for $Q_{1,N}$ to make them linear. Furthermore, as we shall see shortly, Underwood roots $\{\theta_{1,N,q}\}_{q=1}^{N-1}$ are required for the computation of upper bounds on vapor flow variables.

5.1.4 Flow-of-roots Constraints

These cuts are derived by exploiting monotonicity of Underwood constraints, and are added to expedite the convergence of branch-and-bound. We refer the reader

to Chapters 3 and 4 for derivation and computational evidence demonstrating the efficacy of the cuts.

$$\left\{ \left\{ \theta_{i,l,q} - \theta_{i,j,q} \leq M'_q \left[\chi_{i,j} + \sum_{r=1}^{i-1} \sum_{s=i-1}^{j-1} \sigma_{r,s,i,j} + 1 - \sum_{r=i+1}^{j+1} \sigma_{i,j,r,l} \right] \right\}_{l=j+1}^N \right\}_{k=1}^{j-1} \bigg|_{q=i} \quad \forall [i,j] \in \mathcal{S} \quad (\text{W15})$$

Here, $M'_q = \alpha_q - \alpha_{q+1}$ corresponds to the upper bound on the difference between the roots. This concludes the discussion on model constraints, and we now move on to the objective function.

5.1.5 Objective Function and Exergy Constraints

The objective is to minimize the overall exergy loss. Consider the superstructure shown in Figure 5.2(a). The shaded region inside the dashed boundary marks the control volume. Note that heat exchangers are excluded from the control volume (see Figure 5.2(a)) to focus only on the separation process. Exergy, either in the form of heat or work, is neither added nor removed from the control volume. Thus, the total exergy loss is simply the difference between the inflow and outflow of exergy via material streams. The inflow is due to the process feed ($\mathcal{E}_{1,N}$), saturated liquid streams from condensers ($\mathcal{E}_{\text{con},i,j}^{\text{in}}$) and saturated vapor streams from reboilers ($\mathcal{E}_{\text{reb},i,j}^{\text{in}}$). Whereas the outflow is due to pure product streams ($\{\mathcal{E}_{p,p}\}_{p=1}^N$), saturated vapor streams entering condensers ($\mathcal{E}_{\text{con},i,j}^{\text{out}}$) and saturated liquid streams entering reboilers ($\mathcal{E}_{\text{reb},i,j}^{\text{out}}$). The total exergy loss in the system is given by

$$\Delta \mathcal{E}_{\text{loss}} = \left(\mathcal{E}_{1,N} - \sum_{p=1}^N \mathcal{E}_{p,p} \right) + \sum_{(i,j) \in \mathcal{C}} (\mathcal{E}_{\text{con},i,j}^{\text{in}} - \mathcal{E}_{\text{con},i,j}^{\text{out}}) + \sum_{(i,j) \in \mathcal{R}} (\mathcal{E}_{\text{reb},i,j}^{\text{in}} - \mathcal{E}_{\text{reb},i,j}^{\text{out}}). \quad (5.7)$$

Each exergy variable ($\mathcal{E}$) in (5.7) consists of chemical exergy associated with mixture ($\mathcal{E}^M$), thermal ($\mathcal{E}^T$) and pressure ($\mathcal{E}^P$) exergy associated with temperature and

pressure change from a reference state to saturated liquid state, and thermal exergy associated with phase change ($\mathcal{E}^{PC}$) from saturated liquid to the current state; i.e., $\mathcal{E} = \mathcal{E}^M + \mathcal{E}^T + \mathcal{E}^P + \mathcal{E}^{PC}$. Note that $\mathcal{E}^M = 0$ for pure components, and $\mathcal{E}^{PC} = 0$ for a saturated liquid stream. Following the procedure of [29, 63], the terms in (5.7) are evaluated to

$$\begin{aligned} \left(\mathcal{E}_{1,N} - \sum_{p=1}^N \mathcal{E}_{p,p} \right) = & \mathfrak{F} R T_0 \left(\sum_{p=1}^N z_{1,N,p}^{\text{in}} \ln z_{1,N,p}^{\text{in}} \right) + \mathfrak{F} \Delta H \int_{\Phi_{1,N}}^1 \left(1 - \frac{T_0}{T_{1,N}} \right) d\phi \\ & - \sum_{p=1}^N F_p \Delta H \int_{\Phi_{p,p}}^1 \left(1 - \frac{T_0}{T_{p,p}} \right) d\phi, \end{aligned} \quad (5.8a)$$

$$(\mathcal{E}_{\text{con},i,j}^{\text{in}} - \mathcal{E}_{\text{con},i,j}^{\text{out}}) = - F C_{i,j} \Delta H \int_0^1 \left(1 - \frac{T_0}{T_{i,j}^{\text{con}}} \right) d\phi, \quad (5.8b)$$

$$(\mathcal{E}_{\text{reb},i,j}^{\text{in}} - \mathcal{E}_{\text{reb},i,j}^{\text{out}}) = F R_{i,j} \Delta H \int_0^1 \left(1 - \frac{T_0}{T_{i,j}^{\text{reb}}} \right) d\phi, \quad (5.8c)$$

Where $\mathfrak{F} = \sum_{p=1}^N F_p$ is the total flow rate of the process feed, $z_{1,N,p}^{\text{in}} = F_p / \mathfrak{F}$ is the mole fraction of component p in the process feed, R is the universal gas constant and T_0 is the reference temperature. The terms on RHS of (5.8a), respectively, correspond to chemical exergy associated with mixing ($\mathcal{E}_{1,N}^M$), thermal exergy associated with phase change from saturated liquid state to the state corresponding to liquid fraction $\Phi_{1,N}$ ($\mathcal{E}_{1,N}^{PC}$) of process feed, and thermal exergy associated with phase change from saturated liquid state to the state corresponding to liquid fraction $\Phi_{p,p}$ ($\mathcal{E}_{p,p}^{PC}$) of pure component p . The difference in thermal ($\mathcal{E}^T$) and pressure ($\mathcal{E}^P$) exergy contributions arising from process feed and pure products is set to zero *i.e.*, $\mathcal{E}_{1,N}^T - \sum_{p=1}^N \mathcal{E}_{p,p}^T \approx 0$ and $\mathcal{E}_{1,N}^P - \sum_{p=1}^N \mathcal{E}_{p,p}^P \approx 0$. This simplification is found to be valid after rigorous testing on several systems [29].

The only difference between the inlet and outlet terms in (5.8b) and (5.8c) is the phase of the stream. Thus, only thermal exergy corresponding to phase change ($\mathcal{E}_{i,j}^{PC}$) survives in both the equations. Note that, the temperature varies during phase change for mixtures, and thus, it is a function of liquid fraction ϕ . On the other hand, for pure components, temperature remains constant during phase change

i.e., $T_{i,i}^{\text{con}} = T_{i,i}^{\text{reb}} = T_{i,i}$, and it is independent of ϕ . Now, we perform a series of manipulations to eliminate the dependence on temperature.

First, an overall energy balance on the control volume implies that the net condenser duty equals the sum of net reboiler duty and the difference between the enthalpies of process feed and pure components.

$$\sum_{(i,j) \in \mathcal{C}} FC_{i,j} \Delta H = \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \Delta H + \mathfrak{F}(1 - \Phi_{1,n}) \Delta H - \sum_{p=1}^N F_p(1 - \Phi_{p,p}) \Delta H. \quad (5.9)$$

During the energy balance, it is assumed that the sum of enthalpies of pure products at saturated liquid state is roughly the same as the enthalpy of process feed at saturated liquid state. This is also a reasonable approximation that holds for a majority of cases. Next, we multiply (5.9) with the Carnot efficiency factor measured w.r.t the boiling point of the least volatile component ($T_{N,N}$), and rewrite the equation as

$$\begin{aligned} & \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \Delta H \int_0^1 \left(1 - \frac{T_0}{T_{N,N}}\right) d\phi - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \Delta H \int_0^1 \left(1 - \frac{T_0}{T_{N,N}}\right) d\phi \\ & - \mathfrak{F} \Delta H \int_{\Phi_{1,N}}^1 \left(1 - \frac{T_0}{T_{N,N}}\right) d\phi + \sum_{p=1}^N F_p \Delta H \int_{\Phi_{p,p}}^1 \left(1 - \frac{T_0}{T_{N,N}}\right) d\phi = 0. \end{aligned} \quad (5.10)$$

(5.10) is equivalent to (5.9), because the boiling point of the least volatile component ($T_{N,N}$) and the reference temperature (T_0) are constant. Next, we add (5.10) to (5.7), substitute (5.8a) – (5.8c), and rearrange the expression to obtain

$$\begin{aligned} \Delta \mathcal{E}_{\text{loss}} = & \mathfrak{F} R T_0 \left(\sum_{p=1}^N z_{1,N,p}^{\text{in}} \ln z_{1,N,p}^{\text{in}} \right) - \mathfrak{F} T_0 \Delta H \int_{\Phi_{1,N}}^1 \left(\frac{1}{T_{1,N}} - \frac{1}{T_{N,N}} \right) d\phi \\ & + \sum_{p=1}^N F_p T_0 \Delta H \int_{\Phi_{p,p}}^1 \left(\frac{1}{T_{p,p}} - \frac{1}{T_{N,N}} \right) d\phi \\ & + \sum_{(i,j) \in \mathcal{C}} FC_{i,j} T_0 \Delta H \int_0^1 \left(\frac{1}{T_{i,j}^{\text{con}}} - \frac{1}{T_{i,i}} \right) d\phi + \sum_{(i,j) \in \mathcal{C}} FC_{i,j} T_0 \Delta H \int_0^1 \left(\frac{1}{T_{i,i}} - \frac{1}{T_{N,N}} \right) d\phi \\ & - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} T_0 \Delta H \int_0^1 \left(\frac{1}{T_{i,j}^{\text{reb}}} - \frac{1}{T_{j,j}} \right) d\phi - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} T_0 \Delta H \int_0^1 \left(\frac{1}{T_{j,j}} - \frac{1}{T_{N,N}} \right) d\phi. \end{aligned} \quad (5.11)$$

In the above equation, we have added and subtracted the inverse of boiling point of the lightest (*resp.* heaviest) component in terms associated with condensers (*resp.*

reboilers). The reason for this choice will be explained shortly. By combining Dalton's law, Raoult's law and Clausius-Clapeyron equation (both liquid and vapor phases are assumed to be ideal), [28, 29] derived

$$\left. \begin{aligned} \Delta H \left(\frac{1}{T_{j,j}} - \frac{1}{T_{N,N}} \right) &= R \ln \alpha_j \\ \Delta H \left(\frac{1}{T_{i,j}} - \frac{1}{T_{j,j}} \right) &= R \ln \left(\frac{\sum_{p=i}^j \alpha_p x_{i,j,p}}{\alpha_j} \right) \end{aligned} \right\}, \quad (5.12)$$

where $x_{i,j,p}$ is the liquid composition in equilibrium with vapor at temperature $T_{i,j}$. Substituting (5.12) in (5.11) and dividing by RT_0 leads to the scaled total exergy loss in a distillation configuration, solely expressed in terms of relative volatilities and material flows.

$$\begin{aligned} \frac{\Delta \mathcal{E}_{\text{loss}}}{RT_0} &= \mathfrak{F} \left(\sum_{p=1}^N z_{1,N,p}^{\text{in}} \ln z_{1,N,p}^{\text{in}} \right) - \mathfrak{F} \int_{\Phi_{1,N}}^1 \ln \left(\frac{\sum_{p=1}^N \alpha_p x_{1,N,p}^{\text{in}}}{\alpha_N} \right) d\phi + \sum_{p=1}^N F_p (1 - \Phi_{p,p}) \ln \alpha_p \\ &\quad + \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \int_0^1 \ln \left(\frac{\sum_{p=i}^j \alpha_p x_{i,j,p}^{\text{con}}}{\alpha_i} \right) d\phi + \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \ln \alpha_i \\ &\quad - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \int_0^1 \ln \left(\frac{\sum_{p=i}^j \alpha_p x_{i,j,p}^{\text{reb}}}{\alpha_j} \right) d\phi - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \ln \alpha_j. \end{aligned} \quad (5.13)$$

Note that, $x_{i,j,p} = x_{i,j,p}(\phi)$ since $T_{i,j} = T_{i,j}(\phi)$, for $i \neq j$. We emphasize that the temperature independence of (5.13) implies that *the total exergy loss for a mixture, with $\{\alpha_p\}_{p=1}^N$, is the same irrespective of sub-ambient or above-ambient distillation.* This is an astounding result from the foregoing procedure!

The first three terms in (5.13) can be determined prior to solving the MINLP, as they are functions of problem parameters only. Thus, we will exclude those terms from the objective function, and add the value to the optimal solution post-optimization to determine the total exergy loss. Further, we approximate the integral with a quadrature formula. A two-point quadrature is found to be sufficient in all the cases we tested. Nevertheless, for higher accuracy three-point or higher order quadrature formulas can be used. The drawback of using higher order quadrature formulas is

that the number of variables and nonconvex equations increase rapidly. The objective function after the aforementioned changes, simplifies to

$$\begin{aligned} \text{Minimize } \Delta \mathcal{E}'_{\text{loss}} \approx & \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \ln \alpha_i - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \ln \alpha_j \\ & - \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \sum_{g \in \mathcal{G}} \xi_g \ln \left(\frac{\alpha_i}{\sum_{p=i}^j \alpha_p x_{i,j,p,g}^{\text{con}}(\phi_g)} \right) \\ & - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \sum_{g \in \mathcal{G}} \xi_g \ln \left(\frac{\sum_{p=i}^j \alpha_p x_{i,j,p,g}^{\text{reb}}(\phi_g)}{\alpha_j} \right), \end{aligned} \quad (5.14)$$

where ξ_g are the Gauss weights corresponding to quadrature points ϕ_g (see §C.2). $x_{i,j,p,g}^{\text{con}}$ is the liquid phase mole fraction in equilibrium with the vapor, when ϕ_g fraction of the vapor is condensed in a condenser. Similarly, $x_{i,j,p,g}^{\text{reb}}$ is the liquid phase mole fraction in equilibrium with the vapor, when $(1 - \phi_g)$ fraction of the liquid is vaporized in a reboiler. To determine $x_{i,j,p,g}^{\text{con}}$ and $x_{i,j,p,g}^{\text{reb}}$, we perform material balance between the vapor and liquid phases. At equilibrium, total number of moles of component p equals the sum of number of moles of p in liquid (total moles of liquid times the composition of p in liquid) and vapor phases (total moles of vapor times the composition of p in vapor). We divide the material balance equation with the total number of moles, and replace the ratio of total moles of liquid (*resp.* vapor) to total moles with ϕ_g (*resp.* $1 - \phi_g$). Further, we make use of vapor-liquid equilibrium equations to express the composition of vapor phase in terms of the composition of liquid phase. Material balance equations, after making the aforementioned manipulations, are shown below:

$$\left. \begin{aligned} z_{i,j,p}^{\text{con}} &= \phi_g x_{i,j,p,g}^{\text{con}} + (1 - \phi_g) \frac{\alpha_p x_{i,j,p,g}^{\text{con}}}{\sum_{p=i}^j \alpha_p x_{i,j,p,g}^{\text{con}}} \quad \text{for } i \leq p \leq j, \quad g \in \mathcal{G}, \quad (i,j) \in \mathcal{C} \\ z_{i,j,p}^{\text{reb}} &= \phi_g x_{i,j,p,g}^{\text{reb}} + (1 - \phi_g) \frac{\alpha_p x_{i,j,p,g}^{\text{reb}}}{\sum_{p=i}^j \alpha_p x_{i,j,p,g}^{\text{reb}}} \quad \text{for } i \leq p \leq j, \quad g \in \mathcal{G}, \quad (i,j) \in \mathcal{R} \end{aligned} \right\}. \quad (5.15)$$

The LHS of the above equations corresponds to the net mole fraction of component p : moles of p in mixture/total moles of mixture. Since the liquid (*resp.* vapor) reflux from a condenser (*resp.* reboiler) has the same composition as that of the inlet vapor (*resp.* liquid) to the condenser (*resp.* reboiler) (see Figure 5.3), the net composition of component p in the inlet vapor $z_{i,j,p}^{\text{con}}$ (*resp.* inlet liquid $z_{i,j,p}^{\text{reb}}$) is equal

to its composition in distillate (*resp.* residue). Thus, in (5.16), we compute the net composition of component p in the condenser as the ratio of its flowrate in the distillate of pseudocolumn $Q_{i,j+1}$ to the net distillate flowrate, and rearrange it in bilinear form to prevent singularity issues associated with the denominator going to zero. In a similar manner, the net composition of component p in reboiler (i, j) is computed by taking the ratio of its flowrate in the residue of pseudocolumn $Q_{i-1,j}$ to the total residue flowrate.

$$\left. \begin{aligned} \left(\sum_{m=i}^j f_{i,j+1,m}^{\text{rs}} \right) z_{i,j,p}^{\text{con}} &= f_{i,j+1,p}^{\text{rs}} \quad \text{for } i \leq p \leq j, (i, j) \in \mathcal{C} \\ \left(\sum_{m=i}^j f_{i-1,j,m}^{\text{ss}} \right) z_{i,j,p}^{\text{reb}} &= f_{i-1,j,p}^{\text{ss}} \quad \text{for } i \leq p \leq j, (i, j) \in \mathcal{R} \end{aligned} \right\}. \quad (5.16)$$

Although, summation of mole fraction constraints: $\sum_{p=i}^j z_{i,j,p}^{\text{con}} = 1$, $z_{i,j,p}^{\text{reb}} = 1$, $\sum_{p=i}^j x_{i,j,p,g}^{\text{con}} = 1$, and $x_{i,j,p,g}^{\text{reb}} = 1$ are implied from (5.15) and (5.16), adding them explicitly to the model helps, as they are not implied in the relaxation.

Observe that, in (5.14), a negative sign is introduced in front of the third term by changing the argument of $\ln$ function accordingly. The advantage of this manipulation will be apparent when we reformulate the problem. Next, we remark that the argument of all $\ln$ functions in (5.14) belong to the interval $[1, \alpha_i/\alpha_j]$, because $\alpha_j \leq \sum_{p=i}^j \alpha_p x_{i,j,p} \leq \alpha_i$ as $\sum_{p=i}^j x_{i,j,p} = 1$. This is the consequence of manipulations performed in (5.11). Because $\ln$ varies steeply in the interval $(0, 1]$, as a precautionary measure, we made the arguments of all $\ln$ function to lie outside $(0, 1]$.

The MINLP (W) can be solved with (5.14) as the objective function, and by appending (5.15) and (5.16) to the constraint set. After determining the optimum $\Delta \mathcal{E}'_{\text{loss}}$, the total exergy loss ($\Delta \mathcal{E}_{\text{loss}}$) can be obtained as

$$\Delta \mathcal{E}_{\text{loss}} = RT_0 \left[\Delta \mathcal{E}'_{\text{loss}} + \mathfrak{F} \sum_{p=1}^N z_{1,N,p}^{\text{in}} \ln z_{1,N,p}^{\text{in}} - \mathfrak{F} \int_{\Phi_{1,N}}^1 \ln \left(\frac{\sum_{p=1}^N \alpha_p x_{1,N,p}^{\text{in}}}{\alpha_N} \right) d\phi + \sum_{p=1}^N F_p (1 - \Phi_{p,p}) \ln \alpha_p \right]. \quad (5.17)$$

The explicit enumeration-based algorithm introduced by [53] formulates an optimization problem for each configuration in the above manner. As is evident from (5.15)

and (5.16), this approach adds several nonlinear nonconvex equations, which may make it more challenging to solve the MINLP to ϵ -optimum. Therefore, we propose a reformulation to the model to reduce the number of nonconvex terms in the problem significantly.

The purpose of (5.15) and (5.16) is to determine the values of $x_{i,j,p,g}$, required for evaluating $\sum_p \alpha_p x_{i,j,p,g}$ in objective function. However, we are not interested in the explicit values of $x_{i,j,p,g}$ and $z_{i,j,p}$, as they do not appear anywhere else in the problem. Therefore, if there is a simpler alternative to determine $\sum_p \alpha_p x_{i,j,p,g}$ without explicitly computing $x_{i,j,p,g}$ or $z_{i,j,p}$, it would be attractive. This is achieved by performing the following manipulations. We modify the objective function by introducing auxiliary variables $\Psi_{i,j,g}$ and $\Omega_{i,j,g}$ for the arguments of $\ln$ functions

$$\begin{aligned} \text{Minimize } \Delta \mathcal{E}'_{\text{loss}} \approx & \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \ln \alpha_i - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \ln \alpha_j \\ & - \sum_{(i,j) \in \mathcal{C}} FC_{i,j} \sum_{g \in \mathcal{G}} \xi_g \ln \Psi_{i,j,g} - \sum_{(i,j) \in \mathcal{R}} FR_{i,j} \sum_{g \in \mathcal{G}} \xi_g \ln \Omega_{i,j,g}, \end{aligned} \quad (\text{W16})$$

where

$$\Psi_{i,j,g} = \frac{\alpha_i}{\sum_{p=i}^j \alpha_p x_{i,j,p,g}^{\text{con}}}, \quad (5.18a)$$

$$\Omega_{i,j,g} = \frac{\sum_{p=i}^j \alpha_p x_{i,j,p,g}^{\text{reb}}}{\alpha_j}. \quad (5.18b)$$

We substitute (5.18) in (5.15), and rearrange to obtain liquid mole fractions in terms of auxiliary variables:

$$x_{i,j,p,g}^{\text{con}} = \frac{z_{i,j,p}^{\text{con}}}{\phi_g + (1 - \phi_g)(\alpha_p/\alpha_i)\Psi_{i,j,g}}, \quad (5.19a)$$

$$x_{i,j,p,g}^{\text{reb}} = \frac{\alpha_j \Omega_{i,j,g} z_{i,j,p}^{\text{reb}}}{\phi_g \alpha_j \Omega_{i,j,g} + (1 - \phi_g) \alpha_p}. \quad (5.19b)$$

Next, to eliminate $x_{i,j,p,g}^{\text{con}}$, we sum (5.19a) over p from i to j . The left hand side equals one (sum of mole fractions). Further, we eliminate $z_{i,j,p}^{\text{con}}$ by substituting its definition from (5.16). The foregoing manipulations lead to

$$\sum_{p=i}^j \frac{f_{i,j+1,p}^{\text{rs}}}{\phi_g + (1 - \phi_g)(\alpha_p/\alpha_i)\Psi_{i,j,g}} = \sum_{p=i}^j f_{i,j+1,p}^{\text{rs}} \quad g \in \mathcal{G}; (i,j) \in \mathcal{C}. \quad (\text{W17})$$

To eliminate $x_{i,j,p}^{\text{reb}}$, we multiply (5.19b) by α_p , and sum over p (from i to j). The left hand side is replaced with $\Omega_{i,j,g}\alpha_j$ (by definition. See (5.18b)). As before, we eliminate $z_{i,j,p}^{\text{reb}}$ by substituting its definition from (5.16). The foregoing manipulations lead to

$$\sum_{p=i}^j \frac{\alpha_p f_{i-1,j,p}^{\text{ss}}}{\phi_g \alpha_j \Omega_{i,j,g} + (1 - \phi_g) \alpha_p} = \sum_{p=i}^j f_{i-1,j,p}^{\text{ss}} \quad g \in \mathcal{G}; (i, j) \in \mathcal{R}. \quad (\text{W18})$$

We remark that, (W18) can also be derived in a manner similar to (W17) with an additional partial fraction decomposition step. Because (W17) and (W18) govern variables $\Psi_{i,j,g}$ and $\Omega_{i,j,g}$, which are required for the computation of exergy loss, we refer them as *exergy constraints associated with condensers and reboilers*, respectively. Observe that the solution of (W17) (*resp.* (W18)) directly yields the value of $\sum_p \alpha_p x_{i,j,p}^{\text{con}}$ (*resp.* $\sum_p \alpha_p x_{i,j,p}^{\text{reb}}$), without explicitly calculating $x_{i,j,p}^{\text{con}}$ and $z_{i,j,p}^{\text{con}}$ (*resp.* $x_{i,j,p}^{\text{reb}}$ and $z_{i,j,p}^{\text{reb}}$). Therefore, by appending (W17) and (W18) to the constraint set, (5.15) and (5.16) can be eliminated from the model, as well as variables $x_{i,j,p}^{\text{con}}$, $z_{i,j,p}^{\text{con}}$, $x_{i,j,p}^{\text{reb}}$ and $z_{i,j,p}^{\text{reb}}$. As a result, for each heat exchanger and at each quadrature point, one exergy constraint replaces $(j - i + 1)$ vapor-liquid equilibrium equations in (5.15) and $(j - i + 1)$ mole fraction computations in (5.16). Clearly, the proposed reformulation reduces the number of nonlinear nonconvex equations drastically, making the model cleaner and simple. Finally, exergy constraints associated with condensers and reboilers have the following properties, one of which is exploited in the next subsection to derive additional valid cuts.

Property 1. For a given $\{f_{i,j+1,p}^{\text{rs}}\}_{p=i}^j$ (*resp.* $\{f_{i-1,j,p}^{\text{ss}}\}_{p=i}^j$), there exists a unique $\Psi_{i,j,g}$ (*resp.* $\Omega_{i,j,g}$) satisfying (W17) (*resp.* (W18)).

Proof. For $\phi_g \in [0, 1]$ (see §C.2), the LHS of (W17) (*resp.* (W18)) is continuous and decreases monotonically with $\Psi_{i,j,g}$ (*resp.* $\Omega_{i,j,g}$). Further, it can be shown that,

$$\begin{aligned} \sum_{p=i}^j \frac{f_{i,j+1,p}^{\text{rs}}}{\phi_g + (1 - \phi_g)(\alpha_p/\alpha_i)\Psi_{i,j,g}^{\text{lo}}} &\geq \sum_{p=i}^j f_{i,j+1,p}^{\text{rs}} \geq \sum_{p=i}^j \frac{f_{i,j+1,p}^{\text{rs}}}{\phi_g + (1 - \phi_g)(\alpha_p/\alpha_i)\Psi_{i,j,g}^{\text{up}}}, \\ \sum_{p=i}^j \frac{\alpha_p f_{i-1,j,p}^{\text{ss}}}{\phi_g \alpha_j \Omega_{i,j,g}^{\text{lo}} + (1 - \phi_g) \alpha_p} &\geq \sum_{p=i}^j f_{i-1,j,p}^{\text{ss}} \geq \sum_{p=i}^j \frac{\alpha_p f_{i-1,j,p}^{\text{ss}}}{\phi_g \alpha_j \Omega_{i,j,g}^{\text{up}} + (1 - \phi_g) \alpha_p}, \end{aligned}$$

where $\Psi_{i,j,g}^{\text{lo}} = \Omega_{i,j,g}^{\text{lo}} = 1$ and $\Psi_{i,j,g}^{\text{up}} = \Omega_{i,j,g}^{\text{up}} = \alpha_i/\alpha_j$ (see (5.18)). Application of Intermediate value theorem implies that there exists a solution for (W17) (*resp.* (W18)) in the interval $[\Psi_{i,j,g}^{\text{lo}}, \Psi_{i,j,g}^{\text{up}}]$ (*resp.* $[\Omega_{i,j,g}^{\text{lo}}, \Omega_{i,j,g}^{\text{up}}]$), and monotonicity guarantees uniqueness. This property avoids the issues associated with solution multiplicity. $\square$

Property 2. $\Psi_{i,j}$ (*resp.* $\Omega_{i,j}$) decreases (*resp.* increases) monotonically with increase in ϕ .

Proof. First, we prove $\Psi_{i,j}$ decreases monotonically with ϕ . For conciseness, we drop indices ' $i, j+1$ ' and superscript rs from $f_{i,j+1,p}^{\text{rs}}$, drop indices ' i, j ' from $\Psi_{i,j}$, and define a new variable for the denominator of fractions in (W17): $\mathbf{D}_p = \phi + (1-\phi)(\alpha_p/\alpha_i)\Psi$. Differentiating (W17) w.r.t ϕ partially, leads to

$$-\sum_p \frac{f_p}{\mathbf{D}_p^2} \left\{ 1 + (1-\phi) \left(\frac{\alpha_p}{\alpha_i} \right) \frac{\partial \Psi}{\partial \phi} - \left(\frac{\alpha_p}{\alpha_i} \right) \Psi \right\} = 0.$$

Multiplying the above equation with $(1-\phi)$, and rearranging the resulting expression yields

$$\begin{aligned} \frac{\partial \Psi}{\partial \phi} &= \left[-\sum_p \frac{f_p}{\mathbf{D}_p^2} (1-\phi)^2 \left(\frac{\alpha_p}{\alpha_i} \right) \right]^{-1} \left[\sum_p \frac{f_p}{\mathbf{D}_p^2} \left\{ 1 - \phi - (1-\phi) \left(\frac{\alpha_p}{\alpha_i} \right) \Psi \right\} \right] \\ &= \left[-\sum_p \frac{f_p}{\mathbf{D}_p^2} (1-\phi)^2 \left(\frac{\alpha_p}{\alpha_i} \right) \right]^{-1} \left[\sum_p \frac{f_p}{\mathbf{D}_p^2} - \frac{f_p}{\mathbf{D}_p} \right] \left[\frac{\sum_p f_m}{\sum_p f_m} \right] \\ &= \left[-\sum_p \frac{f_p}{\mathbf{D}_p^2} (1-\phi)^2 \left(\frac{\alpha_p}{\alpha_i} \right) \right]^{-1} \left[\left(\sum_p \frac{f_p}{\mathbf{D}_p^2} \right) \left(\sum_p f_p \right) - \left(\sum_p \frac{f_p}{\mathbf{D}_p} \right)^2 \right] \left[\frac{1}{\sum_p f_m} \right]. \end{aligned} \quad (5.20)$$

In the last line, we used $\sum_p f_p = \sum_p f_p/\mathbf{D}_p$ (see (W17)). Using the identity $\left(\sum_p u_p w_p \right)^2 + \sum_p \sum_{l, l < p} (u_p - u_l)^2 w_p w_l = \left(\sum_p w_p \right) \left(\sum_p u_p^2 w_p \right)$ (see §C.3 for proof), with $u_p \equiv 1/\mathbf{D}_p$ and $w_p \equiv f_p$, we express the second term in the last line of (5.20) as

$$\frac{\partial \Psi}{\partial \phi} = \left[-\sum_p \frac{f_p}{\mathbf{D}_p^2} (1-\phi)^2 \left(\frac{\alpha_p}{\alpha_i} \right) \right]^{-1} \left[\sum_p \sum_{l, l < p} \left(\frac{1}{\mathbf{D}_p} - \frac{1}{\mathbf{D}_l} \right)^2 f_p f_l \right] \left[\frac{1}{\sum_p f_m} \right]. \quad (5.21)$$

Observe that the first term on RHS of (5.21) is always negative, while the second and third terms are positive. Therefore, $\frac{\partial \Psi}{\partial \phi} < 0$, indicating that $\Psi_{i,j,g}$ decreases monotonically with increase in ϕ .

Although similar arguments can be used to prove $\Omega_{i,j,g}$ increases monotonically with ϕ , we show a simpler alternative. We recognize that $\Omega \propto \Psi^{-1}$, by definition (see (5.18)). Therefore, $\Psi_{i,j}$ is a monotonically decreasing function of ϕ implies that $\Omega_{i,j}$ increases monotonically with ϕ . $\square$

Remark 5.1. By definition, thermal quality equals zero ($\phi = 0$) for a saturated vapor stream and equals one ($\phi = 1$) for a saturated liquid stream. As a result, ϕ increases monotonically with decrease in temperature *i.e.*, $\phi_1 < \phi_2$ for $T(\phi_1) > T(\phi_2)$. By combining this observation with Property 2, we observe that $\Psi_{i,j}$ (*resp.* $\Omega_{i,j}$) increases (*resp.* decreases) monotonically with increase in temperature.

5.1.6 Exploiting Monotonicity of Exergy Constraints

In our prior work [23], we presented computational evidence illustrating the importance of (W15), derived by exploiting the monotonicity of Underwood constraints, in expediting the convergence of branch-and-bound. We have also shown cases where global solvers fail to converge in a reasonable amount of time, in the absence of (W15). This was because, the monotonicity constraints, though implicit in Underwood constraints, are not implied in the relaxation. Explicitly adding the constraints to the model aids global solvers to use the information in inferring tighter bounds on Underwood roots via feasibility and optimality based bound tightening techniques. Based on this evidence, we expect that the constraints derived by exploiting monotonicity of exergy constraints (Property 2) have potential in expediting the convergence of branch-and-bound, as they will not be implied in the relaxation.

We present the constraints for two-point quadrature formula. The extension to higher order quadrature formulas is straightforward. Exergy constraint in (W17) (*resp.* (W18)) is written at both quadrature points ϕ_1 and ϕ_2 (where $\phi_1 < \phi_2$). See

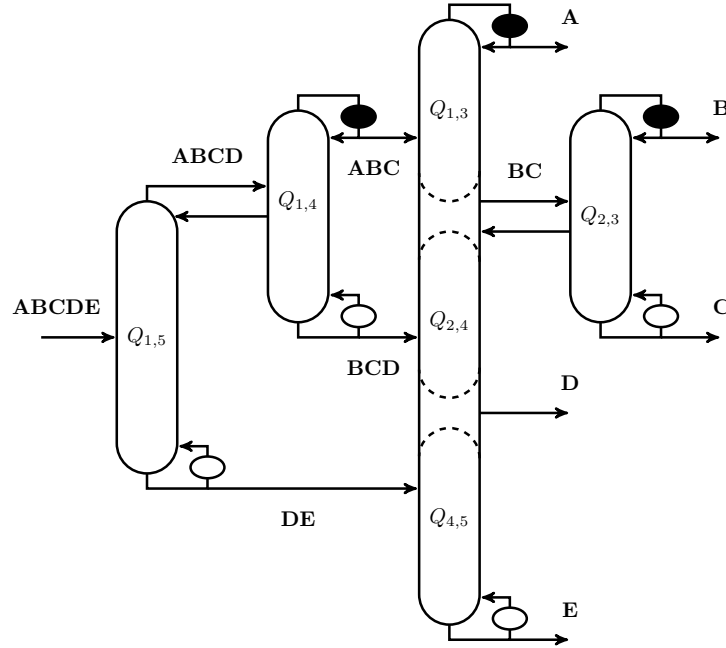


Fig. 5.4. A five-component configuration motivating the need for constraints in section 2.7

§C.2). Because $\Psi_{i,j,g}$ (*resp.* $\Omega_{i,j,g}$) is a monotonically decreasing (*resp.* increasing) function of ϕ (see Property 2), we have

$$\left. \begin{aligned} \Psi_{i,j,1} &\geq \Psi_{i,j,2}, & \forall & (i,j) \in \mathcal{C} \\ \Omega_{i,j,1} &\leq \Omega_{i,j,2}, & \forall & (i,j) \in \mathcal{R} \end{aligned} \right\} \quad (\text{W19})$$

We shall see in the next subsection, that these constraints are not only useful, but necessary for the model.

5.1.7 Relation Between Temperatures of Condensers and Reboilers

The motivation for the inclusion of the following constraints is better understood with an example. Consider a five-component mixture of hydrocarbons: $A \equiv$ Propylene ($F_1 = 5$, $\alpha_1 = 11.761$), $B \equiv$ Propane ($F_2 = 5$, $\alpha_2 = 9.801$), $C \equiv$ 1-Butene ($F_3 = 32.5$, $\alpha_3 = 3.63$), $D \equiv$ n-Butane ($F_4 = 25$, $\alpha_4 = 3$) and $E \equiv$ n-Pentane ($F_5 = 32.5$, $\alpha_5 = 1$). Flowrate of each component (F_p) in the process feed ($[1, N]$) is given

in kmol/h, and the relative volatility information is taken from [46]. The remaining problem parameters are taken to be $\Phi_{1,N} = 1$ and $\{\Phi_{p,p}\}_{p=1}^N = \{1, 1, 1, 1, 1\}$. We fix the discrete variables ($\zeta_{i,j}$, $\chi_{i,j}$ and $\rho_{i,j}$) to configuration shown in Figure 5.4, and solve the MINLP (W). The optimum objective function value ($\Delta\mathcal{E}'_{\text{loss}}$) is -84.671 making the first term on the right hand side of (5.17) negative. Further, the second term of (5.17) is also negative, the third and fourth terms for the chosen parameters evaluate to zero; implying that the total exergy loss for the configuration at optimum operation is negative. Physically, negative exergy loss implies that the exergy is generated in the system, because of which, the net outflow of exergy from the system is greater than the net inflow of exergy into the system. In other words, work is produced by distilling the hydrocarbon mixture with the configuration in Figure 5.4. Obviously, such solutions are infeasible as they violate physics of the distillation process. To understand why such solutions are feasible to the model, we examine the optimum operating conditions (material flows) more closely.

The net component, vapor and liquid flows in the rectifying and stripping sections of all pseudocolumns at optimum are shown in Table 5.2. In particular, observe the net component flows in the stripping section of pseudocolumn $Q_{1,4}$. The net component flowrate, and thus, the net flowrate of residue is zero. As a result, exergy constraints for reboiler (2, 4) are satisfied trivially; making variables $\Omega_{2,4,1}$ and $\Omega_{2,4,2}$, which are a measure of temperature, become unrestricted. This makes the optimizer push the temperature of the reboiler (2, 4) to the lowest value possible, by pushing both $\Omega_{2,4,1}$ and $\Omega_{2,4,2}$ to their upper bound (see Remark 1 for dependence of temperature on $\Omega_{i,j}$). In fact, for the configuration in Figure 5.4, the temperature of condenser (1, 3) at optimum is higher than the temperature of reboiler (2, 4), enabling extraction of work by running a heat engine between the heat exchangers.

Table 5.2.

Material flowrates (in kmol/h) corresponding to the optimal operation of configuration in Figure. Material flowrates (in kmol/h) in condensers and reboilers at optimum operation: $FC_{1,1} = 63.659$, $FC_{2,2} = 20.912$, $FC_{1,3} = 1122.545$, $FR_{3,3} = 30.822$, $FR_{5,5} = 53.75$, $FR_{2,4} = 801.818$ and $FR_{4,5} = 57.5$. The remaining variables can be determined by performing mass balances, and are not shown in the table. The problem parameters are taken to be $N = 5$, $\{F_p\}_{p=1}^N = \{5, 5, 32.5, 25, 32.5\}$, $\{\alpha_p\}_{p=1}^N = \{11.761, 9.801, 3.63, 3, 1\}$, $\Phi_{1,n} = 1$ and $\{\Phi_{p,p}\}_{p=1}^N = \{1, 1, 1, 1, 1\}$. In the table, we write ‘-’ to indicate that the variable is not defined in the problem. The objective function value ($\Delta\mathcal{E}'_{\text{loss}}$) at optimum operation: -84.671 .

Material flows in rectifying sections								Material flows in stripping sections						
$Q_{i,j}$	$f_{i,j,1}^{\text{rs}}$	$f_{i,j,2}^{\text{rs}}$	$f_{i,j,3}^{\text{rs}}$	$f_{i,j,4}^{\text{rs}}$	$f_{i,j,5}^{\text{rs}}$	$V_{i,j}^{\text{rs}}$	$L_{i,j}^{\text{rs}}$	$f_{i,j,1}^{\text{ss}}$	$f_{i,j,2}^{\text{ss}}$	$f_{i,j,3}^{\text{ss}}$	$f_{i,j,4}^{\text{ss}}$	$f_{i,j,5}^{\text{ss}}$	$V_{i,j}^{\text{ss}}$	$L_{i,j}^{\text{ss}}$
$Q_{1,5}$	5	5	32.5	0	-	320.727	278.227	-	0	0	25	32.5	320.727	378.227
$Q_{1,4}$	5	5	32.5	-	-	1122.545	1080.045	-	0	0	0	-	801.818	801.818
$Q_{1,3}$	5	0	-	-	-	63.659	58.659	-	5	32.5	-	-	63.659	101.159
$Q_{2,4}$	-	0	0	-	-	53.75	53.75	-	-	0	0	-	53.75	53.75
$Q_{4,5}$	-	-	-	25	-	53.75	28.75	-	-	-	-	32.5	53.75	86.25
$Q_{2,3}$	-	5	-	-	-	20.912	15.912	-	-	32.5	-	-	30.822	63.322

In contrast, when the net distillate and residue flows are nonzero, variables $\Psi_{i,j}$ and $\Omega_{i,j}$ are restricted by exergy constraints. As such, Underwood method ensures that the distillate from a column is “lighter” than the residue from the same column; solutions making the condenser of a pseudocolumn hotter than its reboiler are infeasible to the model. In the explicit enumeration approach of [53], the authors impose a small, but nonzero, lower bound on distillate and residue flows. Consequently, the authors never encountered situations which violated the physics of distillation process. However, in our model, we optimize over the entire search space of configurations, and thus nonzero lower bounds on net component flows cannot be imposed. This is the reason why our model admits the solution in Table 5.2, while it is infeasible to the model of [53].

We resolve this issue by explicitly imposing constraints which ensure that the condenser of a pseudocolumn is always colder than its reboiler. These constraints are redundant when the net distillate and residue flows are nonzero, but are required otherwise. Consider pseudocolumn $Q_{i,j}$. The hottest (*resp.* coldest) temperature in condenser $(i, j - 1)$ (*resp.* reboiler $(i + 1, j)$) is the dew point ($T_{i,j-1}^{dew}$) (*resp.* bubble point ($T_{i+1,j}^{bub}$)) of distillate (*resp.* residue). In any distillation column $T_{i,j-1}^{dew} \leq T_{i+1,j}^{bub}$ or $(1/T_{i,j-1}^{dew}) \geq (1/T_{i+1,j}^{bub})$, which we rewrite as

$$\left(\frac{1}{T_{i,j-1}^{dew}} - \frac{1}{T_{i,i}} \right) + \left(\frac{1}{T_{i,i}} - \frac{1}{T_{N,N}} \right) \geq \left(\frac{1}{T_{i+1,j}^{bub}} - \frac{1}{T_{j,j}} \right) + \left(\frac{1}{T_{j,j}} - \frac{1}{T_{N,N}} \right). \quad (5.22)$$

Using (5.12) and (5.18), we rewrite the above constraint as

$$-\ln(\Psi_{i,j-1}^{dew}) + \ln \alpha_i \geq \ln(\Omega_{i+1,j}^{bub}) + \ln \alpha_j, \quad (5.23)$$

which is further simplified to

$$\Psi_{i,j-1}^{dew} \Omega_{i+1,j}^{bub} \leq \frac{\alpha_i}{\alpha_j}. \quad (5.24)$$

Here, $\Psi_{i,j-1}^{dew} = \Psi_{i,j-1}(\phi = 0)$ (*resp.* $\Omega_{i+1,j}^{bub} = \Omega_{i+1,j}(\phi = 1)$) is determined by solving an equation of the form (W17) (*resp.* (W18)) with ϕ_g substituted to 0 (*resp.* 1). Clearly, (5.24) requires introduction of additional variables and nonlinear constraints.

However, we exploit Property 2, and derive a relaxed version of (5.24) in terms of problem variables.

Because $\Psi_{i,j-1}$ decreases monotonically with increase in ϕ , and $0 < \phi_1$, $\Psi_{i,j-1}(\phi = 0) > \Psi_{i,j-1}(\phi = \phi_1)$ or $\Psi_{i,j-1}^{dew} > \Psi_{i,j-1,1}$. Similarly, $\Omega_{i+1,j}(\phi = 1) > \Omega_{i+1,j}(\phi = \phi_2)$ or $\Omega_{i+1,j}^{bub} > \Omega_{i+1,j,2}$ because $1 > \phi_2$ and $\Omega_{i+1,j}$ increases monotonically with increase in ϕ . From the foregoing arguments, $\Psi_{i,j-1,1}\Omega_{i+1,j,2} < \Psi_{i,j-1}^{dew}\Omega_{i+1,j}^{bub}$, which in conjunction with (5.24) leads to

$$\Psi_{i,j-1,1}\Omega_{i+1,j,2} \leq \frac{\alpha_i}{\alpha_j}, \quad \forall \quad [i, j] \in \mathcal{P}. \quad (\text{W20})$$

Obviously, this raises the following question: Does the relaxation of (5.24) add physically infeasible solutions to the model? The answer is no, and the reason is stated as a proposition below.

Proposition 5.3. (W20) along with monotonicity constraints in (W19) imply that the exergy associated with unit heat at reboiler of a pseudocolumn is greater than the exergy associated with unit heat at condenser of the same pseudocolumn.

Proof. Consider pseudocolumn $Q_{i,j}$. We need to show

$$\int_0^1 \left(1 - \frac{T_0}{T_{i+1,j}^{\text{reb}}} \right) d\phi \geq \int_0^1 \left(1 - \frac{T_0}{T_{i,j-1}^{\text{con}}} \right) \phi,$$

As in (5.22) and (5.23), we add and subtract appropriate terms, use (5.12) and (5.18) to rewrite the above inequality as shown below

$$\int_0^1 (-\ln \Omega_{i+1,j} - \ln \Psi_{i,j-1} + \ln \alpha_i - \ln \alpha_j) d\phi \geq 0.$$

Next, we approximate the integral with two-point Gauss quadrature formula to obtain

$$\Psi_{i,j-1,1}\Psi_{i,j-1,2}\Omega_{i+1,j,1}\Omega_{i+1,j,2} \leq \left(\frac{\alpha_i}{\alpha_j} \right)^2. \quad (5.25)$$

We show that (W20) and (W19) imply the inequality in (5.25). $\Psi_{i,j-1,1} \geq \Psi_{i,j-1,2}$ implies $\Psi_{i,j-1,1}\Omega_{i+1,j,2} \geq \Psi_{i,j-1,2}\Omega_{i+1,j,2}$, and $\Omega_{i+1,j,2} \geq \Omega_{i+1,j,1}$ implies $\Omega_{i+1,j,2}\Psi_{i,j-1,1} \geq \Omega_{i+1,j,1}\Psi_{i,j-1,1}$. Combining the arguments leads to $\Omega_{i+1,j,1}\Psi_{i,j-1,1}\Psi_{i,j-1,2}\Omega_{i+1,j,2} \leq (\Omega_{i+1,j,2}\Psi_{i,j-1,1})^2$, which, along with (W20), implies (5.25). $\square$

This concludes the formulation (W). It can be readily solved to ϵ -global optimum with off-the-shelf solvers like BARON.

5.2 Computational Studies

The purpose of this section is twofold. First, with a case study, we present two ways of using the MINLP formulation (W): identification of *simple* (requiring fewer number of column sections) yet exergetically favorable configurations, and determining good retrofits. Second, with several four and five component examples from literature, we investigate the solution performance to changes in problem parameters *viz.*, composition of process feed $\{F_p\}_{p=1}^N$ and relative volatilities $\{\alpha_p\}_{p=1}^N$.

5.2.1 Case Study: Shale Gas Separation

Due to advances in technology such as hydraulic fracturing and horizontal drilling, the production of shale gas has increased rapidly in the United States. Shale gas contains considerable amount of Natural Gas Liquids (NGLs): mixture of C2, C3, C4, C5⁺. For example, shale gas from Eagle Ford contains roughly 24 mol % of NGLs [68,69]. Motivated by applications such as gasoline blending (C5⁺), production of olefins (feedstock for production of several value-added chemicals) etc., it is desired to recover NGLs from shale gas. For this task, distillation is attractive owing to (i) the relative ease of separation of consecutive components, (ii) ability to produce high purity products and (iii) ability to process large quantities of feed. Since the lighter hydrocarbons boil at sub-ambient temperatures, a sub-ambient distillation configuration, which uses work rather than heat, is required for the portion of flowsheet separating lighter hydrocarbons. Minimization of exergy losses is crucial for such flowsheets, as the work input is proportional to the total exergy losses. Therefore, the formulation developed in the previous section is useful for this application.

Table 5.3.
Composition of shale gas from Eagle Ford

	Component	Composition (mol%)	Relative Volatility
<i>A</i>	C1	75.52	24
<i>B</i>	C2	14	4.21
<i>C</i>	C3	5.49	2.08
<i>D</i>	C4	2.82	1.32
<i>E</i>	C5 ⁺	2.17	1

As an example, we consider a shale gas mixture from Eagle Ford, whose composition is shown in Table 5.3. We remark that shale gas also contains acid gases (CO_2 and H_2S), N_2 and H_2O . However, these gases are generally removed prior to distillation. Therefore, we borrowed the shale gas composition from [68,69], removed acid gases, N_2 and H_2O , and normalized the composition of hydrocarbons. We refer the reader to §C.4 for a description on the computation of relative volatilities reported in Table 5.3. We consider the case, where the process feed is fed as a saturated vapor ($\Phi_{1,N} = 0$) and the pure components are withdrawn as saturated vapors ($\{\Phi_{p,p}\}_{p=1}^N = \{0, 0, 0, 0, 0\}$). We use BARON 18.5.1 on GAMS 25.1 to solve MINLPs to ϵ -global optimum, with relative tolerance for convergence (ϵ_r) set to 1%.

First, we determine the best *sharp-split configuration* for this application. Of all regular-column configurations, sharp-split configurations have the least number of column sections and inter-column transfer streams; making them attractive from operational standpoint. A characteristic of sharp-split configurations is that they have exactly $(N - 2)$ submixture streams, one process feed stream ($[1, N]$) and N pure component product streams ($\{[p, p]\}_{p=1}^N$). Therefore, to tailor the search space to contain only sharp-split configurations, we add

$$\sum_{[i,j] \in \mathcal{T}} \zeta_{i,j} = (N - 2) + (N + 1). \quad (5.26)$$

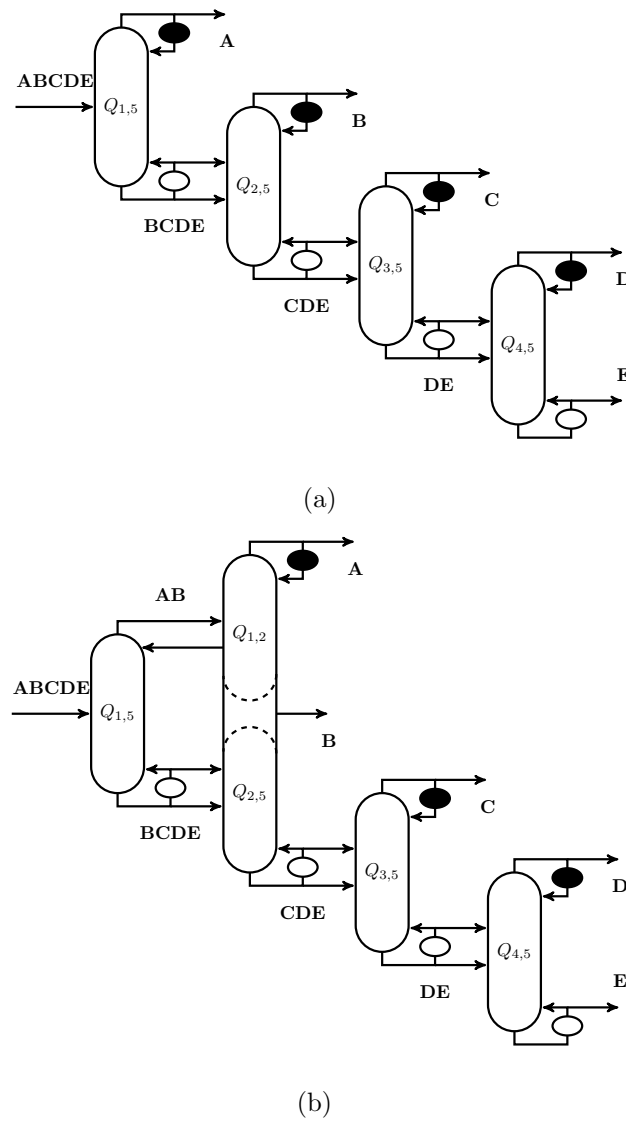


Fig. 5.5. (a) Conventional configuration for shale gas separation (b) A plausible retrofit to reduce exergy loss. Operating conditions at optimum operation are shown in Table 5.4.

to (W). Solving the MINLP to ϵ -optimality yields the configuration shown in Figure 5.5(a). This configuration, commonly referred as the direct-split configuration, is currently the preferred choice for shale gas separation in industry. Our model, despite several simplifying assumptions, determined it to be the best among all 112 sharp-split

configurations in less than a minute. This illustrates that the simplifying assumptions are reasonable and the model results in worthwhile solutions.

However, it is well known that sharp-split configurations are not always attractive in terms of overall energy requirement. Substantial benefits can be obtained by including sloppy splits in a configuration. Therefore, to assess the penalty from using a sharp-split configuration, we solve (W) over the entire space of regular-column configurations, and compare the optimal exergy loss with that of the direct-split configuration. The optimal configuration among all the regular-column configurations has an exergy loss of $29.96 RT_0$ kJ/h. On the other hand, the total exergy loss for the direct-split configuration at optimum operation is $47.73 RT_0$ kJ/h, which is 59.3% higher than the optimal configuration. This shows that there is a large potential in reducing the overall exergy losses, and in turn the overall energy requirement, by introducing sloppy splits in the configuration.

Since *ab initio* redesign of an existing plant may not be economical, the process designer would be interested in a plausible retrofit for the direct-split configuration. For this purpose, we look for configurations that are structurally similar to the direct-split configuration. We add

$$\begin{aligned} \zeta_{2,5} = \zeta_{3,5} = \zeta_{4,5} &= 1 \\ \sum_{[i,j] \in \mathcal{T}} \zeta_{i,j} &= (N-2) + (N+1) + 1 \end{aligned} \tag{5.27}$$

to (W). The first constraint enforces that the desired configuration contains submixtures $BCDE$, CDE and DE . Whereas, the second constraint is added to limit the increase in number of column sections to two. This way, we seek configurations that may provide substantial benefits with minimal increase in operational complexity. The optimal solution is shown in Figure 5.5(b), and its exergy loss is $36.45 RT_0$ kJ/h. Table 5.4 lists the material flows in rectifying and stripping sections of all pseudo-columns at optimum operation for both configurations in Figure 5.5. In particular, we point out that the duty of condenser (1, 1) is the same for both the configurations.

Table 5.4.

Optimal material flowrates (in kmol/h) in rectifying and stripping sections for configurations in Figure 5.5. For Figure 5.5(a), material flowrates (in kmol/h) in condensers and reboilers at optimum operation: $FC_{1,1} = 45.54$, $FC_{2,2} = 17.64$, $FC_{3,3} = 6.23$, $FC_{4,4} = 4.03$, $FR_{2,5} = 41.73$, $FR_{3,5} = 17.68$, $FR_{4,5} = 8.77$ and $FR_{5,5} = 3.1$. For Figure 5.5(b), material flowrates (in kmol/h) in condensers and reboilers at optimum operation: $FC_{1,1} = 45.54$, $FC_{3,3} = 6.23$, $FC_{4,4} = 4.03$, $FR_{2,5} = 27.04$, $FR_{3,5} = 14.73$, $FR_{4,5} = 8.77$ and $FR_{5,5} = 3.1$. As before, we write ‘-’ to indicate that the variable is not defined in the problem.

Material flows in rectifying sections							Material flows in stripping sections							
Direct-split configuration (Figure 5.5(a))														
$Q_{i,j}$	$f_{i,j,1}^{rs}$	$f_{i,j,2}^{rs}$	$f_{i,j,3}^{rs}$	$f_{i,j,4}^{rs}$	$f_{i,j,5}^{rs}$	$V_{i,j}^{rs}$	$L_{i,j}^{rs}$	$f_{i,j,1}^{ss}$	$f_{i,j,2}^{ss}$	$f_{i,j,3}^{ss}$	$f_{i,j,4}^{ss}$	$f_{i,j,5}^{ss}$	$V_{i,j}^{ss}$	$L_{i,j}^{ss}$
$Q_{1,5}$	75.52	0	0	0	-	116.52	41	-	14	5.49	2.82	2.17	16.52	41
$Q_{2,5}$	-	14	0	0	-	42.51	28.51	-	-	5.49	2.82	2.17	19.03	29.51
$Q_{3,5}$	-	-	5.49	0	-	24.61	19.12	-	-	-	2.82	2.17	16.1	21.09
$Q_{4,5}$	-	-	-	2.82	-	20.58	17.76	-	-	-	-	2.17	15.59	17.76

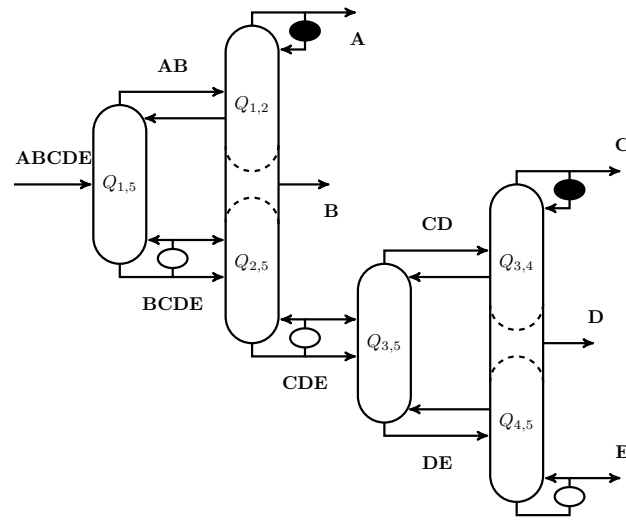
Best retrofit for Direct-split configuration (Figure 5.5(b))														
$Q_{i,j}$	$f_{i,j,1}^{rs}$	$f_{i,j,2}^{rs}$	$f_{i,j,3}^{rs}$	$f_{i,j,4}^{rs}$	$f_{i,j,5}^{rs}$	$V_{i,j}^{rs}$	$L_{i,j}^{rs}$	$f_{i,j,1}^{ss}$	$f_{i,j,2}^{ss}$	$f_{i,j,3}^{ss}$	$f_{i,j,4}^{ss}$	$f_{i,j,5}^{ss}$	$V_{i,j}^{ss}$	$L_{i,j}^{ss}$
$Q_{1,5}$	75.52	8.45	0	0	-	108.12	24.15	-	5.55	5.49	2.82	2.17	8.12	24.15
$Q_{1,2}$	75.52	-	-	-	-	116.52	41	-	8.45	-	-	-	8.4	16.85
$Q_{2,5}$	-	5.55	0	0	-	22.4	16.85	-	-	5.49	2.82	2.17	13.52	24
$Q_{3,5}$	-	-	5.49	0	-	24.61	19.12	-	-	-	2.82	2.17	16.1	21.09
$Q_{4,5}$	-	-	-	2.82	-	20.58	17.76	-	-	-	-	2.17	15.59	17.76

Further, in Figure 5.5(b), the liquid reflux produced at condenser (1, 1) is used efficiently between $Q_{1,5}$, $Q_{1,2}$ and $Q_{2,5}$. On the other hand, in Figure 5.5(a), the required liquid reflux for the columns $Q_{1,5}$ and $Q_{2,5}$ is produced independently at condensers (1, 1) and (2, 2), respectively; thereby arising the need for additional refrigeration at condenser (2, 2). For this reason, the configuration in Figure 5.5(b) is attractive over the direct-split configuration. Moreover, it is structurally similar to Figure 5.5(a), and thus, amenable for retrofitting.

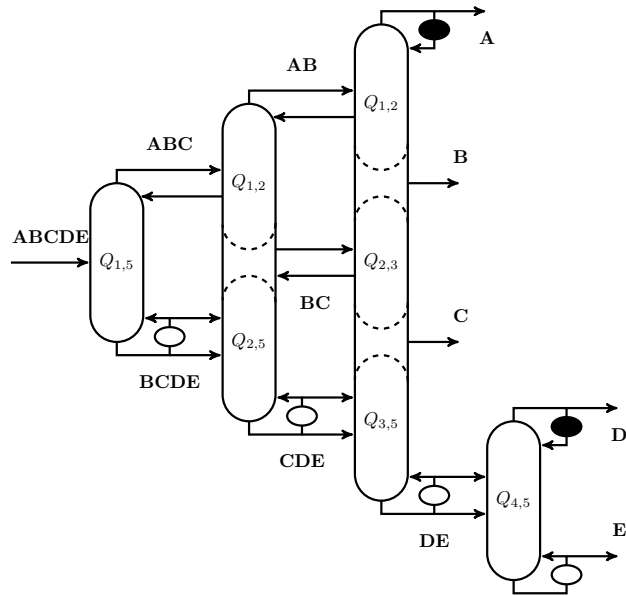
Nonetheless, the configuration in Figure 5.5(b) has 21.66% higher exergy loss than that of the optimal solution over the entire search space. Alternative solutions that provide additional benefit may be determined by relaxing the limitation on the number of column sections and re-solving (W). Figures 5.6(a) and 5.6(b) show the optimal configurations with four and six additional sections than the direct split configuration. The optimal exergy loss for Figure 5.6(a) is $35.62 RT_0$ kJ/h (18.9% higher than the optimal configuration over the entire search space) and Figure 5.6(b) is $31.4 RT_0$ kJ/h (4.8% higher than the optimal configuration over the entire search space). Now, the process designer has to assess trade-offs between the benefit from introducing additional sections in configurations vs. increased operational complexity, in order to determine the best configuration for the separation of shale gas.

5.2.2 Examples from Literature

In addition to the shale gas example discussed in the previous subsection, we also tested the model (W) on several four and five component examples taken from the literature. Problem parameters are reported in the third and fourth columns of Table 5.5. The remaining parameters are taken to be $\Phi_{1,n} = 1$ and $\{\Phi_{p,p} = 1\}_{p=1}^N$ for all cases. We chose BARON 18.5.1 on GAMS 25.1 to solve the MINLP (W), with all options, except pDo, set at their defaults. We observed that BARON can solve (W) faster, when pDo is set to -1 *i.e.*, when all variables are probed. Finally, we chose 1% for relative tolerance for convergence (ϵ_r) and 10000 s for time limit as the termination



(a)



(b)

Fig. 5.6. (a) Conventional configuration for shale gas separation (b) A plausible retrofit to reduce exergy loss.

criteria. All computations were done on Dell Optiplex 5040 with 16 GB RAM Intel Core i7-6700 3.4 GHz processor. The results are summarized in Table 5.5. Evidently, BARON could solve all the cases considered to 1%-global optimum within the

Table 5.5.

Examples from literature. The remaining problem parameters are taken be $\Phi_{1,N} = 1$ and $\{\Phi_{p,p} = 1\}_{p=1}^N$. Case 1 is taken from [10], Cases 2, 3 and 4 from [45], Cases 6, 8 and 10 from [9], Cases 7, 9 and 11 from [17]. Relative volatility information for Cases 12, 13, 14 and 15 are taken from [46]. All BARON options except pDo are set at their defaults (pDo is set to -1). Relative tolerance for convergence (ϵ_r) is set to 1%.

Case	N	$\{F_p\}_{p=1}^N$	$\{\alpha_p\}_{p=1}^N$	Optimum $\Delta\mathcal{E}_{loss}/RT_0$	# of Iterations	Comp. Time (s)
1	4	{17.8, 40.4, 27.7, 14.1}	{3.413, 2.146, 1.398, 1}	74.05	71	36
2	4	{30, 40, 25, 5}	{29.07, 12.81, 2.35, 1}	87.88	51	22
3	4	{5, 25, 40, 30}	{24.19, 4.599, 2.19, 1}	70.68	53	26
4	4	{3, 55, 12, 30}	{27.88, 12.28, 4.76, 1}	109.45	82	45
5	5	{20, 20, 20, 20, 20}	{39.0625, 15.625, 6.25, 2.5, 1}	100.97	2449	9396
6	5	{10, 10, 40, 30, 10}	{8.9, 5.7, 3.2, 1.55, 1}	84.35	1036	2905
7	5	{10, 20, 30, 20, 20}	{8.9, 5.7, 3.2, 1.55, 1}	84.71	1142	3462
8	5	{10, 10, 40, 30, 10}	{4.1, 3.6, 2.1, 1.42, 1}	79.14	715	2031
9	5	{20, 30, 20, 20, 10}	{4.1, 3.6, 2.1, 1.42, 1}	96.62	252	540
10	5	{15, 15, 40, 20, 10}	{10.5, 4.04, 1.76, 1.31, 1}	71.01	834	1849
11	5	{4, 6, 50, 35, 5}	{10.5, 4.04, 1.76, 1.31, 1}	69.96	415	1139
12	5	{5, 15, 25, 20, 35}	{7.98, 3.99, 3, 1.25, 1}	77.64	169	585
13	5	{25, 15, 25, 20, 20}	{13.72, 3.92, 3.267, 1.21, 1}	94.26	211	565
14	5	{5, 32.5, 25, 32.5, 5}	{11.761, 9.801, 3.63, 3, 1}	90.97	381	1021
15	5	{5, 5, 32.5, 25, 32.5}	{11.761, 9.801, 3.63, 3, 1}	67.07	410	978

specified time limit. This illustrates that model (W) performs well to changes in problem parameters, and it can be used to quickly identify attractive solutions with off-the-shelf global solvers.

5.3 Conclusion

In this work, we proposed a novel MINLP formulation for identification of thermodynamically-efficient distillation configurations, for separation of non-azeotropic multicomponent mixtures. Thermodynamic analysis is crucial, especially for work-driven systems like heat-pump assisted distillations, which also include cryogenic separations. The highlights of this work are summarized in the following. First, we proposed a new model for the space of admissible distillation configurations, and proved it to be strictly tighter than the intersection of prior formulations. Second, using techniques described in [28, 29], we formulated exergy loss calculations that do not depend on temperature of mixtures, explicitly. Third, the model in its default form has several nonlinear nonconvex equations. We proposed a simple variable elimination strategy, that collapses a system of equations describing material balance and vapor-liquid equilibrium onto a single equation. The approach reduces the number of nonconvex equations and simplifies the model significantly. Fourth, we described the properties satisfied by the derived equations, and exploited them while deriving additional valid cuts to the problem. Finally, the model was used to identify attractive configurations for shale gas separation. We have also shown through numerical examples from literature that the model performs well with changes in problem parameters. Designing distillation sequences is challenging. This model empowers process designers to quickly screen through thousands of alternatives, and identify attractive solutions worthy of further exploration.

6. FUTURE DIRECTION

In this work, we have developed novel MINLP-based approaches for systematically identifying the best distillation configuration along with its optimal operating condition for a given separation. Despite significant energy savings possible, and decades of work, the problem has continued to resist solution. By performing a careful mathematical analysis, we have addressed the key challenges and developed the first valid formulation. Through extensive computational experiments, we have demonstrated that the proposed approach improves the state-of-the-art. However, this is simply the first step towards making industrial distillations more sustainable. Below, we describe a few possible ways to build up on the current work.

6.1 Heat Integration

For various reasons (e.g., ease of operation), industrial practitioners might continue to run above-ambient separations using heat. In that case, heat integration is crucial to minimize the overall heat duty. Here, heat integration refers to utilization of the heat rejected from the condenser of one distillation column in the reboiler of another distillation column. For example, consider the configuration in Figure 6.1(a). Since the boiling point of component C is greater than that of component B , the heat duty of the condenser of the column $Q3$ (see Figure caption) can be rejected to the reboiler of the column $Q2$. This heat integration is always feasible, unless the column $Q2$ is operated at a much higher pressure than the column $Q3$. In general, whenever the lightest component in the mixture that is being condensed is the same or heavier than the heaviest component in the mixture that is being boiled, then the heat integration between the condenser and the reboiler is always feasible. Such

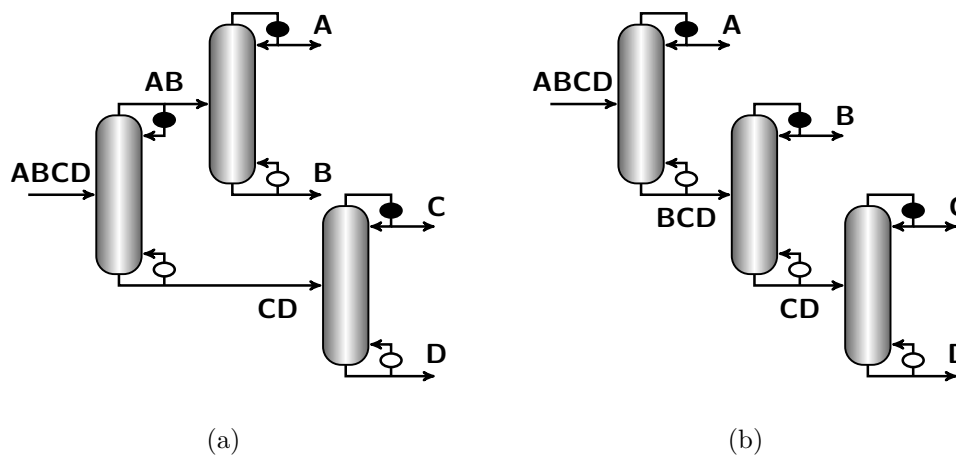


Fig. 6.1. Four-component configurations. We refer the leftmost column as $Q1$, the middle column as $Q2$ and the rightmost column as $Q3$.

heat integrations can be included in the proposed MINLP approach directly without additional constraints to check for feasibility.

Next, some heat integrations are feasible only when the composition of certain streams are favorable. For example, consider the configuration in Figure 6.1(b). In general, the temperature of the reboiler BCD will not be less than that of the condenser C . However, when the submixture BCD is very rich in B , it is possible to heat integrate the reboiler BCD and the condenser C . Mathew et al. [70] pointed out that such heat integrations can reduce the overall heat duty well below that of the fully thermally coupled configuration. An immediate question that arises is the following. How do we identify such non-intuitive heat integrations? Of course, if the temperature of the streams were available, then the feasibility can be inferred from the temperatures. However, computation of the temperature requires the solution of complex thermodynamic nonlinear and nonconvex equations. In their presence, solving the MINLP to the desired optimality gap can be challenging. A more promising approach is to use the shortcut criteria proposed by Mathew et al. [71]. As demonstrated therein, off-the-shelf solvers like BARON are able to solve the optimization problem to the desired optimality gap for specific configurations. Therefore, by incor-

porating their shortcut criteria, the proposed formulation can be extended to identify the optimal heat integrated distillation configuration.

6.2 Improvement to the Proposed Algorithm

In Chapter 4, we proposed a simple discretization-based solution procedure to solve the MINLP. Below, we describe two possible ways to further improve the algorithm.

First, consider the performance profile in Figure 4.7(a). In the beginning, the proposed algorithm solves fewer cases than BARON. The analysis of the results revealed the following. In some cases, the lower bound has reached the global optimum in the first two iterations and the algorithm did not find the optimal solution (upper bound). Nevertheless, the current algorithm solves the computationally expensive MIP relaxation again in order to find a new initial point for the local search. This can be avoided by collecting multiple points, for the local search, in each iteration using the solution pool feature of Gurobi. This improves the possibility of finding the optimal solutions very quickly, and reduces the number of iterations (and computational time) to solve the MINLP.

Second, the main inefficiency of the algorithm is that it solves the MIP relaxation from scratch at every iteration. This increases the computational time per iteration substantially beyond the sixth or seventh iteration. Since refinement of discretization does not worsen the MIP relaxation, it is worthwhile to investigate if there is a way to utilize the solution from the previous iteration to solve the next iteration more efficiently. If successful, this can reduce the computational time substantially, given that the MILP solvers currently need more than one hour per iteration beyond the sixth iteration.

6.3 Application to Design and Operation

After identifying an attractive configuration for a given separation, the next step involves a detailed design. One of the underlying assumptions in the formulation is that a distillation column is allowed to operate at pinch *i.e.*, at its minimum reflux ratio. In reality, due to economic considerations, a distillation column is always operated above the minimum reflux ratio. Because of this, the actual composition of the distillate and the residue product streams may deviate from that predicted by the MINLP framework. This is especially true when the column performs a sloppy split, since the product composition is sensitive to both the number of trays, the feed location and the actual reflux ratio. The detailed design of a configuration is usually performed using a process simulator, such as Aspen Plus. However, as described in Chapter 1, process simulators face convergence issues and often get trapped in a sub-optimal solution. It would be interesting to develop a method that guides the process simulators to obtain a detailed design by using the optimal solution obtained from the MINLP approach. Another potential direction is to explore the practicality of the proposed approach in designing controllers for optimal operation.

REFERENCES

REFERENCES

- [1] J. L. Humphrey, *Separation process technology*. McGraw-Hill (canada), 1997.
- [2] J. J. Siirola, "The impact of shale gas in the chemical industry," *AIChE Journal*, vol. 60, no. 3, pp. 810–819, 2014.
- [3] T. Ridha, Y. Li, E. Gençer, J. J. Siirola, J. T. Miller, F. H. Ribeiro, and R. Agrawal, "Valorization of shale gas condensate to liquid hydrocarbons through catalytic dehydrogenation and oligomerization," *Processes*, vol. 6, no. 9, 2018.
- [4] G. Madenoor Ramapriya, A. Selvarajah, L. E. Jimenez Cucaita, J. Huff, M. Tawarmalani, and R. Agrawal, "Short-cut methods versus rigorous methods for performance-evaluation of distillation configurations," *Industrial & Engineering Chemistry Research*, 2018.
- [5] V. H. Shah and R. Agrawal, "A matrix method for multicomponent distillation sequences," *AIChE journal*, vol. 56, no. 7, pp. 1759–1775, 2010.
- [6] R. Sargent and K. Gaminibandara, "Introduction: approaches to chemical process synthesis," 1976.
- [7] R. Agrawal, "Synthesis of distillation column configurations for a multicomponent separation," *Industrial & engineering chemistry research*, vol. 35, no. 4, pp. 1059–1071, 1996.
- [8] —, "Synthesis of multicomponent distillation column configurations," *AIChE journal*, vol. 49, no. 2, pp. 379–401, 2003.
- [9] J. A. Caballero and I. E. Grossmann, "Structural considerations and modeling in the synthesis of heat-integrated-thermally coupled distillation sequences," *Industrial & engineering chemistry research*, vol. 45, no. 25, pp. 8454–8474, 2006.
- [10] A. Giridhar and R. Agrawal, "Synthesis of distillation configurations. ii: A search formulation for basic configurations," *Computers & chemical engineering*, vol. 34, no. 1, pp. 84–95, 2010.
- [11] M. J. Andrecovich and A. Westerberg, "An milp formulation for heat-integrated distillation sequence synthesis," *AIChE journal*, vol. 31, no. 9, pp. 1461–1474, 1985.
- [12] Z. Fidkowski and L. Królikowski, "Minimum energy requirements of thermally coupled distillation systems," *AIChE Journal*, vol. 33, no. 4, pp. 643–653, 1987.
- [13] Z. T. Fidkowski and R. Agrawal, "Multicomponent thermally coupled systems of distillation columns at minimum reflux," *AIChE journal*, vol. 47, no. 12, pp. 2713–2724, 2001.

- [14] I. J. Halvorsen and S. Skogestad, "Minimum energy consumption in multicomponent distillation. 3. more than three products and generalized petlyuk arrangements," *Industrial & Engineering Chemistry Research*, vol. 42, no. 3, pp. 616–629, 2003.
- [15] Z. Fidkowski and L. Krolikowski, "Thermally coupled system of distillation columns: optimization procedure," *AIChE journal*, vol. 32, no. 4, pp. 537–546, 1986.
- [16] R. Agrawal and Z. T. Fidkowski, "New thermally coupled schemes for ternary distillation," *AIChE Journal*, vol. 45, no. 3, pp. 485–496, 1999.
- [17] J. A. Caballero and I. E. Grossmann, "Design of distillation sequences: from conventional to fully thermally coupled distillation systems," *Computers & chemical engineering*, vol. 28, no. 11, pp. 2307–2329, 2004.
- [18] J. A. Caballero, J. A. Reyes-Labarta, and I. E. Grossmann, "Synthesis of integrated distillation systems," *Computer Aided Chemical Engineering*, vol. 14, pp. 59–64, 2003.
- [19] J. A. Caballero and I. E. Grossmann, "Synthesis of complex thermally coupled distillation systems including divided wall columns," *AIChE Journal*, vol. 59, no. 4, pp. 1139–1159, 2013.
- [20] —, "Optimal synthesis of thermally coupled distillation sequences using a novel milp approach," *Computers & Chemical Engineering*, vol. 61, pp. 118–135, 2014.
- [21] M. Tawarmalani and N. V. Sahinidis, "A polyhedral branch-and-cut approach to global optimization," *Mathematical Programming*, vol. 103, no. 2, pp. 225–249, 2005.
- [22] U. Nallasivam, V. H. Shah, A. A. Shenvi, J. Huff, M. Tawarmalani, and R. Agrawal, "Global optimization of multicomponent distillation configurations: 2. enumeration based global minimization algorithm," *AIChE Journal*, vol. 62, no. 6, pp. 2071–2086, 2016.
- [23] R. Tumbalam Gooty, R. Agrawal, and M. Tawarmalani, "An minlp formulation for the optimization of multicomponent distillation configurations," *Computers & Chemical Engineering*, vol. 125, pp. 13–30, 2019.
- [24] J. Caballero and I. Grossmann, "Optimal separation sequences based on distillation: From conventional to fully thermally coupled systems," Modification of: 06:56:09, November 23. 2009, available from CyberInfrastructure for MINLP [www.minlp.org, a collaboration of Carnegie Mellon University and IBM Research] at: www.minlp.org/library/problem/index.php?i=69.
- [25] R. Tumbalam Gooty, R. Agrawal, and M. Tawarmalani, "Advances in minlp to identify energy-efficient distillation configurations," *arxiv:2010.12113*, Under Review.
- [26] H. D. Sherali and A. Alameddine, "A new reformulation-linearization technique for bilinear programming problems," *Journal of Global optimization*, vol. 2, no. 4, pp. 379–410, 1992.

- [27] R. Tumbalam Gooty, T. J. Mathew, M. Tawarmalani, and R. Agrawal, "An minlp formulation to identify thermodynamically-efficient distillation configurations," *In preparation*.
- [28] K. Glinos and M. Malone, "Net work consumption in distillation—short-cut evaluation and applications to synthesis," *Computers & chemical engineering*, vol. 13, no. 3, pp. 295–305, 1989.
- [29] R. Agrawal and D. M. Herron, "Optimal thermodynamic feed conditions for distillation of ideal binary mixtures," *AIChE journal*, vol. 43, no. 11, pp. 2984–2996, 1997.
- [30] T. J. Mathew, M. Tawarmalani, and R. Agrawal, "A tighter relation for recoveries in multicomponent distillation," Working paper.
- [31] A. Underwood, "Fractional distillation of multicomponent mixtures," *Chem. Eng. Prog.*, vol. 44, no. 8, pp. 603–614, 1948.
- [32] I. J. Halvorsen, "Minimum energy requirements in complex distillation arrangements," 2001.
- [33] I. J. Halvorsen and S. Skogestad, "Minimum energy consumption in multicomponent distillation. 1. v min diagram for a two-product column," *Industrial & Engineering Chemistry Research*, vol. 42, no. 3, pp. 596–604, 2003.
- [34] —, "Minimum energy consumption in multicomponent distillation. 2. three-product petlyuk arrangements," *Industrial & engineering chemistry research*, vol. 42, no. 3, pp. 605–615, 2003.
- [35] U. Nallasivam, V. H. Shah, A. A. Shenvi, M. Tawarmalani, and R. Agrawal, "Global optimization of multicomponent distillation configurations: 1. need for a reliable global optimization algorithm," *AIChE Journal*, vol. 59, no. 3, pp. 971–981, 2013.
- [36] R. A. Stubbs and S. Mehrotra, "A branch-and-cut method for 0-1 mixed convex programming," *Mathematical programming*, vol. 86, no. 3, pp. 515–532, 1999.
- [37] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part i—convex underestimating problems," *Mathematical programming*, vol. 10, no. 1, pp. 147–175, 1976.
- [38] N. A. Carlberg and A. W. Westerberg, "Temperature-heat diagrams for complex columns. 3. underwood's method for the petlyuk configuration," *Industrial & engineering chemistry research*, vol. 28, no. 9, pp. 1386–1397, 1989.
- [39] A. Giridhar and R. Agrawal, "Synthesis of distillation configurations: I. characteristics of a good search space," *Computers & chemical engineering*, vol. 34, no. 1, pp. 73–83, 2010.
- [40] M. R. Kılınç and N. V. Sahinidis, "Exploiting integrality in the global optimization of mixed-integer nonlinear programming problems with baron," *Optimization Methods and Software*, vol. 33, no. 3, pp. 540–562, 2018.

- [41] G. Gamrath, T. Fischer, T. Gally, A. M. Gleixner, G. Hendel, T. Koch, S. J. Maher, M. Miltenberger, B. Müller, M. E. Pfetsch, C. Puchert, D. Rehfeldt, S. Schenker, R. Schwarz, F. Serrano, Y. Shinano, S. Vigerske, D. Weninger, M. Winkler, J. T. Witt, and J. Witzig, “The scip optimization suite 3.2,” ZIB, Takustr.7, 14195 Berlin, Tech. Rep. 15-60, 2016.
- [42] V. H. Shah and R. Agrawal, “Are all thermal coupling links between multicomponent distillation columns useful from an energy perspective?” *Industrial & Engineering Chemistry Research*, vol. 50, no. 3, pp. 1770–1777, 2010.
- [43] G. Angulo, S. Ahmed, S. S. Dey, and V. Kaibel, “Forbidden vertices,” *Mathematics of Operations Research*, vol. 40, no. 2, pp. 350–360, 2014.
- [44] R. Tumbalam Gooty, P. Mobed, M. Tawarmalani, and R. Agrawal, “Optimal multicomponent distillation column sequencing: Software and case studies,” in *Computer-Aided Chemical Engineering*. Elsevier, 2018, vol. 44, pp. 223–228.
- [45] J. K. Kim and P. C. Wankat, “Quaternary distillation systems with less than $n-1$ columns,” *Industrial & engineering chemistry research*, vol. 43, no. 14, pp. 3838–3846, 2004.
- [46] V. Nadgir and Y. Liu, “Studies in chemical process design and synthesis: Part v: A simple heuristic method for systematic synthesis of initial sequences for multicomponent separations,” *AIChE journal*, vol. 29, no. 6, pp. 926–934, 1983.
- [47] F. M. Houry, *Predicting the performance of multistage separation processes*. CRC Press, 1999.
- [48] A. Del Pia and A. Khajavirad, “The multilinear polytope for acyclic hypergraphs,” *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1049–1076, 2018.
- [49] L. Gurobi Optimization, “Gurobi optimizer reference manual,” 2018. [Online]. Available: <http://www.gurobi.com>
- [50] N. A. Carlberg and A. W. Westerberg, “Temperature-heat diagrams for complex columns. 2. underwood’s method for side strippers and enrichers,” *Industrial & engineering chemistry research*, vol. 28, no. 9, pp. 1379–1386, 1989.
- [51] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Siam, 2001, vol. 2.
- [52] A. MOSEK, “Mosek modeling cookbook,” 2020.
- [53] Z. Jiang, Z. Chen, J. Huff, A. A. Shenvi, M. Tawarmalani, and R. Agrawal, “Global minimization of total exergy loss of multicomponent distillation configurations,” *AIChE Journal*, vol. 65, no. 11, p. e16737, 2019.
- [54] M. Tawarmalani, “Inclusion certificates and simultaneous convexification of functions,” 2010. [Online]. Available: http://www.optimization-online.org/DB_FILE/2010/09/2722.pdf
- [55] D. Davarnia, J.-P. P. Richard, and M. Tawarmalani, “Simultaneous convexification of bilinear functions over polytopes with application to network interdiction,” *SIAM Journal on Optimization*, vol. 27, no. 3, pp. 1801–1833, 2017.

- [56] A. D. Rikun, "A convex envelope formula for multilinear functions," *Journal of Global Optimization*, vol. 10, no. 4, pp. 425–437, 1997.
- [57] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [58] Z. Jiang, T. J. Mathew, H. Zhang, J. Huff, U. Nallasivam, M. Tawarmalani, and R. Agrawal, "Global optimization of multicomponent distillation configurations: Global minimization of total cost for multicomponent mixture separations," *Computers & Chemical Engineering*, vol. 126, pp. 249–262, 2019.
- [59] T. J. Kotas, *The exergy method of thermal plant analysis*. Elsevier, 2013.
- [60] F. B. Petlyuk, "Thermodynamically optimal method for separating multicomponent mixtures," *Int. Chem. Eng.*, vol. 5, pp. 555–561, 1965.
- [61] A. Gomez-Munoz and J. Seader, "Synthesis of distillation trains by thermodynamic analysis," *Computers & chemical engineering*, vol. 9, no. 4, pp. 311–341, 1985.
- [62] R. Agrawal and Z. T. Fidkowski, "Improved direct and indirect systems of columns for ternary distillation," *AIChE journal*, vol. 44, no. 4, pp. 823–830, 1998.
- [63] —, "Are thermally coupled distillation columns always thermodynamically more efficient for ternary distillations?" *Industrial & engineering chemistry research*, vol. 37, no. 8, pp. 3444–3454, 1998.
- [64] —, "Thermodynamically efficient systems for ternary distillation," *Industrial & engineering chemistry research*, vol. 38, no. 5, pp. 2065–2074, 1999.
- [65] E. Rév, M. Emtir, Z. Szitkai, P. Mizsey, and Z. Fonyó, "Energy savings of integrated and coupled distillation systems," *Computers & Chemical Engineering*, vol. 25, no. 1, pp. 119–140, 2001.
- [66] O. A. Flores, J. C. Cardenas, S. Hernández, and V. Rico-Ramírez, "Thermodynamic analysis of thermally coupled distillation sequences," *Industrial & engineering chemistry research*, vol. 42, no. 23, pp. 5940–5945, 2003.
- [67] Y. Demirel, "Thermodynamic analysis of separation systems," *Separation science and technology*, vol. 39, no. 16, pp. 3897–3942, 2004.
- [68] D. L. George and E. B. Bowles, "Shale gas measurement and associated issues," *Pipeline & Gas Journal*, vol. 238, no. 7, pp. 38–41, 2011.
- [69] C. He and F. You, "Shale gas processing integrated with ethylene production: novel process designs, exergy analysis, and techno-economic analysis," *Industrial & Engineering Chemistry Research*, vol. 53, no. 28, pp. 11 442–11 459, 2014.
- [70] T. J. Mathew, R. Tumbalam Gooty, M. Tawarmalani, and R. Agrawal, "110th anniversary: Thermal coupling via heat transfer: A potential route to simple distillation configurations with lower heat duty," *Industrial & Engineering Chemistry Research*, vol. 58, no. 47, pp. 21 671–21 678, 2019.

- [71] ———, “Thermal coupling via heat transfer: Shortcut criteria for feasibility of heat integration in multicomponent distillation,” *In preparation*.
- [72] Y. Crama, “Concave extensions for nonlinear 0–1 maximization problems,” *Mathematical Programming*, vol. 61, no. 1-3, pp. 53–60, 1993.
- [73] R. Rado, “Theorems on linear combinatorial topology and general measure,” *Annals of Mathematics*, pp. 228–270, 1943.

APPENDICES

A. TEST SET

The test set for computational experiments is borrowed from [39]. The current state-of-the-art methods can handle design problem involving four components. However, they are often unable to scale to five components, which are practically relevant and remains challenging. In this study, we focus on five component separations, i.e., $N = 5$.

The parameter settings are generated in the following manner. For every $a \in \{1, \dots, 2^N - 1\}$, we first construct N -digit binary representation of a , denoted as $\text{bin}(a)$. Let $\text{bin}(a)(p)$ denote the p^{th} digit of $\text{bin}(a)$. We define two sets: $\mathcal{D}_0 = \{p : \text{bin}(a)(p) = 0\}$ and $\mathcal{D}_1 = \{p : \text{bin}(a)(p) = 1\}$. $\text{bin}(a)(p) = 0$ indicates that component p is lean in the mixture, and its composition is set to 5%. On the other hand, $\text{bin}(a)(p) = 1$ indicates that component p is abundant in the mixture. We consider the case, where all abundant components are present in equal proportions. Therefore, for a given a , the feed composition $\{F_p^a\}_{p=1}^N$ is obtained as

$$F_p^a = \begin{cases} 5 & \text{if } p \in \mathcal{D}_0 \\ \frac{100-5 \times |\mathcal{D}_0|}{|\mathcal{D}_1|} & \text{if } p \in \mathcal{D}_1 \end{cases} \quad \forall \quad p \in \{1, \dots, N\} \quad (\text{A.1})$$

In a similar manner, for every $b \in \{0, \dots, 2^{N-1} - 1\}$, we first construct $(N-1)$ -digit binary representation of b . Here, $\text{bin}(b)(p) = 0$ (resp. $\text{bin}(b)(p) = 1$) indicates that the separation between component p and $p+1$ is easy (resp. difficult). We take relative volatility value of 2.5 and 1.1 for an easy and difficult separation, respectively. For a given b , expressing all relative volatilities w.r.t to the heaviest component, we have $\alpha_N^b = 1$ and

$$\alpha_p^b = \prod_{q=p}^{N-1} [2.5 (1 - \text{bin}(b)(q)) + 1.1 \text{bin}(b)(q)] \quad \forall \quad p \in \{1, \dots, N-1\} \quad (\text{A.2})$$

The parameter settings for **Case(a,b)** are then given by $N = 5$, $\{F_p^a\}_{p=1}^N$, $\{\alpha_p^b\}_{p=1}^N$, $\Phi_{1,N} = \Phi_{1,1} = \dots = \Phi_{N,N} = 1$. Since $a \in \{1, \dots, 2^N - 1\}$ and $b \in \{0, \dots, 2^{N-1} - 1\}$, total number of cases in the test set is $(2^5 - 1) \times 2^4 = 496$.

B. MISSING PROOFS AND DERIVATIONS

B.1 Proof of Proposition 4.1

Let, $\mathcal{B} = [0, 1]^n \times [0, 1]^n$. Since the set $S = \{(x, z) \in \mathcal{B} \mid z_j = \prod_{p=1}^j x_p, j = 1, \dots, n\}$ is compact, $\text{Conv}(S)$ is compact and, by Krein-Milman Theorem, is the convex hull of its extreme points. Therefore, we determine the extreme points of S , and take their disjunctive union to obtain $\text{Conv}(S)$. When $(x_2, \dots, x_n)$ in S are restricted to $(\bar{x}_2, \dots, \bar{x}_n) \in [0, 1]^{n-1}$, then the set S is convex and its extreme points are such that $x_1 \in \{0, 1\}$. Let S_1 and $\tilde{S}_2$ denote the set S restricted to $x_1 = 0$ and $x_1 = 1$, respectively, *i.e.*, $S_1 = \{(x, z) \in \mathcal{B} \mid x_1 = 0, z_j = 0, j = 1, \dots, n\}$ and $\tilde{S}_2 = \{(x, z) \in \mathcal{B} \mid x_1 = z_1 = 1, z_j = \prod_{p=2}^j x_p, j = 2, \dots, n\}$. Observe that S_1 is convex, and $\tilde{S}_2$ is nonconvex. Next, when $(x_3, \dots, x_n)$ in $\tilde{S}_2$ are restricted to $(\bar{x}_3, \dots, \bar{x}_n) \in [0, 1]^{n-2}$, then $\tilde{S}_2$ is convex and its extreme points are such that $x_2 \in \{0, 1\}$. Let S_2 and $\tilde{S}_3$ denote the set $\tilde{S}_2$ restricted to $x_2 = 0$ and $x_2 = 1$, respectively, *i.e.*, $S_2 = \{(x, z) \in \mathcal{B} \mid x_1 = z_1 = 1, x_2 = z_2 = \dots, z_n = 0\}$ and $\tilde{S}_3 = \{(x, z) \in \mathcal{B} \mid x_1 = z_1 = x_2 = z_2 = 1, z_j = \prod_{p=3}^j x_p, j = 3, \dots, n\}$. As before, S_2 is convex and $\tilde{S}_3$ is nonconvex. Repeating the argument leads to sets $S_3, \dots, S_{n+1}$, where $S_i = \{(x, z) \in \mathcal{B} \mid x_1 = z_1 = \dots = x_{i-1} = z_{i-1} = 1, x_i = z_i = \dots, z_n = 0\}$ for $i = 3, \dots, n$ and $S_{n+1} = \tilde{S}_{n+1} = \{x_1 = z_1 = \dots, x_n = z_n = 1\}$. The sets S_1 through S_{n+1} contain the extreme points of convex hull of S . Therefore, $\text{Conv}(S) = \text{Conv}(S_1 \cup S_2 \cup \dots \cup S_{n+1})$, where $S_1 \cup S_2 \cup \dots \cup S_{n+1}$ is given below

$$\left[\begin{array}{l} x_1 = 0 \\ z_1 = \dots = z_n = 0 \\ 0 \leq x_j \leq 1, j = 2, \dots, n \end{array} \right] \bigvee_{i=2}^n \left[\begin{array}{l} x_1 = \dots = x_{i-1} = 1, x_i = 0 \\ z_1 = \dots = z_{i-1} = 1 \\ z_i = \dots = z_n = 0 \\ 0 \leq x_j \leq 1, j = i+1, \dots, n \end{array} \right] \bigvee \left[\begin{array}{l} x_1 = \dots = x_n = 1 \\ z_1 = \dots = z_n = 1 \end{array} \right].$$

Application of disjunctive programming technique leads to

$$\text{Conv}(S) = \left\{ \begin{array}{ll} x_j^i = \lambda^i, & \text{for } j = 1, \dots, i-1; \quad i = 2, \dots, n+1 \\ x_i^i = 0, & \text{for } i = 1, \dots, n \\ 0 \leq x_j^i \leq \lambda^i, & \text{for } j = i+1, \dots, n; \quad i = 1, \dots, n-1 \\ x_j = \sum_{i=1}^{n+1} x_j^i, & \text{for } j = 1, \dots, n \\ z_j = \sum_{i=j+1}^{n+1} \lambda^i, & \text{for } j = 1, \dots, n \\ \sum_{i=1}^{n+1} \lambda^i = 1, \lambda^i \geq 0, & \text{for } i = 1, \dots, n+1 \end{array} \right\}, \quad (\text{B.1})$$

where $\{x_j^i\}_{i=1}^{n+1}$ are to be regarded as linearization of $x_j \cdot \lambda^i$. We eliminate $\{x_j^i\}_{i=j}^{n+1}$ by direct substitution (see (B.1)). This leads to $x_j = \sum_{i=1}^{j-1} x_j^i + \sum_{i=j+1}^{n+1} \lambda^i$, or $\sum_{i=1}^{j-1} x_j^i \leq x_j - \sum_{i=j+1}^{n+1} \lambda^i \leq \sum_{i=1}^{j-1} x_j^i$, where $\{x_j^i\}_{i=1}^{j-1}$ are constrained by $0 \leq x_j^i \leq \lambda^i$. Now, using Fourier-Motzkin elimination, we eliminate $\{x_j^i\}_{i=1}^{j-1}$ to obtain $0 \leq x_j - \sum_{i=j+1}^{n+1} \lambda^i \leq \sum_{i=1}^{j-1} \lambda^i$, or $\sum_{i=j+1}^{n+1} \lambda^i \leq x_j \leq \sum_{i=1}^{j-1} \lambda^i + \sum_{i=j+1}^{n+1} \lambda^i = 1 - \lambda^j$. This transforms (B.1) to

$$\text{Conv}(S) = \left\{ \begin{array}{ll} \sum_{i=j+1}^{n+1} \lambda^i \leq x_j \leq 1 - \lambda^j, & \text{for } j = 1, \dots, n \\ z_j = \sum_{i=j+1}^{n+1} \lambda^i, & \text{for } j = 1, \dots, n \\ \sum_{i=1}^{n+1} \lambda^i = 1, \lambda^i \geq 0, & \text{for } i = 1, \dots, n+1 \end{array} \right\}. \quad (\text{B.2})$$

Next, we determine λ^i in terms of z_j . From $z_j = \sum_{i=j+1}^{n+1} \lambda^i$ for $j = 1, \dots, n$ and $\sum_{i=1}^{n+1} \lambda^i = 1$, $z_n = \lambda^{n+1}$, $z_{n-1} = \lambda^n + \lambda^{n+1}$ or $z_{n-1} - z_n = \lambda^n$, $z_{n-2} = \lambda^{n-1} + \lambda^n + \lambda^{n+1}$ or $z_{n-2} - z_{n-1} = \lambda^{n-1}$, $\dots$, $z_1 - z_2 = \lambda^2$, and $\lambda^1 = 1 - \sum_{i=2}^{n+1} \lambda^i = 1 - z_1$. Using these relations, we eliminate λ^i variables from (B.2) to obtain

$$\text{Conv}(S) = \left\{ \begin{array}{ll} z_1 \leq x_1 \leq z_1 \\ z_j \leq x_j \leq 1 - z_{j-1} + z_j, & \text{for } j = 2, \dots, n \\ z_n \geq 0, (1 - z_1) \geq 0, \\ z_{j-1} - z_j \geq 0, & \text{for } j = 2, \dots, n \end{array} \right\}. \quad (\text{B.3})$$

Observe that the same set of inequalities result from recursive McCormick relaxation of $z_j = z_{j-1} \cdot x_j$ for $j = 2, \dots, n$. Therefore, the convex hull of set S can be constructed by a recursive application of McCormick procedure on $z_j = z_{j-1} \cdot x_j$, $j = 2, \dots, n$. $\square$

B.2 Proof of Remark 4.2

We show the proof for $\tau_{i,k,j}$ variables, and the proof for $\beta_{i,l,j}$ variables is similar. By Remark 4.1, the convex hull of $\nu_{i,k,j} = \prod_{n=k}^j (1 - \zeta_{i,n})$ over $(\zeta_{i,k}, \dots, \zeta_{i,j}) \in [0, 1]^{j-k+1}$, given by

$$\nu_{i,k,j} \geq \max\{0, -\zeta_{i,k} - \dots - \zeta_{i,j} + 1\} \quad (\text{B.4a})$$

$$\nu_{i,k,j} \leq \min\{1 - \zeta_{i,k}, \dots, 1 - \zeta_{i,j}\}, \quad (\text{B.4b})$$

is implied from (A3), for every $[i, j] \in \mathcal{P}$, $\llbracket k \rrbracket_i^{j-1}$. We use the above inequalities, in addition to (A2) and (A3), for the proof. We consider two cases: $k+1 < j$ and $k+1 = j$. When, $k+1 < j$, the convex hull of $\tau_{i,k,j} = \zeta_{i,k}(1 - \zeta_{i,k+1}) \dots (1 - \zeta_{i,j-1})\zeta_{i,j}$ over $(\zeta_{i,k}, \dots, \zeta_{i,j}) \in [0, 1]^{j-k+1}$ is given by [72]

$$\tau_{i,k,j} \geq 0, \quad (\text{B.5a})$$

$$\tau_{i,k,j} \geq \zeta_{i,k} - \zeta_{i,k+1} - \dots - \zeta_{i,j-1} + \zeta_{i,j} - 1, \quad (\text{B.5b})$$

$$\tau_{i,k,j} \leq \zeta_{i,k}, \quad (\text{B.5c})$$

$$\tau_{i,k,j} \leq 1 - \zeta_{i,n}, \quad \llbracket n \rrbracket_{k+1}^{j-1}, \quad (\text{B.5d})$$

$$\tau_{i,k,j} \leq \zeta_{i,j}. \quad (\text{B.5e})$$

On the other hand, when $k+1 = j$, the convex hull of $\tau_{i,k,j} = \tau_{i,j-1,j} = \zeta_{i,j-1}\zeta_{i,j}$ over $(\zeta_{i,j-1}, \zeta_{i,j}) \in [0, 1]^2$ is given by

$$\tau_{i,j-1,j} \geq \max\{0, \zeta_{i,j-1} + \zeta_{i,j} - 1\}, \quad (\text{B.6a})$$

$$\tau_{i,j-1,j} \leq \min\{\zeta_{i,j-1}, \zeta_{i,j}\}. \quad (\text{B.6b})$$

In the following, we present the proof only for $k+1 < j$, and point out that the proof for the case $k+1 = j$ is similar.

(B.5a): From (A3), $\nu_{i,k,j-1} + \nu_{i,k+1,j} - \nu_{i,k+1,j-1} \leq \nu_{i,k,j} \implies 0 \leq \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j} \stackrel{(A2)}{=} \tau_{i,k,j}$.

$$\begin{aligned}
\text{(B.5b): } \tau_{i,k,j} &\stackrel{\text{(A2)}}{=} \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j} \stackrel{\text{(B.4a)}}{\geq} -\zeta_{i,k+1} - \cdots - \zeta_{i,j-1} + \\
&\stackrel{\text{(B.4b)}}{\geq} -\zeta_{i,k+1} - \cdots - \zeta_{i,j-1} + 1 - (1 - \zeta_{i,k}) - (1 - \zeta_{i,j-1}) = \\
&\zeta_{i,k} - \zeta_{i,k+1} - \cdots - \zeta_{i,j-1} + \zeta_{i,j} - 1.
\end{aligned}$$

$$\begin{aligned}
\text{(B.5c): } \tau_{i,k,j} &\stackrel{\text{(A2)}}{=} \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j} \stackrel{\text{(A3)}}{\leq} \nu_{i,k+1,j-1} - \nu_{i,k,j-1} \\
&\stackrel{\text{(A3)}}{\leq} 1 - \nu_{i,k,k} \stackrel{\text{(A3)}}{=} \zeta_{i,k}.
\end{aligned}$$

$$\begin{aligned}
\text{(B.5d): } \tau_{i,k,j} &\stackrel{\text{(A2)}}{=} \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j} \stackrel{\text{(A3)}}{\leq} \nu_{i,k+1,j-1} - \nu_{i,k,j-1} \stackrel{\text{(B.4a)}}{\leq} \\
&\stackrel{\text{(B.4b)}}{\leq} 1 - \zeta_{i,n}, \text{ for } k+1 \leq n \leq j-1.
\end{aligned}$$

$$\begin{aligned}
\text{(B.5e): } \tau_{i,k,j} &\stackrel{\text{(A2)}}{=} \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j} \stackrel{\text{(A3)}}{\leq} \nu_{i,k+1,j-1} - \nu_{i,k+1,j} \\
&\stackrel{\text{(A3)}}{\leq} 1 - \nu_{i,j,j} \stackrel{\text{(A3)}}{=} \zeta_{i,j}. \quad \square
\end{aligned}$$

B.3 Proof of Proposition 4.6

Definition B.1. Let, $\mathcal{D} = (V, A)$ be a digraph and $b \in \mathbb{R}^{|V|}$. A function $f : A \rightarrow \mathbb{R}$ is called as b -transshipment if $\text{excess}_f(v_i) := f\langle\delta^{\text{in}}(v_i)\rangle - f\langle\delta^{\text{out}}(v_i)\rangle = b(v_i) \ \forall v_i \in V$, where $\delta^{\text{in}}(v_i) \subseteq A$ (resp. $\delta^{\text{out}}(v_i) \subseteq A$) is the set of all arcs entering (resp. leaving) the vertex v_i , and $f\langle\delta(v_i)\rangle := \sum_{a \in \delta(v_i)} f(a)$. In our case, the function $f(a)$ evaluates the flow along the arc a .

Lemma B.1 ([73]). Let $\mathcal{D} = (V, A)$ be a digraph, and let $b : V \rightarrow \mathbb{R}$ with $b\langle V \rangle = 0$. Then there exists a b -transshipment $f \geq \mathbf{0}$ if and only if $b\langle U \rangle \leq 0$ for each $U \subseteq V$ with $\delta^{\text{in}}(U) = \emptyset$.

We now use Lemma B.1 to prove Proposition 4.6.

Consider the digraph $\mathcal{D} = (V, A)$, where $V = D_6 \cup D_7$ and $A = (D_6 \times D_7) \setminus \{(N+1, 0)\}$ (see §4.2.2 and Figure 4.3 for definition of D_6 and D_7). We have discarded the arc from $N+1 \in D_6$ to $0 \in D_7$, because the flow along that arc is zero (see (4.10)).

Observe that, for every $n \in D_6$, $b(n) = \text{excess}_\psi(n) = -\sum_{m \in D_7} \psi_{i,n,m,j} = -\tau_{i,j,n}$ (see Figure 4.3). Similarly, for every $m \in D_7$, $b(m) = \text{excess}_\psi(m) = \sum_{n \in D_6} \psi_{i,n,m,j} = \beta_{m,i,j}$. Then, $b\langle V \rangle = \sum_{n=j+1}^{N+1} b(n) + \sum_{m=0}^{i-1} b(m) = -\sum_{n=j+1}^{N+1} \tau_{i,j,n} + \sum_{m=0}^{i-1} \beta_{m,i,j} = 0$ (from definition of $S_{i,j}$). From Lemma B.1, a b -transshipment $\psi \geq \mathbf{0}$ exists if and only if $b\langle U \rangle \leq 0$ for each $U \subseteq V$ with $\delta^{\text{in}}(U) = \emptyset$. For every $U \subseteq D_6 \subset V$, $b\langle U \rangle \leq 0$ is satisfied trivially. On the other hand, U cannot be chosen to be a subset of D_7 , because for every $U \subseteq D_7$, $\delta^{\text{in}}(U) \neq \emptyset$. Therefore, in order to derive non-trivial inequalities, we must choose subsets of V containing vertices of both D_6 and D_7 .

Let $U = (D_6 \setminus \{N+1\}) \cup \{0\}$. Note that $\delta^{\text{in}}(U) = \emptyset$. Then, a b -transshipment $\psi \geq \mathbf{0}$ exists if and only if $b\langle U \rangle = -\sum_{n=j+1}^N \tau_{i,j,n} + \beta_{0,i,j} \leq 0$, or

$$\beta_{0,i,j} \leq \sum_{n=j+1}^N \tau_{i,j,n}. \quad (\text{B.7})$$

It can be verified that for every other subset $U \subseteq V$ satisfying $\delta^{\text{in}}(U) = \emptyset$, the inequality ensuring $b\langle U \rangle \leq 0$ is implied from $\sum_{m=0}^{i-1} \beta_{m,i,j} = \sum_{n=j+1}^{N+1} \tau_{i,j,n}$. Therefore,

$$\text{proj}_{(\tau,\beta)}(S_{i,j}) = \left\{ (\text{B.7}); \sum_{m=0}^{i-1} \beta_{m,i,j} = \sum_{n=j+1}^{N+1} \tau_{i,j,n}; \tau_{i,j,n} \geq 0, \llbracket n \rrbracket_{j+1}^{N+1}; \beta_{m,i,j} \geq 0, \llbracket m \rrbracket_0^{i-1} \right\}. \quad (\text{B.8})$$

Indeed, $\psi_{i,n,m,j}$ can be defined to verify that (B.8) is the projection of $S_{i,j}$.

Def1: Define $\psi_{i,N+1,0,j} = 0$.

Def2: For $1 \leq m \leq i-1$, define

$$\psi_{i,N+1,m,j} = \begin{cases} \tau_{i,j,N+1} \cdot \frac{\beta_{m,i,j}}{\sum_{m=1}^{i-1} \beta_{m,i,j}}, & \text{if } \sum_{m=1}^{i-1} \beta_{m,i,j} > 0 \\ 0, & \text{if } \sum_{m=1}^{i-1} \beta_{m,i,j} = 0. \end{cases}$$

Since $\sum_{m=0}^{i-1} \beta_{m,i,j} = \sum_{n=j+1}^{N+1} \tau_{i,j,n}$, (B.7) implies $\tau_{i,j,N+1} \leq \sum_{m=1}^{i-1} \beta_{m,i,j}$. Then, the above definition guarantees that $\psi_{i,N+1,m,j} \leq \beta_{m,i,j}$ for every $1 \leq m \leq i-1$, and $\sum_{m=1}^{i-1} \psi_{i,N+1,m,j} = \tau_{i,j,N+1}$.

Def3: For $j+1 \leq n \leq N$, define

$$\psi_{i,n,0,j} = \begin{cases} \beta_{0,i,j} \cdot \frac{\tau_{i,j,n}}{\sum_{n=j+1}^N \tau_{i,j,n}}, & \text{if } \sum_{n=j+1}^N \tau_{i,j,n} > 0 \\ 0, & \text{if } \sum_{n=j+1}^N \tau_{i,j,n} = 0. \end{cases}$$

Because $\beta_{0,i,j} \leq \sum_{n=j+1}^N \tau_{i,j,n}$ (see (B.7)), the above definition guarantees that $\psi_{i,n,0,j} \leq \tau_{i,j,n}$ for every $j+1 \leq n \leq N$, and $\sum_{n=j+1}^N \psi_{i,n,0,j} = \beta_{0,i,j}$.

Def4: For every $1 \leq m \leq i-1$ and $j+1 \leq n \leq N$, define

$$\psi_{i,n,m,j} = \begin{cases} \frac{(\tau_{i,j,n} - \psi_{i,n,0,j}) \cdot (\beta_{m,i,j} - \psi_{i,N+1,m,j})}{\sum_{n=j+1}^N (\tau_{i,j,n} - \psi_{i,n,0,j})}, & \text{if } \sum_{n=j+1}^N (\tau_{i,j,n} - \psi_{i,n,0,j}) \\ 0, & \text{if } \sum_{n=j+1}^N (\tau_{i,j,n} - \psi_{i,n,0,j}) = 0. \end{cases}$$

Since $(\beta_{m,i,j} - \psi_{i,N+1,m,j}) \geq 0$ (see **Def2**) and $(\tau_{i,j,n} - \psi_{i,n,0,j}) \geq 0$ (see **Def3**), the above definition guarantees $\psi_{i,n,m,j} \geq 0$ for every $1 \leq m \leq i-1$ and $j+1 \leq n \leq N$. Next, it can be shown that $\sum_{n=j+1}^N (\tau_{i,j,n} - \psi_{i,n,0,j}) = \sum_{m=1}^{i-1} (\beta_{m,i,j} - \psi_{i,N+1,m,j})$ from $\sum_{n=j+1}^{N+1} \tau_{i,j,n} = \sum_{m=0}^{i-1} \beta_{m,i,j}$, $\sum_{m=1}^{i-1} \psi_{i,N+1,m,j} = \tau_{i,j,N+1}$ (see **Def2**), and $\sum_{n=j+1}^N \psi_{i,n,0,j} = \beta_{0,i,j}$ (see **Def3**). Then, the above definition guarantees that $\sum_{n=j+1}^N \psi_{i,n,m,j} = \beta_{m,i,j} - \psi_{i,N+1,m,j}$ and $\sum_{m=1}^{i-1} = \tau_{i,j,n} - \psi_{i,n,0,j}$.

B.4 Derivation of $\text{Conv}(\mathcal{F}_p)$

Let $X := \{(f_p^{\text{in}}, f_p^{\text{rs}}, f_p^{\text{ss}}) \in [0, F_p]^3 \mid f_p^{\text{in}} = f_p^{\text{rs}} + f_p^{\text{ss}}\}$. Then, the extreme points of the polytope X are $v^1 = (0, 0, 0)$, $v^2 = (F_p, F_p, 0)$ and $v^3 = (F_p, 0, F_p)$. From Proposition 4.9, the convex hull of $\mathcal{F}_p$ is obtained as $\text{Conv}(\mathcal{F}_p) = \text{proj}_{(f, \theta, H, \underline{f}\theta)} \{(\text{B.9})\}$, where

$$w^i \geq T_p^*(\lambda^i, \theta^i), \quad i = 1, 2, 3 \quad (\text{B.9a})$$

$$w^i \leq \lambda^i T_p(\theta^{\text{lo}}) + \left[\frac{T_p(\theta^{\text{up}}) - T_p(\theta^{\text{lo}})}{\theta^{\text{up}} - \theta^{\text{lo}}} \right] (\theta^i - \lambda^i \theta^{\text{lo}}), \quad i = 1, 2, 3 \quad (\text{B.9b})$$

$$\lambda^i \theta^{\text{lo}} \leq \theta^i \leq \lambda^i \theta^{\text{up}}, \quad i = 1, 2, 3 \quad (\text{B.9c})$$

$$H_p^{\text{in}} = F_p w^2 + F_p w^3, \quad H_p^{\text{rs}} = F_p w^2, \quad H_p^{\text{ss}} = F_p w^3, \quad (\text{B.9d})$$

$$\underline{f}_p^{\text{in}} = F_p \theta^2 + F_p \theta^3, \quad \underline{f}_p^{\text{rs}} = F_p \theta^2, \quad \underline{f}_p^{\text{ss}} = F_p \theta^3, \quad (\text{B.9e})$$

$$f_p^{\text{in}} = F_p \lambda^2 + F_p \lambda^3, \quad f_p^{\text{rs}} = F_p \lambda^2, \quad f_p^{\text{ss}} = F_p \lambda^3, \quad (\text{B.9f})$$

$$w = w^1 + w^2 + w^3, \quad \theta = \theta^1 + \theta^2 + \theta^3, \quad (\text{B.9g})$$

$$\lambda^1 + \lambda^2 + \lambda^3 = 1, \quad \lambda^1, \lambda^2, \lambda^3 \geq 0. \quad (\text{B.9h})$$

We solve linear equations and obtain auxiliary variables in terms of problem variables as $(\lambda^2, \theta^2, w^2) = (f_p^{\text{rs}}/F_p, \underline{f}\theta_p^{\text{rs}}/F_p, H_p^{\text{rs}}/F_p)$, $(\lambda^3, \theta^3, w^3) = (f_p^{\text{ss}}/F_p, \underline{f}\theta_p^{\text{ss}}/F_p, H_p^{\text{ss}}/F_p)$, $(\lambda^1, \theta^1, w^1) = (1 - \lambda^2 - \lambda^3, \theta - \theta^2 - \theta^3, w - w^2 - w^3) = ((F_p - f_p^{\text{in}})/F_p, (F_p\theta - \underline{f}\theta_p^{\text{in}})/F_p, (F_pw - H_p^{\text{in}})/F_p)$ (from first equation in (B.9d), (B.9e), and (B.9f)). Using these relations, all variables can be eliminated from the hull description, except w , which is constrained by

$$T_p^*(\lambda^1, \theta^1) \leq \frac{F_pw - H_p^{\text{in}}}{F_p} \leq \lambda^1 T_p(\theta^{\text{lo}}) + \left(\frac{T_p(\theta^{\text{up}}) - T_p(\theta^{\text{lo}})}{\theta^{\text{up}} - \theta^{\text{lo}}} \right) (\theta^1 - \lambda^1 \theta^{\text{lo}}).$$

We eliminate w using Fourier-Motzkin elimination to obtain $T_p^*(\lambda^1, \theta^1) \leq \lambda^1 T_p(\theta^{\text{lo}}) + \left[\frac{T_p(\theta^{\text{up}}) - T_p(\theta^{\text{lo}})}{\theta^{\text{up}} - \theta^{\text{lo}}} \right] (\theta^1 - \lambda^1 \theta^{\text{lo}})$. The resulting constraint is redundant, so we do not impose it explicitly. This leads to the convex hull description described in §4.3.2.

B.5 Proof of Corollary 4.2

Here, x lies in the polytope $x^{\text{lo}} \leq x \leq x^{\text{up}}$, whose extreme points are x^{lo} and x^{up} . Application of Proposition 4.9 yields $\text{Conv}(S) = \text{proj}_{(x, y, z, \underline{xy})} \{(x, y, z, \underline{xy}, w, y^1, y^2, w^1, w^2, \lambda^1, \lambda^2)\}$

$$w^1 \geq g^*(\lambda^1, y^1), \quad w^2 \geq g^*(\lambda^2, y^2), \quad (\text{B.10a})$$

$$w^i \leq \lambda^i g(y^{\text{lo}}) + \left[\frac{g(y^{\text{up}}) - g(y^{\text{lo}})}{y^{\text{up}} - y^{\text{lo}}} \right] (y^i - \lambda^i y^{\text{lo}}), \quad i = 1, 2, \quad (\text{B.10b})$$

$$\lambda^1 y^{\text{lo}} \leq y^1 \leq \lambda^1 y^{\text{up}}, \quad \lambda^2 y^{\text{lo}} \leq y^2 \leq \lambda^2 y^{\text{up}}, \quad (\text{B.10c})$$

$$z = x^{\text{lo}} w^1 + x^{\text{up}} w^2, \quad \underline{xy} = x^{\text{lo}} y^1 + x^{\text{up}} y^2, \quad w = w^1 + w^2, \quad (\text{B.10d})$$

$$y = y^1 + y^2, \quad x = x^{\text{lo}} \lambda^1 + x^{\text{up}} \lambda^2, \quad \lambda^1 + \lambda^2 = 1, \quad \lambda^1, \lambda^2 \geq 0. \quad (\text{B.10e})$$

We remove the equality $w = w^1 + w^2$ to project out w . Solving the linear equations yields $\lambda^1 = (x^{\text{up}} - x)/(x^{\text{up}} - x^{\text{lo}})$, $\lambda^2 = (x - x^{\text{lo}})/(x^{\text{up}} - x^{\text{lo}})$, $y^1 = (x^{\text{up}} y - \underline{xy})/(x^{\text{up}} - x^{\text{lo}})$ and $y^2 = (\underline{xy} - x^{\text{lo}} y)/(x^{\text{up}} - x^{\text{lo}})$. Using these equations, we substitute out auxiliary variables y^1 , y^2 , λ^1 and λ^2 . Finally, eliminating variables w^1 and w^2 using Fourier-Motzkin elimination yields the convex hull description in the Proposition. The outer-approximation of the convex hull follows directly from Remark 4.7. $\square$

B.6 Proof of Proposition 4.12

We assume w.l.o.g that

$$\theta^{\text{lo}} \leq \alpha_1 - F_1/H_1^{\text{up}}. \quad (\text{B.11})$$

Otherwise, we update $F_1 = H_1^{\text{up}}(\alpha_1 - \theta^{\text{lo}})$ (See Figure B.1).

We begin by determining the extreme points of the convex hull of $\mathcal{H}_1$. When θ is restricted to $\bar{\theta} \in [\theta^{\text{lo}}, \alpha_1]$, the set $\mathcal{H}_1 = \{(f_1, \theta, H_1, \underline{f\theta}_1) \mid 0 \leq f_1 \leq \min\{F_1, H_1^{\text{up}}(\alpha_1 - \bar{\theta})\}, \theta = \bar{\theta}, \underline{f\theta}_1 = f_1 \cdot \bar{\theta}, H_1 = \{f_1/(\alpha_1 - \bar{\theta}), \text{ if } \bar{\theta} < \alpha_1; H_1 \in [0, H_1^{\text{up}}], \text{ if } \bar{\theta} = \alpha_1\}\}$ can be expressed as an affine transform of $0 \leq f_1 \leq \min\{F_1, H_1^{\text{up}}(\alpha_1 - \bar{\theta})\}$ whose extreme points are $f_1 \in \{0, \min\{F_1, H_1^{\text{up}}(\alpha_1 - \bar{\theta})\}\}$. Therefore, the extreme points of $\text{Conv}(\mathcal{H}_1)$ are contained in the set of points where $f_1 = 0$, or $f_1 = F_1$ and $\theta^{\text{lo}} \leq \theta \leq (\alpha_1 - F_1/H_1^{\text{up}})$, or $f_1 = H_1^{\text{up}}(\alpha_1 - \theta)$ and $(\alpha_1 - F_1/H_1^{\text{up}}) \leq \theta \leq \alpha_1$ (see Figure B.1). Let,

1. S^a be $\mathcal{H}_1$ restricted to $f_1 = 0$ i.e., $S^a = \{(f_1, \theta, H_1, \underline{f\theta}_1) \mid f_1 = 0, \theta^{\text{lo}} \leq \theta \leq \alpha_1, \underline{f\theta}_1 = 0, H_1 = 0 \text{ if } \theta < \alpha_1; H_1 \in [0, H_1^{\text{up}}], \text{ if } \theta = \alpha_1\}$ (see Figure B.1).
2. S^b be $\mathcal{H}_1$ restricted to $f_1 = H_1^{\text{up}}(\alpha_1 - \theta)$ and $(\alpha_1 - F_1/H_1^{\text{up}}) \leq \theta \leq \alpha_1$ i.e., $S^b = \{(f_1, \theta, H_1, \underline{f\theta}_1) \mid f_1 = H_1^{\text{up}}(\alpha_1 - \theta), (\alpha_1 - F_1/H_1^{\text{up}}) \leq \theta \leq \alpha_1, H_1 = H_1^{\text{up}}, \underline{f\theta}_1 = H_1^{\text{up}}(\alpha_1 - \theta)\theta\}$ (see Figure B.1).
3. S^c be $\mathcal{H}_1$ restricted to $f_1 = F_1$ and $\theta^{\text{lo}} \leq \theta \leq (\alpha_1 - F_1/H_1^{\text{up}})$ i.e., $S^c = \{(f_1, \theta, H_1, \underline{f\theta}_1) \mid f_1 = F_1, \theta^{\text{lo}} \leq \theta \leq (\alpha_1 - F_1/H_1^{\text{up}}), H_1 = F_1/(\alpha_1 - \theta), \underline{f\theta}_1 = F_1 \cdot \theta\}$ (see Figure B.1).

By Krein-Milman theorem, $\text{Conv}(\mathcal{H}_1) = \text{Conv}(S^a \cup S^b \cup S^c) = \text{Conv}(\text{Conv}(S^a) \cup \text{Conv}(S^b) \cup \text{Conv}(S^c))$, where

$$\text{Conv}(S^a) = \left\{ (f_1, \theta, H_1, \underline{f\theta}_1) \left| \begin{array}{l} f_1 = 0, \underline{f\theta}_1 = 0 \\ 0 \leq H_1 \leq H_1^{\text{up}} \left(\frac{\theta - \theta^{\text{lo}}}{\alpha_1 - \theta^{\text{lo}}} \right) \\ \theta^{\text{lo}} \leq \theta \leq \alpha_1 \end{array} \right. \right\}, \quad (\text{B.12})$$

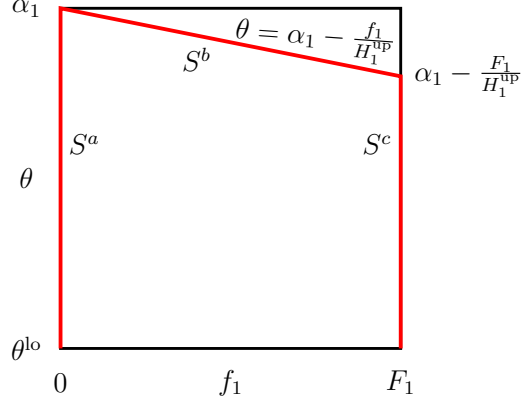


Fig. B.1. (f_1, θ) domain for §B.6. The extreme points of $\text{Conv}(\mathcal{H}_1)$ are contained in points in red.

$$\text{Conv}(S^b) = \left\{ (f_1, \theta, H_1, \underline{f\theta}_1) \left| \begin{array}{l} f_1 = H_1^{\text{up}}(\alpha_1 - \theta), H_1 = H_1^{\text{up}} \\ H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) (\alpha_1 - \theta) \leq \underline{f\theta}_1 \\ \underline{f\theta}_1 \leq H_1^{\text{up}} \theta (\alpha_1 - \theta) \\ \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \leq \theta \leq \alpha_1 \end{array} \right. \right\}, \quad (\text{B.13})$$

$$\text{Conv}(S^c) = \left\{ (f_1, \theta, H_1, \underline{f\theta}_1) \left| \begin{array}{l} f_1 = F_1, \underline{f\theta}_1 = F_1 \cdot \theta \\ \frac{F_1}{\alpha_1 - \theta} \leq H_1 \\ H_1 \leq \frac{F_1}{\alpha_1 - \theta^{\text{lo}}} + \frac{H_1^{\text{up}}}{\alpha_1 - \theta^{\text{lo}}} (\theta - \theta^{\text{lo}}) \\ \theta^{\text{lo}} \leq \theta \leq \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \end{array} \right. \right\}. \quad (\text{B.14})$$

Disjunctive union of $\text{Conv}(S^a)$, $\text{Conv}(S^b)$ and $\text{Conv}(S^c)$ leads to (4.39). $\square$

B.7 Relaxation of (4.39)

Since (4.39) introduces many variables, we derive a relaxation of $\text{Conv}(\mathcal{H}_1)$ instead. Let, $\bar{\theta}^r \in [\theta^{\text{lo}}, \alpha_1]$ for $r = 1, \dots, R$. First, we outer approximate $\text{Conv}(S^b)$ and $\text{Conv}(S^c)$ as shown below:

$$\text{Conv}_{OA}(S^2) = \left\{ (f_1, \theta, H_1, \underline{f}\theta_1) \left| \begin{array}{l} f_1 = H_1^{\text{up}}(\alpha_1 - \theta), H_1 = H_1^{\text{up}} \\ H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) (\alpha_1 - \theta) \leq \underline{f}\theta_1 \\ \underline{f}\theta_1 \leq \min \left\{ H_1^{\text{up}} \bar{\theta}^r (\alpha_1 - \bar{\theta}^r) + H_1^{\text{up}} (\alpha_1 - 2\bar{\theta}^r) (\theta - \bar{\theta}^r) \right\}_{r=1}^R \\ \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \leq \theta \leq \alpha_1 \end{array} \right. \right\},$$

$$\text{Conv}_{OA}(S^3) = \left\{ (f_1, \theta, H_1, \underline{f}\theta_1) \left| \begin{array}{l} f_1 = F_1, \underline{f}\theta_1 = F_1 \cdot \theta \\ H_1 \geq \max \left\{ \frac{F_1}{\alpha_1 - \bar{\theta}^r} + \frac{F_1}{(\alpha_1 - \bar{\theta}^r)^2} (\theta - \bar{\theta}^r) \right\}_{r=1}^R \\ H_1 \leq \frac{F_1}{\alpha_1 - \theta^{\text{lo}}} + \frac{H_1^{\text{up}}}{\alpha_1 - \theta^{\text{lo}}} (\theta - \theta^{\text{lo}}) \\ \theta^{\text{lo}} \leq \theta \leq \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \end{array} \right. \right\}.$$

Next, we take the disjunctive union of $\text{Conv}(S^a)$, $\text{Conv}_{OA}(S^b)$ and $\text{Conv}_{OA}(S^c)$ to obtain

$$H_1 \geq H_1^{\text{up}} \lambda^b + \max \left\{ \frac{F_1 \lambda^c}{\alpha_1 - \bar{\theta}^r} + \frac{F_1}{(\alpha_1 - \bar{\theta}^r)^2} (\theta^3 - \bar{\theta}^r \lambda^c) \right\}_{r=1}^R \quad (\text{B.15a})$$

$$H_1 \leq H_1^{\text{up}} \left(\frac{\theta^a - \theta^{\text{lo}} \lambda^a}{\alpha_1 - \theta^{\text{lo}}} \right) + H_1^{\text{up}} \lambda^b + \frac{F_1 \lambda^c}{\alpha_1 - \theta^{\text{lo}}} + \frac{H_1^{\text{up}}}{\alpha_1 - \theta^{\text{lo}}} (\theta^c - \theta^{\text{lo}} \lambda^c) \quad (\text{B.15b})$$

$$\underline{f}\theta_1 \geq H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c \quad (\text{B.15c})$$

$$\underline{f}\theta_1 \leq \min \left\{ H_1^{\text{up}} \bar{\theta}^r (\alpha_1 - \bar{\theta}^r) \lambda^b + H_1^{\text{up}} (\alpha_1 - 2\bar{\theta}^r) (\theta^b - \bar{\theta}^r \lambda^b) \right\}_{r=1}^R + F_1 \theta^c \quad (\text{B.15d})$$

$$f_1 = H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) + F_1 \lambda^c \quad (\text{B.15e})$$

$$\theta = \theta^a + \theta^b + \theta^c \quad (\text{B.15f})$$

$$\theta^{\text{lo}} \lambda^a \leq \theta^a \leq \alpha_1 \lambda^a \quad (\text{B.15g})$$

$$\left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \lambda^b \leq \theta^b \leq \alpha_1 \lambda^b \quad (\text{B.15h})$$

$$\theta^{\text{lo}} \lambda^c \leq \theta^c \leq \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \lambda^c \quad (\text{B.15i})$$

$$\lambda^a + \lambda^b + \lambda^c = 1, \quad \lambda^a, \lambda^b, \lambda^c \geq 0 \quad (\text{B.15j})$$

In the following, we derive relaxed version of each inequality in terms of problem variables.

$$H_1 \geq H_1^{\text{up}} \lambda^b + \frac{F_1 \lambda^c}{\alpha_1 - \bar{\theta}^r} + \frac{F_1}{(\alpha_1 - \bar{\theta}^r)^2} (\theta^c - \bar{\theta}^r \lambda^c) \quad (\text{B.15a})$$

$$\begin{aligned} &= \frac{(\alpha_1 - 2\bar{\theta}^r) f_1 + H_1^{\text{up}} \bar{\theta}^r (\alpha_1 - \bar{\theta}^r) \lambda^b}{(\alpha_1 - \bar{\theta}^r)^2} \\ &\quad + \frac{H_1^{\text{up}} (\alpha_1 - 2\bar{\theta}^r) (\theta^2 - \bar{\theta}^r \lambda^b) + F_1 \theta^c}{(\alpha_1 - \bar{\theta}^r)^2} \end{aligned} \quad (\text{B.15e})$$

$$\geq \frac{(\alpha_1 - 2\bar{\theta}^r) f_1 + \underline{f} \theta_1}{(\alpha_1 - \bar{\theta}^r)^2} \quad (\text{B.15d})$$

$$\begin{aligned} &= \frac{f_1}{(\alpha_1 - \bar{\theta}^r)} + \frac{1}{(\alpha_1 - \bar{\theta}^r)^2} (\underline{f} \theta_1 - \bar{\theta}^r f_1) \\ &= f_1 T_1(\bar{\theta}^r) + T_1'(\bar{\theta}^r) (\underline{f} \theta_1 - \bar{\theta}^r f_1), \end{aligned}$$

$$\begin{aligned} H_1 &\leq \frac{H_1^{\text{up}} (\theta^a - \theta^{\text{lo}} \lambda^a)}{\alpha_1 - \theta^{\text{lo}}} \\ &\quad + \frac{H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b + \theta^b - \theta^{\text{lo}} \lambda^b)}{\alpha_1 - \theta^{\text{lo}}} \\ &\quad + \frac{F_1 \lambda^c + H_1^{\text{up}} (\theta^c - \theta^{\text{lo}} \lambda^c)}{\alpha_1 - \theta^{\text{lo}}} \end{aligned} \quad (\text{B.15b})$$

$$= \frac{f_1}{\alpha_1 - \theta^{\text{lo}}} + H_1^{\text{up}} \left(\frac{\theta - \theta^{\text{lo}}}{\alpha_1 - \theta^{\text{lo}}} \right) \quad (\text{B.15e}), (\text{B.15f}), (\text{B.15j}),$$

$$\underline{f} \theta_1 \geq H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c \quad (\text{B.15c})$$

$$= \theta^{\text{lo}} H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) + H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} - \theta^{\text{lo}} \right) (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c$$

$$= \theta^{\text{lo}} f_1 + H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} - \theta^{\text{lo}} \right) (\alpha_1 \lambda^b - \theta^b)$$

$$+ F_1 (\theta^c - \theta^{\text{lo}} \lambda^c) \quad (\text{B.15e}), (\text{B.11})$$

$$\geq \theta^{\text{lo}} f_1 \quad (\text{B.15h}), (\text{B.15i}),$$

$$\underline{f\theta}_1 \geq H_1^{\text{up}} \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c \quad (\text{B.15c})$$

$$\begin{aligned} &= \alpha_1 H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) - F_1 (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c \\ &= \alpha_1 (f_1 - F_1 \lambda^c) - F_1 (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c \end{aligned} \quad (\text{B.15e})$$

$$= \alpha_1 f_1 - \alpha_1 F_1 (1 - \lambda^a) + F_1 (\theta - \theta^a) \quad (\text{B.15f}), (\text{B.15j})$$

$$\geq \alpha_1 f_1 + F_1 \theta - F_1 \alpha_1 \quad (\text{B.15g}),$$

$$\begin{aligned} \underline{f\theta}_1 &\leq \left\{ H_1^{\text{up}} \bar{\theta}^r (\alpha_1 - \bar{\theta}^r) \lambda^b \right. \\ &\quad \left. + H_1^{\text{up}} (\alpha_1 - 2\bar{\theta}^r) (\theta^b - \bar{\theta}^r \lambda^b) \right\}_{\bar{\theta}^r = \alpha_1} + F_1 \theta^c \end{aligned} \quad (\text{B.15d})$$

$$\begin{aligned} &= \alpha_1 H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) + F_1 \theta^c \\ &= \alpha_1 f_1 - \alpha_1 F_1 \lambda^c + F_1 \theta^c \end{aligned} \quad (\text{B.15e})$$

$$\leq \alpha_1 f_1 \quad (\text{B.15i}),$$

$$\begin{aligned} \underline{f\theta}_1 &\leq \left\{ H_1^{\text{up}} \bar{\theta}^r (\alpha_1 - \bar{\theta}^r) \lambda^b \right. \\ &\quad \left. + H_1^{\text{up}} (\alpha_1 - 2\bar{\theta}^r) (\theta^b - \bar{\theta}^r \lambda^b) \right\}_{\bar{\theta}^r = \alpha_1 - \frac{F_1}{H_1^{\text{up}}}} + F_1 \theta^c \quad (\text{B.15d}) \\ &= F_1 \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \lambda^b + \left(\alpha_1 - 2\frac{F_1}{H_1^{\text{up}}} \right) [H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) - F_1 \lambda^b] + F_1 \theta^c \\ &= \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) [H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) - F_1 \lambda^b] + F_1 \theta^b + F_1 \theta^c \\ &\leq \theta^{\text{lo}} [H_1^{\text{up}} (\alpha_1 \lambda^b - \theta^b) - F_1 \lambda^b] + F_1 \theta^b + F_1 \theta^c \quad (\text{B.15h}), (\text{B.11}) \\ &\leq \theta^{\text{lo}} [f_1 - F_1 \lambda^b - F_1 \lambda^c] + F_1 \theta^b + F_1 \theta^c \quad (\text{B.15e}) \\ &\leq \theta^{\text{lo}} [f_1 - F_1 \lambda^a - F_1 \lambda^b - F_1 \lambda^c] \\ &\quad + F_1 \theta^a + F_1 \theta^b + F_1 \theta^c \quad (\text{B.15g}) \\ &\leq F_1 \theta + \theta^{\text{lo}} f_1 - F_1 \theta^{\text{lo}} \quad (\text{B.15f}), (\text{B.15j}), \end{aligned}$$

$$\theta = \theta^a + \theta^b + \theta^c \quad (\text{B.15f})$$

$$\leq \alpha_1 \lambda^a + \theta^b + \left(\alpha_1 - \frac{F_1}{H_1^{\text{up}}} \right) \lambda^c \quad (\text{B.15g}), (\text{B.15i})$$

$$\leq \alpha_1 - \frac{f_1}{H_1^{\text{up}}} \quad (\text{B.15e}), (\text{B.15j}).$$

B.8 Derivation of MIP Representation of Piecewise Relaxation of $\mathcal{F}_1$

Let the domain of Underwood root be partitioned as $\mathcal{I} = \{[\Theta^0, \Theta^1], \dots, [\Theta^{|\mathcal{I}|-1}, \Theta^{|\mathcal{I}|}]\}$, such that $\alpha_2 = \Theta^0 \leq \dots \leq \Theta^{|\mathcal{I}|} = \alpha_1$. We express the piecewise relaxation of $\mathcal{F}_1$, given by $\bigcup_{t=1}^{|\mathcal{I}|-1} \text{Conv}_{OA}(\mathcal{F}_{1,t}) \cup \mathcal{F}_{1,|\mathcal{I}|,\text{Relax}}$, as the following disjunction:

$$\bigvee_{t=1}^{|\mathcal{I}|-1} \left[\begin{array}{l} H_1^{\text{rs}} \geq f_1^{\text{rs}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1})(\underline{f}\theta_1^{\text{rs}} - \Theta^{t-1} f_1^{\text{rs}}) \\ H_1^{\text{rs}} \geq f_1^{\text{rs}} T_1(\Theta^t) + T_1'(\Theta^t)(\underline{f}\theta_1^{\text{rs}} - \Theta^t f_1^{\text{rs}}) \\ H_1^{\text{ss}} \geq f_1^{\text{ss}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1})(\underline{f}\theta_1^{\text{ss}} - \Theta^{t-1} f_1^{\text{ss}}) \\ H_1^{\text{ss}} \geq f_1^{\text{ss}} T_1(\Theta^t) + T_1'(\Theta^t)(\underline{f}\theta_1^{\text{ss}} - \Theta^t f_1^{\text{ss}}) \\ H_1^{\text{rs}} \leq f_1^{\text{rs}} T_1(\Theta^{t-1}) + \left[\frac{T_1(\Theta^{t-1}) - T_1(\Theta^t)}{\Theta^{t-1} - \Theta^t} \right] (\underline{f}\theta_1^{\text{rs}} - \Theta^{t-1} f_1^{\text{rs}}) \\ H_1^{\text{ss}} \leq f_1^{\text{ss}} T_1(\Theta^{t-1}) + \left[\frac{T_1(\Theta^{t-1}) - T_1(\Theta^t)}{\Theta^{t-1} - \Theta^t} \right] (\underline{f}\theta_1^{\text{ss}} - \Theta^{t-1} f_1^{\text{ss}}) \\ (F_1 - f_1^{\text{in}})\Theta^{t-1} \leq (F_1\theta - \underline{f}\theta_1^{\text{in}}) \leq (F_1 - f_1^{\text{in}})\Theta^t \\ f_1^{\text{rs}}\Theta^{t-1} \leq \underline{f}\theta_1^{\text{rs}} \leq f_1^{\text{rs}}\Theta^t, \quad f_1^{\text{ss}}\Theta^{t-1} \leq \underline{f}\theta_1^{\text{ss}} \leq f_1^{\text{ss}}\Theta^t \\ H_1^{\text{in}} = H_1^{\text{rs}} + H_1^{\text{ss}}, \quad \underline{f}\theta_1^{\text{in}} = \underline{f}\theta_1^{\text{rs}} + \underline{f}\theta_1^{\text{ss}}, \quad f_1^{\text{in}} = f_1^{\text{rs}} + f_1^{\text{ss}} \end{array} \right]$$

$$\bigvee_{t=|\mathcal{I}|} \left[\begin{array}{l} H_1^{\text{rs}} \geq f_1^{\text{rs}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_1^{\text{rs}} - \Theta^{t-1} f_1^{\text{rs}}) \\ H_1^{\text{ss}} \geq f_1^{\text{ss}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_1^{\text{ss}} - \Theta^{t-1} f_1^{\text{ss}}) \\ H_1^{\text{in}} \geq f_1^{\text{in}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_1^{\text{in}} - \Theta^{t-1} f_1^{\text{in}}) \\ H_1^{\text{rs}} \leq \frac{f_1^{\text{rs}}}{\alpha_1 - \Theta^{t-1}} + (H_1^{\text{rs}})^{\text{up}} \left[\frac{\theta - \Theta^{t-1}}{\alpha_1 - \Theta^{t-1}} \right] \\ H_1^{\text{ss}} \leq \frac{f_1^{\text{ss}}}{\alpha_1 - \Theta^{t-1}} + (H_1^{\text{ss}})^{\text{up}} \left[\frac{\theta - \Theta^{t-1}}{\alpha_1 - \Theta^{t-1}} \right] \\ H_1^{\text{in}} \leq \frac{f_1^{\text{in}}}{\alpha_1 - \Theta^{t-1}} + (H_1^{\text{in}})^{\text{up}} \left[\frac{\theta - \Theta^{t-1}}{\alpha_1 - \Theta^{t-1}} \right] \\ (F_1 - f_1^{\text{in}})\Theta^{t-1} \leq (F_1\theta - \underline{f}\theta_1^{\text{in}}) \leq (F_1 - f_1^{\text{in}})\Theta^t \\ f_1^{\text{rs}}\Theta^{t-1} \leq \underline{f}\theta_1^{\text{rs}} \leq f_1^{\text{rs}}\Theta^t, \quad f_1^{\text{ss}}\Theta^{t-1} \leq \underline{f}\theta_1^{\text{ss}} \leq f_1^{\text{ss}}\Theta^t \\ H_1^{\text{in}} = H_1^{\text{rs}} + H_1^{\text{ss}}, \quad \underline{f}\theta_1^{\text{in}} = \underline{f}\theta_1^{\text{rs}} + \underline{f}\theta_1^{\text{ss}}, \quad f_1^{\text{in}} = f_1^{\text{rs}} + f_1^{\text{ss}} \end{array} \right]. \quad (\text{B.16})$$

In $\text{Conv}_{OA}(\mathcal{F}_{1,t})$, we choose the extreme points of the partition, $\bar{\theta} = \Theta^{t-1}$ and $\bar{\theta} = \Theta^t$, for linearization; and in $\mathcal{F}_{1,|\mathcal{I}|,\text{Relax}}$, we choose only $\bar{\theta} = \Theta^{t-1}$ since $T_1(\cdot)$ is not defined at $\bar{\theta} = \Theta^{|\mathcal{I}|}$. In order to derive an MIP representation that is reasonable in size, we make the following simplifications to the set $\mathcal{F}_{1,|\mathcal{I}|,\text{Relax}}$. First, observe that the third inequality in $\mathcal{F}_{1,|\mathcal{I}|,\text{Relax}}$ is implied from the first two inequalities and $H_1^{\text{in}} = H_1^{\text{rs}} + H_1^{\text{ss}}$, so we drop it from the set. Next, if $(H_1^{\text{in}})^{\text{up}} > (H_1^{\text{rs}})^{\text{up}} + (H_1^{\text{ss}})^{\text{up}}$, we reduce $(H_1^{\text{in}})^{\text{up}}$ to $(H_1^{\text{rs}})^{\text{up}} + (H_1^{\text{ss}})^{\text{up}}$ because of fourth and fifth inequalities and $H_1^{\text{in}} = H_1^{\text{rs}} + H_1^{\text{ss}}$. Otherwise, we relax the sixth inequality by letting $(H_1^{\text{in}})^{\text{up}} = (H_1^{\text{rs}})^{\text{up}} + (H_1^{\text{ss}})^{\text{up}}$. Then, the sixth inequality is implied from the fourth and fifth inequalities, so we drop it from the set. Next, using disjunctive programming techniques, we obtain

$$H_{1,t}^{\text{rs}} \geq \max \left\{ f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}), \right. \\ \left. f_{1,t}^{\text{rs}} T_1(\Theta^t) + T'_1(\Theta^t)(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^t f_{1,t}^{\text{rs}}) \right\}, \quad \llbracket t \rrbracket_1^{|\mathcal{I}|-1} \quad (\text{B.17a})$$

$$H_{1,t}^{\text{rs}} \geq f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}), \quad t = |\mathcal{I}| \quad (\text{B.17b})$$

$$H_{1,t}^{\text{ss}} \geq \max \left\{ f_{1,t}^{\text{ss}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}), \right. \\ \left. f_{1,t}^{\text{ss}} T_1(\Theta^t) + T'_1(\Theta^t)(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^t f_{1,t}^{\text{ss}}) \right\}, \quad \llbracket t \rrbracket_1^{|\mathcal{I}|-1} \quad (\text{B.17c})$$

$$H_{1,t}^{\text{ss}} \geq f_{1,t}^{\text{ss}} T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}), \quad t = |\mathcal{I}| \quad (\text{B.17d})$$

$$H_{1,t}^{\text{rs}} \leq f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + \left[\frac{T_1(\Theta^t) - T_1(\Theta^{t-1})}{\Theta^t - \Theta^{t-1}} \right] (\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}), \quad \llbracket t \rrbracket_1^{|\mathcal{I}|-1} \quad (\text{B.17e})$$

$$H_{1,t}^{\text{rs}} \leq \frac{f_{1,t}^{\text{rs}}}{\alpha_1 - \Theta^{t-1}} + (H_1^{\text{rs}})^{\text{up}} \left[\frac{\theta_t - \Theta^{t-1} \mu_t}{\Theta^t - \Theta^{t-1}} \right], \quad t = |\mathcal{I}| \quad (\text{B.17f})$$

$$H_{1,t}^{\text{ss}} \leq f_{1,t}^{\text{ss}} T_1(\Theta^{t-1}) + \left[\frac{T_1(\Theta^t) - T_1(\Theta^{t-1})}{\Theta^t - \Theta^{t-1}} \right] (\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}), \quad \llbracket t \rrbracket_1^{|\mathcal{I}|-1} \quad (\text{B.17g})$$

$$H_{1,t}^{\text{ss}} \leq \frac{f_{1,t}^{\text{ss}}}{\alpha_1 - \Theta^{t-1}} + (H_1^{\text{ss}})^{\text{up}} \left[\frac{\theta_t - \Theta^{t-1} \mu_t}{\Theta^t - \Theta^{t-1}} \right], \quad t = |\mathcal{I}| \quad (\text{B.17h})$$

$$(F_1 \mu_t - f_{1,t}^{\text{in}}) \Theta^{t-1} \leq (F_1 \theta_t - \underline{f}\theta_{1,t}^{\text{in}}) \leq (F_1 \mu_t - f_{1,t}^{\text{in}}) \Theta^t, \quad \llbracket t \rrbracket_1^{|\mathcal{I}|} \quad (\text{B.17i})$$

$$f_{1,t}^{\text{rs}} \Theta^{t-1} \leq \underline{f}\theta_{1,t}^{\text{rs}} \leq f_{1,t}^{\text{rs}} \Theta^t, \quad f_{1,t}^{\text{ss}} \Theta^{t-1} \leq \underline{f}\theta_{1,t}^{\text{ss}} \leq f_{1,t}^{\text{ss}} \Theta^t, \quad \llbracket t \rrbracket_1^{|\mathcal{I}|} \quad (\text{B.17j})$$

$$H_{1,t}^{\text{in}} = H_{1,t}^{\text{rs}} + H_{1,t}^{\text{ss}}, \quad \underline{f}\theta_{1,t}^{\text{in}} = \underline{f}\theta_{1,t}^{\text{rs}} + \underline{f}\theta_{1,t}^{\text{ss}}, \quad f_{1,t}^{\text{in}} = f_{1,t}^{\text{rs}} + f_{1,t}^{\text{ss}}, \quad \llbracket t \rrbracket_1^{|\mathcal{I}|} \quad (\text{B.17k})$$

$$H_1^{\text{in}} = \sum_{t=1}^{|\mathcal{I}|} H_{1,t}^{\text{in}}, \quad H_1^{\text{rs}} = \sum_{t=1}^{|\mathcal{I}|} H_{1,t}^{\text{rs}}, \quad H_1^{\text{ss}} = \sum_{t=1}^{|\mathcal{I}|} H_{1,t}^{\text{ss}}, \quad (\text{B.17l})$$

$$\underline{f}\theta_1^{\text{in}} = \sum_{t=1}^{|\mathcal{I}|} \underline{f}\theta_{1,t}^{\text{in}}, \quad \underline{f}\theta_1^{\text{rs}} = \sum_{t=1}^{|\mathcal{I}|} \underline{f}\theta_{1,t}^{\text{rs}}, \quad \underline{f}\theta_1^{\text{ss}} = \sum_{t=1}^{|\mathcal{I}|} \underline{f}\theta_{1,t}^{\text{ss}}, \quad (\text{B.17m})$$

$$f_1^{\text{in}} = \sum_{t=1}^{|\mathcal{I}|} f_{1,t}^{\text{in}}, \quad f_1^{\text{rs}} = \sum_{t=1}^{|\mathcal{I}|} f_{1,t}^{\text{rs}}, \quad f_1^{\text{ss}} = \sum_{t=1}^{|\mathcal{I}|} f_{1,t}^{\text{ss}}, \quad \theta = \sum_{t=1}^{|\mathcal{I}|} \theta_t, \quad (\text{B.17n})$$

$$\sum_{t=1}^{|\mathcal{I}|} \mu_t = 1, \quad \mu_t \geq 0, \quad \llbracket t \rrbracket_1^{|\mathcal{I}|} \quad (\text{B.17o})$$

Here, μ_t are the convex multipliers in disjunctive programming, and variables with subscript t are to be regarded as linearizations of products of the corresponding variables with μ_t . For example, $\theta_{1,t}^{\text{in}}$ linearizes $\theta^{\text{in}} \mu_t$. To control the problem size, we project out $H_{1,t}^{\text{in}}$, $\underline{f}\theta_{1,t}^{\text{in}}$ and $f_{1,t}^{\text{in}}$ variables by substitution. Next, we eliminate $H_{1,t}^{\text{rs}}$ and $H_{1,t}^{\text{ss}}$ variables using Fourier-Motzkin. This leads to

$$H_1^{\text{rs}} \geq \sum_{t=1}^{|\mathcal{I}|-1} \max \left\{ f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1}) (\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}), \right. \\ \left. f_{1,t}^{\text{rs}} T_1(\Theta^t) + T_1'(\Theta^t) (\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^t f_{1,t}^{\text{rs}}) \right\} + f_{1,t}^{\text{rs}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1}) (\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}) \quad (\text{B.18a})$$

$$H_1^{\text{ss}} \geq \sum_{t=1}^{|\mathcal{I}|-1} \max \left\{ f_{1,t}^{\text{ss}} T_1(\Theta^{t-1}) + T_1'(\Theta^{t-1}) (\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}), \right.$$

$$f_{1,t}^{\text{ss}}T_1(\Theta^t) + T'_1(\Theta^t)(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^t f_{1,t}^{\text{ss}}) \Big\} + f_{1,t}^{\text{ss}}T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{ss}} - \Theta^{t-1} f_{1,t}^{\text{ss}}) \quad (\text{B.18b})$$

$$(4.41\text{e}) - (4.41\text{k}) \quad (\text{B.18c})$$

Now, we observe that each linear function in (B.18a) and (B.18b) is nonnegative. For example, consider $f_{1,t}^{\text{rs}}T_1(\Theta^{t-1}) + T'_1(\Theta^{t-1})(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}})$ in (B.18a). Here, $f_{1,t}^{\text{rs}}T_1(\Theta^{t-1}) \geq 0$, $T'_1(\Theta^{t-1}) \geq 0$, and $(\underline{f}\theta_{1,t}^{\text{rs}} - \Theta^{t-1} f_{1,t}^{\text{rs}}) \geq 0$ (see (4.41h)). We use this observation, and relax (B.18a) and (B.18b) to (4.41a)–(4.41d). Finally, we require the solution to lie in a single partition by imposing integrality constraint on μ_t variables.

B.9 Derivation of MIP representation of Piecewise Relaxation of $\mathcal{V}$

For convenience, we replace (4.32a) and (4.32b) in $\text{Conv}(\mathcal{V})$ with $U^{\text{rs}} - U^{\text{ss}} = \Upsilon^{\text{rs}} - \Upsilon^{\text{ss}}$ and $\underline{U}\theta^{\text{rs}} - \underline{U}\theta^{\text{ss}} = \underline{\Upsilon}\theta^{\text{rs}} - \underline{\Upsilon}\theta^{\text{ss}}$. Note that this still captures $\text{Conv}(\mathcal{V})$, since the former can be derived by a linear combination of the latter. Next, we use disjunctive programming to construct the convex hull of piecewise relaxation of $\mathcal{V} = \bigcup_{t=1}^{|\mathcal{I}|} \text{Conv}(\mathcal{V}_t)$.

$$[t]_1^{|\mathcal{I}|}, \quad U_t^{\text{rs}} - U_t^{\text{ss}} = \Upsilon_t^{\text{rs}} - \Upsilon_t^{\text{ss}}, \quad (\text{B.19a})$$

$$[t]_1^{|\mathcal{I}|}, \quad \underline{U}\theta_t^{\text{rs}} - \underline{U}\theta_t^{\text{ss}} = \underline{\Upsilon}\theta_t^{\text{rs}} - \underline{\Upsilon}\theta_t^{\text{ss}}, \quad (\text{B.19b})$$

$$[t]_1^{|\mathcal{I}|}, \quad 0 \leq (\cdot)\theta_t - \Theta^{t-1}(\cdot)_t \leq (\cdot)^{\text{up}}(\theta_t - \Theta^{t-1}\mu_t), \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}, \quad (\text{B.19c})$$

$$[t]_1^{|\mathcal{I}|}, \quad 0 \leq \Theta^t(\cdot)_t - (\cdot)\theta_t \leq (\cdot)^{\text{up}}(\theta^{\text{up}}\mu_t - \theta_t), \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\}, \quad (\text{B.19d})$$

$$(\cdot)\theta = \sum_{t=1}^{|\mathcal{I}|} (\cdot)\theta_t, \quad (\cdot) = \sum_{t=1}^{|\mathcal{I}|} (\cdot)_t, \quad \forall (\cdot) \in \{U^{\text{rs}}, U^{\text{ss}}, \Upsilon^{\text{rs}}, \Upsilon^{\text{ss}}\} \quad (\text{B.19e})$$

$$\sum_{t=1}^{|\mathcal{I}|} \theta_t = \theta, \quad \sum_{t=1}^{|\mathcal{I}|} \mu_t = 1, \quad \mu_t \geq 0, \quad [t]_1^{|\mathcal{I}|} \quad (\text{B.19f})$$

Here, μ_t are disjunctive programming variables, and variables $\underline{U}\theta_t^{\text{rs}}$, U_t^{rs} are to be regarded as the linearizations of $\underline{U}\theta^{\text{rs}} \cdot \mu_t$, $U^{\text{rs}} \cdot \mu_t$, respectively. To the above, we append the redundant constraint $\underline{U}\theta^{\text{rs}} - \underline{U}\theta^{\text{ss}} = \underline{\Upsilon}\theta^{\text{rs}} - \underline{\Upsilon}\theta^{\text{ss}}$, which is derived by

adding all the equations in (B.19a), and using (B.19e). Then, we relax (B.19) by discarding all the equations in (B.19b). Next, we eliminate variables of the form $\underline{U}\theta_t$ and θ_t in the following manner. For notational convenience, we present the elimination process assuming we have three partitions. Consider

$$0 \leq \underline{U}\theta_t - \Theta^{t-1}U_t \leq U^{\text{up}}(\theta_t - \Theta^{t-1}\mu_t), \quad t = 1, 2, 3 \quad (\text{B.20a})$$

$$0 \leq \Theta^t U_t - \underline{U}\theta_t \leq U^{\text{up}}(\Theta^t \mu_t - \theta_t), \quad t = 1, 2, 3 \quad (\text{B.20b})$$

$$\underline{U}\theta = \underline{U}\theta_1 + \underline{U}\theta_2 + \underline{U}\theta_3, \quad \theta = \theta_1 + \theta_2 + \theta_3 \quad (\text{B.20c})$$

First, we substitute out $\underline{U}\theta_1$ by $\underline{U}\theta - \underline{U}\theta_2 - \underline{U}\theta_3$. Then, we rearrange the inequalities governing $\underline{U}\theta_2$ in the following manner:

$$\left. \begin{aligned} -(\Theta^0 U_1 - \underline{U}\theta + \underline{U}\theta_3) - U^{\text{up}}(\theta_1 - \Theta^0 \mu_1) &\leq \underline{U}\theta_2 \leq -(\Theta^0 U_1 - \underline{U}\theta + \underline{U}\theta_3) \\ -(\Theta^1 U_1 - \underline{U}\theta + \underline{U}\theta_3) &\leq \underline{U}\theta_2 \leq U^{\text{up}}(\Theta^1 \mu_1 - \theta_1) - (\Theta^1 U_1 - \underline{U}\theta + \underline{U}\theta_3) \\ \Theta^1 U_2 &\leq \underline{U}\theta_2 \leq U^{\text{up}}(\theta_2 - \Theta^1 \mu_2) + \Theta^1 U_2 \\ \Theta^2 U_2 - U^{\text{up}}(\Theta^2 \mu_2 - \theta_2) &\leq \underline{U}\theta_2 \leq \Theta^2 U_2 \end{aligned} \right\} \quad (\text{B.21})$$

Now, we eliminate $\underline{U}\theta_2$ using Fourier-Motzkin. We write (L1R3) to denote first inequality from the left hand side, and third inequality from the right hand side.

$$(\text{L1R1}) \text{ and } (\text{L2R2}): \quad \Theta^0 \mu_1 \leq \theta_1 \leq \Theta^1 \mu_1$$

$$(\text{L2R1}) \text{ and } (\text{L1R2}): \quad 0 \leq U_1 \leq \mu_1 U^{\text{up}}$$

$$(\text{L3R3}) \text{ and } (\text{L4R4}): \quad \Theta^1 \mu_2 \leq \theta_2 \leq \Theta^2 \mu_2$$

$$(\text{L3R4}) \text{ and } (\text{L4R3}): \quad 0 \leq U_2 \leq \mu_2 U^{\text{up}}$$

$$\begin{aligned} (\text{L1R3}) \text{ and } (\text{L3R1}): \quad &-(\Theta^0 U_1 + \Theta^1 U_2 - \underline{U}\theta) - U^{\text{up}}(\theta_1 + \theta_2 - \Theta^0 \mu_1 - \Theta^1 \mu_2) \\ &\leq \underline{U}\theta_3 \leq -(\Theta^0 U_1 + \Theta^1 U_2 - \underline{U}\theta) \end{aligned}$$

$$\begin{aligned} (\text{L2R4}) \text{ and } (\text{L4R2}): \quad &-(\Theta^1 U_1 + \Theta^2 U_2 - \underline{U}\theta) \\ &\leq \underline{U}\theta_3 \leq U^{\text{up}}(\Theta^1 \mu_1 + \Theta^2 \mu_2 - \theta_1 - \theta_2) - (\Theta^1 U_1 - \Theta^2 U_2 - \underline{U}\theta) \end{aligned}$$

$$(\text{L1R4}) \text{ and } (\text{L4R1}): \quad -(\Theta^0 U_1 - \underline{U}\theta) - \Theta^2 U_2 - U^{\text{up}}(\theta_1 - \Theta^0 \mu_1)$$

$$\begin{aligned}
&\leq \underline{U}\theta_3 \leq -(\Theta^0 U_1 - \underline{U}\theta) - \Theta^2 U_2 + U^{\text{up}}(\Theta^2 \mu_2 - \theta_2) \\
\text{(L2R3) and (L3R2):} \quad &-(\Theta^1 U_1 - \underline{U}\theta) - \Theta^1 U_2 - U^{\text{up}}(\theta_2 - \Theta^1 \mu_2) \\
&\leq \underline{U}\theta_3 \leq U^{\text{up}}(\Theta^1 \mu_1 - \theta_1) - \Theta^1 U_2 - (\Theta^1 U_1 - \underline{U}\theta)
\end{aligned}$$

We relax the set by discarding inequalities obtained from (L1R4), (L4R1), (L2R3) and (L3R2). The inequalities obtained from (L1R3), (L3R1), (L2R4) and (L4R2) have the same form as the four inequalities in (B.21). As before, we eliminate $\underline{U}\theta_3$ using Fourier-Motzkin, and discard inequalities obtained from (L1R4), (L4R1), (L2R3) and (L3R2). This leads to

$$0 \leq \underline{U}\theta - \sum_{t=1}^3 \Theta^{t-1} U_t \leq U^{\text{up}} \left(\theta - \sum_{t=1}^3 \Theta^{t-1} \mu_t \right) \quad (\text{B.22a})$$

$$0 \leq \sum_{t=1}^3 \Theta^t U_t - \underline{U}\theta \leq U^{\text{up}} \left(\sum_{t=1}^3 \Theta^t \mu_t - \theta \right) \quad (\text{B.22b})$$

$$\theta = \sum_{t=1}^3 \theta_t, \quad \Theta^{t-1} \mu_t \leq \theta_t \leq \Theta^t \mu_t, \quad 0 \leq U_t \leq U^{\text{up}} \mu_t, \quad t = 1, 2, 3 \quad (\text{B.22c})$$

In this manner, we eliminate all variables of the form $\underline{(\cdot)}\theta_t$ from (B.19). Then, we eliminate all θ_t variables, which are now constrained only by (B.22c), using Fourier-Motzkin. This leads to $\sum_{t=1}^{|\mathcal{I}|} \Theta^{t-1} \mu^t \leq \theta \leq \sum_{t=1}^{|\mathcal{I}|} \Theta^t \mu^t$. Since it is implied from (B.22a) and (B.22b), we do not impose it explicitly. Finally, we require the solution to lie in a single partition by imposing integrality constraint on μ_t variables.

C. MISSING IDENTITIES AND DERIVATIONS

C.1 Derivation of the Search Space Formulation

Constraints (A2) through (A8) in Chapter 4 are described below. Here, we write $\{p\}_{p=1}^n$ and $\llbracket p \rrbracket_1^n$ as a shorthand notation for $\{1, \dots, n\}$ and $\forall p \in \{1, \dots, n\}$, respectively.

$$\begin{aligned} \text{for } [i, j] \in \mathcal{P}, \quad & \begin{cases} \tau_{i,k,j} = \nu_{i,k+1,j-1} - \nu_{i,k,j-1} - \nu_{i,k+1,j} + \nu_{i,k,j}, & \llbracket k \rrbracket_i^{j-1} \\ \beta_{i,l,j} = \omega_{i+1,l-1,j} - \omega_{i+1,l,j} - \omega_{i,l-1,j} + \omega_{i,l,j}, & \llbracket l \rrbracket_{i+1}^j \end{cases} \quad (\text{A2}) \\ \text{for } [i, j] \in \mathcal{P}, \quad & \begin{cases} \nu_{i,j,j} = \omega_{i,i,j} = 1 - \zeta_{i,j} \\ \nu_{i,k,j} \geq \max\{0, \nu_{i,k,m} + \nu_{i,n,j} - \nu_{i,n,m}\}, & \llbracket n \rrbracket_{k+1}^{m+1}, \llbracket m \rrbracket_k^{j-1}, \llbracket k \rrbracket_i^{j-1} \\ \nu_{i,k,j} \leq \min\{\nu_{i,k,m}, \nu_{i,n,j}\}, & \llbracket n \rrbracket_{k+1}^{m+1}, \llbracket m \rrbracket_k^{j-1}, \llbracket k \rrbracket_i^{j-1} \\ \omega_{i,l,j} \geq \max\{0, \omega_{i,m,j} + \omega_{n,l,j} - \omega_{n,m,j}\}, & \llbracket n \rrbracket_{i+1}^{m+1}, \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^j \\ \omega_{i,l,j} \leq \min\{\omega_{i,m,j}, \omega_{n,l,j}\}, & \llbracket n \rrbracket_{i+1}^{m+1}, \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^j \end{cases} \quad (\text{A3}) \end{aligned}$$

$$\zeta_{1,N} = \zeta_{1,1} = \dots = \zeta_{N,N} = 1 \quad (\text{A4})$$

$$\text{for } [i, j] \in \mathcal{P}, \quad \begin{cases} \sum_{l=i+1}^{k+1} \sigma_{i,k,l,j} = \tau_{i,k,j}, & \llbracket k \rrbracket_i^{j-1} \\ \sum_{k=l-1}^{j-1} \sigma_{i,k,l,j} = \beta_{i,k,l,j}, & \llbracket l \rrbracket_{i+1}^j \\ \sigma_{i,k,l,j} \geq 0, & \llbracket l \rrbracket_{i+1}^{k+1}, \llbracket k \rrbracket_i^{j-1} \end{cases} \quad (\text{A5})$$

$$\begin{cases} \chi_{i,j} \leq \zeta_{i,j} - \sum_{k=1}^{i-1} \beta_{k,i,j}, & \forall (i, j) \in \mathcal{C} \\ \rho_{i,j} \leq \zeta_{i,j} - \sum_{l=j+1}^n \tau_{i,j,l}, & \forall (i, j) \in \mathcal{R} \end{cases} \quad (\text{A6})$$

$$\begin{cases} \chi_{i,i} \geq \zeta_{i,i} - \sum_{k=1}^{i-1} \beta_{k,i,i}, & \forall (i,i) \in \mathcal{C} \\ \rho_{i,i} \geq \zeta_{i,i} - \sum_{l=i+1}^n \tau_{i,i,l}, & \forall (i,i) \in \mathcal{R} \end{cases} \quad (\text{A7})$$

$$\text{for } [i,j] \in \mathcal{T} \setminus [1,n], \quad \zeta_{i,j} \leq \sum_{k=1}^{i-1} \beta_{k,i,j} + \sum_{l=j+1}^n \tau_{i,j,l} \quad (\text{A8})$$

In the above, $\nu_{i,k,j}$ and $\omega_{i,l,j}$ are linearizations of $\prod_{m=k}^j (1 - \zeta_{i,m})$ and $\prod_{m=i}^l (1 - \zeta_{m,j})$, respectively. $\tau_{i,k,j} = 1$ (*resp.* $\beta_{i,l,j} = 1$) if and only if stream $[i,k]$ (*resp.* $[l,j]$) is produced as distillate (*resp.* residue) from stream $[i,j]$. Note that in (A6) and (A7), we used $\beta_{0,i,j} = \zeta_{i,j} - \sum_{k=1}^{i-1} \beta_{k,i,j}$ and $\tau_{i,j,n+1} = \zeta_{i,j} - \sum_{l=j+1}^n \tau_{i,j,l}$. For more details, see [25]. Using (C.1) and (C.2) (see (3) and (4) in [25] for derivation), we substitute out $\nu_{i,k,j}$ and $\omega_{i,l,j}$ variables.

$$\text{for } [i,j] \in \mathcal{T}, \quad \nu_{i,k,j} = \begin{cases} 0, & \text{for } k = i, \\ 1 - \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s}, & \text{for } i+1 \leq k \leq j \end{cases} \quad (\text{C.1})$$

$$\text{for } [i,j] \in \mathcal{T}, \quad \omega_{i,l,j} = \begin{cases} 1 - \sum_{r=i}^l \sum_{s=l+1}^j \beta_{r,s,j}, & \text{for } i \leq l \leq j-1 \\ 0, & \text{for } l = j \end{cases} \quad (\text{C.2})$$

We substitute (C.1) and (C.2) in (A2) to obtain

$$\text{for } [i,j] \in \mathcal{P}, \quad \begin{cases} \tau_{i,k,j} = \tau_{i,k,j} \\ \omega_{i,l,j} = \omega_{i,l,j} \end{cases} \quad (\text{C.3})$$

Since the above constraints satisfy trivially, we do not include them in the model.

Next, we substitute (C.1) and (C.2) in (A3) to obtain

$$\text{for } [i, j] \in \mathcal{P}, \quad \left\{ \begin{array}{l} \sum_{r=i}^{j-1} \tau_{i,r,j} = \sum_{s=i+1}^j \beta_{i,s,j} = \zeta_{i,j} \\ \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \leq \min \left\{ 1, \sum_{s=k}^m \sum_{r=i}^{k-1} \tau_{i,r,s} + \sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s} - \sum_{s=n}^m \sum_{r=i}^{n-1} \tau_{i,r,s} \right\}, \\ \quad \quad \quad \llbracket n \rrbracket_{k+1}^{m+1}, \llbracket m \rrbracket_k^{j-1}, \llbracket k \rrbracket_{i+1}^{j-1} \\ \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \geq \max \left\{ \sum_{s=k}^m \sum_{r=i}^{k-1} \tau_{i,r,s}, \sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s} \right\}, \quad \llbracket n \rrbracket_{k+1}^{m+1}, \llbracket m \rrbracket_k^{j-1}, \llbracket k \rrbracket_{i+1}^{j-1} \\ \sum_{r=i}^l \sum_{s=l+1}^j \beta_{r,s,j} \leq \min \left\{ 1, \sum_{r=i}^m \sum_{s=m+1}^j \beta_{r,s,j} + \sum_{r=n}^l \sum_{s=l+1}^j \beta_{r,s,j} - \sum_{r=n}^m \sum_{s=m+1}^j \beta_{r,s,j} \right\}, \\ \quad \quad \quad \llbracket n \rrbracket_{i+1}^{m+1}, \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^{j-1} \\ \sum_{r=i}^l \sum_{s=l+1}^j \beta_{r,s,j} \geq \max \left\{ \sum_{r=i}^m \sum_{s=m+1}^j \beta_{r,s,j}, \sum_{r=n}^l \sum_{s=l+1}^j \beta_{r,s,j} \right\}, \quad \llbracket n \rrbracket_{i+1}^{m+1}, \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^{j-1} \end{array} \right. \quad (\text{C.4})$$

We express the second constraint in (C.4) as the intersection of two inequalities

$$\sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \leq 1 \text{ and } \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \leq \sum_{s=k}^m \sum_{r=i}^{k-1} \tau_{i,r,s} + \sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s} - \sum_{s=n}^m \sum_{r=i}^{n-1} \tau_{i,r,s}.$$

By noting that $k \leq m < j$ and $k < n \leq m+1 \leq j$, the second inequality can be

simplified as $0 \leq \sum_{s=m+1}^j \sum_{r=k}^{n-1} \tau_{i,r,s}$. This inequality is trivially satisfied because of

(A5), so we do not include it in the model. Similarly, we express the third constraint

in (C.4) as the intersection of two inequalities $\sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \geq \sum_{s=k}^m \sum_{r=i}^{k-1} \tau_{i,r,s}$ and

$\sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \geq \sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s}$. By recognizing $k \leq m < j$, the first inequality can be simplified to $0 \leq \sum_{m+1}^j \sum_{r=i}^{k-1} \tau_{i,r,s}$. This inequality is redundant, because of

(A5). We repeat the arguments with the last two constraints in (C.4), and discard the redundant inequalities. This simplifies (C.4) to

$$\text{for } [i, j] \in \mathcal{P}, \quad \begin{cases} \sum_{r=i}^{j-1} \tau_{i,r,j} = \sum_{s=i+1}^j \beta_{i,s,j} = \zeta_{i,j} \\ \sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s} \leq \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \leq 1, \quad \llbracket n \rrbracket_{k+1}^j, \llbracket k \rrbracket_{i+1}^{j-1} \\ \sum_{r=i}^m \sum_{s=m+1}^j \beta_{r,s,j} \leq \sum_{r=i}^l \sum_{s=l+1}^j \beta_{r,s,j} \leq 1, \quad \llbracket m \rrbracket_i^{l-1}, \llbracket l \rrbracket_{i+1}^{j-1} \end{cases} \quad (\text{C.5})$$

The second constraint in (C.5) for streams of the form $[i, N]$ is given by

$$\sum_{s=n}^N \sum_{r=i}^{n-1} \tau_{i,r,s} \leq \sum_{s=k}^N \sum_{r=i}^{k-1} \tau_{i,r,s} \leq 1, \quad \llbracket n \rrbracket_{k+1}^N, \llbracket k \rrbracket_{i+1}^{N-1} \quad (\text{C.6})$$

Consider a stream $[i, j]$ such that $j < N$. Then, we observe that $\sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \leq \sum_{s=k}^N \sum_{r=i}^{k-1} \tau_{i,r,s}$. This implies that the constraint $\sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s} \leq 1$ in (C.5) is redundant for $j < N$, since it can be derived from (C.6). Next, we observe that the first inequality in (C.6) can be expressed as $\sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s} + \sum_{s=j+1}^N \sum_{r=k}^{n-1} \tau_{i,r,s} \leq \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s}$. Since $\tau_{i,r,s} \geq 0$ from (A5), the constraint $\sum_{s=n}^j \sum_{r=i}^{n-1} \tau_{i,r,s} \leq \sum_{s=k}^j \sum_{r=i}^{k-1} \tau_{i,r,s}$ in (C.5) is implied from (C.6) for $j < N$. In a similar manner, it can be shown that it suffices to impose the third constraint in (C.5) only for streams of the form $[1, j]$ *i.e.*, $i = 1$. Therefore, the projection of (A2) and (A3) in $\tau - \beta$ space is given by

$$\begin{cases} \sum_{r=i}^{j-1} \tau_{i,r,j} = \sum_{s=i+1}^j \beta_{i,s,j} = \zeta_{i,j}, \quad \forall [i, j] \in \mathcal{P} \\ \sum_{s=n}^N \sum_{r=i}^{n-1} \tau_{i,r,s} \leq \sum_{s=k}^N \sum_{r=i}^{k-1} \tau_{i,r,s} \leq 1, \quad \llbracket n \rrbracket_{k+1}^N, \llbracket k \rrbracket_{i+1}^{N-1} \\ \sum_{r=1}^m \sum_{s=m+1}^j \beta_{r,s,j} \leq \sum_{r=1}^l \sum_{s=l+1}^j \beta_{r,s,j} \leq 1, \quad \llbracket m \rrbracket_1^{l-1}, \llbracket l \rrbracket_2^{j-1} \end{cases} \quad (\text{C.7})$$

Next, we use (A5) to project out $\tau_{i,k,j}$ and $\beta_{i,l,j}$ variables from (C.7), (A6) – (A8) to obtain the constraints (W1) – (W6).

C.2 Gauss Quadrature Formula

$$\int_0^1 f(\phi) d\phi \approx \sum_{g \in \mathcal{G}} \xi_g f(\phi_g) \quad (\text{C.8})$$

where $\mathcal{G} = \{1, 2\}$, $\xi_1 = \xi_2 = 1/2$, $\phi_1 = 0.211325$ and $\phi_2 = 0.788675$ for two point quadrature formula.

C.3 Proof of Identity

$$\left(\sum_k u_k w_k \right)^2 + \sum_k \sum_{l, l < k} (u_k - u_l)^2 w_k w_l = \left(\sum_l w_l \right) \left(\sum_k u_k^2 w_k \right) \quad (\text{C.9})$$

Consider the left hand side

$$\begin{aligned} \text{LHS} &= \left(\sum_k u_k w_k \right)^2 + \sum_k \sum_{l, l < k} (u_k - u_l)^2 w_k w_l \\ &= \sum_k u_k^2 w_k^2 + 2 \sum_k \sum_{l, l < k} u_k u_l w_k w_l + \sum_k \sum_{l, l < k} (u_k^2 + u_l^2 - 2u_k u_l) w_k w_l \\ &= \sum_k u_k^2 w_k^2 + \sum_k \sum_{l, l < k} u_k^2 w_k w_l + \sum_k \sum_{l, l < k} u_l^2 w_k w_l \\ &= \sum_k u_k^2 w_k^2 + \sum_k \sum_{l, l < k} u_k^2 w_k w_l + \sum_l \sum_{k, k > l} u_k^2 w_l w_k \\ &= \sum_k \sum_{l, k=l} u_k^2 w_k w_l + \sum_k \sum_{l, k > l} u_k^2 w_k w_l + \sum_k \sum_{l, k < l} u_k^2 w_k w_l \\ &= \sum_k \sum_l u_k^2 w_k w_l \\ &= \left(\sum_l w_l \right) \left(\sum_k u_k^2 w_k \right) = \text{RHS} \end{aligned}$$

C.4 Relative Volatilities for the Case Study

We remark on the computation of relative volatilities used for the case study in §5.2. Shale gas is treated as a mixture of five components C1 through C5. For

this system, large number of data points are required to span the entire composition space. Therefore, to reduce the computational effort, we chose to fit the surrogate for each consecutive binary pairs of representative species. Figure C.1 shows the true VLE (solid blue curve) and the best surrogate (dashed red curve) in least-square error sense for Methane-Ethane, Ethane-Propane, Propane-Butane, Butane-Pentane pairs. Next, the required relative volatilities are computed as $\alpha_1 = \alpha_{12}\alpha_{23}\alpha_{34}\alpha_{45}$, $\alpha_2 = \alpha_{23}\alpha_{34}\alpha_{45}$, $\alpha_3 = \alpha_{34}\alpha_{45}$, $\alpha_4 = \alpha_{45}$ and $\alpha_5 = 1$ (see Figure title for α_{12} , α_{23} , α_{34} , and α_{45} values).

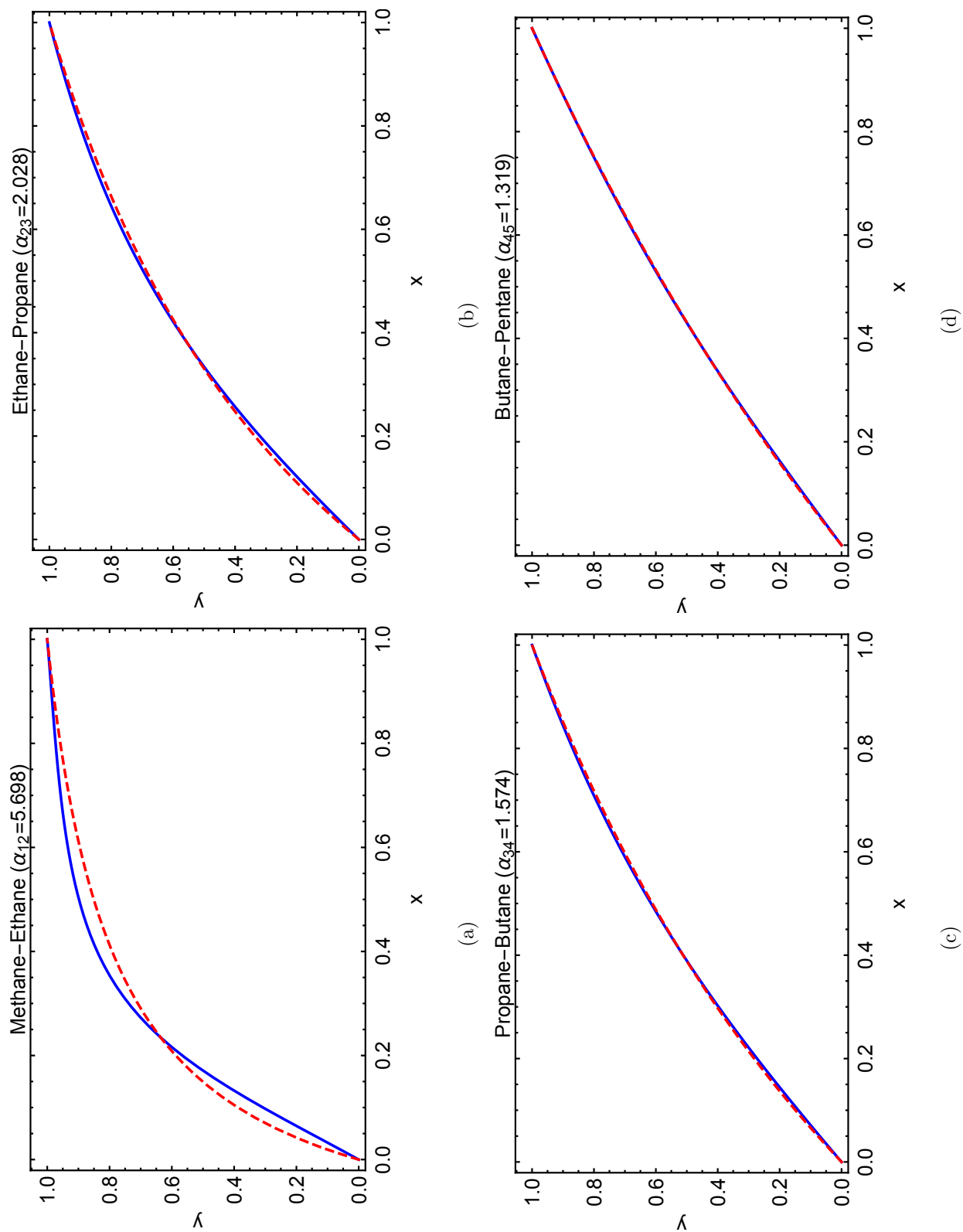


Fig. C.1. True VLE (solid blue curve) and the best surrogate (dashed red curve) in least square sense for (a) Methane-Ethane (b) Ethane-Propane (c) Propane-Butane and (d) Butane-Pentane pairs. For each pair, the surrogate model is taken to be $y = \alpha_{ij}x/(\alpha_{ij}x + 1 - x)$.

VITA

VITA

Radhakrishna Tumbalam Gooty received Bachelor of Technology with Honors and Master of Technology (under the Dual Degree Program) in Chemical Engineering from the Indian Institute of Technology Madras, India, in 2015. Later, in the Fall of 2015, he joined Purdue University to pursue a PhD in chemical engineering under the guidance of Prof. Rakesh Agrawal and Prof. Mohit Tawarmalani. For his thesis, he developed novel mixed integer nonlinear programs and global optimization algorithms to reliably identify energy-efficient distillation configurations. Over the course of the PhD program, he received multiple awards from the department, the graduate school, and prestigious international conferences. He received the Centennial Fellowship from the Davidson school of chemical engineering, the Bilsland Dissertation Fellowship from Purdue University graduate school, PSE Young Researcher award from the International Symposium on Process Systems Engineering (PSE) 2018, the AIChE separations division graduate student research award, and he was one of finalists for the AIChE CAST (Computing and Systems Technology) Director's student presentation award.