Property Based Process and Product Synthesis and Design

Eden, Mario Richard; Gani, Rafiqul; Jørgensen, Sten Bay

Publication date:
2003

Document Version
Early version, also known as pre-print

Property Based Process and Product Synthesis and Design

Ph.D. Thesis

Mario Richard Eden
CAPEC
Department of Chemical Engineering
Technical University of Denmark

November 2003

ISBN 87–90142–97–7
Printed by Book Partner, Nørhaven Digital, Copenhagen, Denmark

Preface

This thesis is submitted as partial fulfillment of the requirements for the Ph.D.–degree at Danmarks Tekniske Universitet (Technical University of Denmark). The work has been carried out at Institut for Kemiteknik (Department of Chemical Engineering) from June 1999 to November 2003 under the supervision of Professor Sten Bay Jørgensen and Professor Rafiqul Gani. Financial support from the Nordic Energy Research Program (Process Integration Section) and the Danish Department of Energy is highly appreciated.

I would like to thank all the students and coworkers at the Computer Aided Process Engineering Center (CAPEC) at the Department of Chemical Engineering at the Technical University of Denmark for their support and many fun hours of research. In particular, Mr. René Skotte, Dr. Peter M. Harper and Dr. Martin Hostrup deserve special recognition for their enduring friendship and fruitful collaboration through the years.

Very special thanks go to Professor Sten Bay Jørgensen and Professor Rafiqul Gani for giving me more opportunities than ordinarily extended to Ph.D. students, not only has this helped my research for this thesis but also taught me how to conduct international research. I have thoroughly enjoyed being a part of the dynamic environment in CAPEC and feel very grateful for the diverse experiences I have been exposed to. Being deeply involved in the organization of ESCAPE-11, an international conference with more than 300 participants, was a most gratifying experience, that I am convinced will be beneficial in my future career. I sincerely appreciate the leadership, vision and guidance of both Professor Sten Bay Jørgensen and Professor Rafiqul Gani throughout this project and will fondly remember the professional as well as personal discussions carried out at the department and on numerous travels abroad.

Also, a most sincere thanks goes to Dr. Mahmoud M. El-Halwagi for inviting me to do research with him at Auburn University. Working closely with such a visionary and innovative researcher was very fruitful and I enjoyed the time in Auburn immensely, academically as well as personally. Not only did Dr. El-Halwagi invite me to join a creative working environment, he also became my friend, for which I am very thankful. I also thank all the faculty, staff and students at the Department of Chemical Engineering at Auburn University for their warm welcome, kindness and support. My deepest appreciation goes to Dr. Christopher B. Roberts, his wife Tracy and their two daughters Heather and Natalie. The same heartfelt appreciation goes to Mr. Scott Besong, his wife Mary-Kathryn and their children Lauren and Jack. Words can not adequately describe the amazing hospitality and unconditional friendships extended to me by these two families. During my two stays in Auburn, all of these people

ensured that I felt at home in the loveliest village on the plains. To all of you..... WAR EAGLE!

Finally, I would like to thank my family and friends for their support and understanding during the writing of this thesis. I am grateful to my parents, Asta and Richard Eden for instilling in me a desire for learning and for always supporting and encouraging my pursuit of academic excellence. My brother Jens and his girlfriend Helene deserve a heartfelt thank you for proofreading this thesis and for always being there to lift my spirits when required. Last, but not least I extend my appreciation to Mr. Carsten Jensen, his fiancé Lone and their two sons Tobias and Magnus for countless dinners and enjoyable times over the years. Friends like these are hard to come by.

This thesis is dedicated to the memory of my father,
who passed away prematurely January 23, 1998.

Tempora mutantur, et nos mutamur in illis.

Lyngby, November 2003

Mario Richard Eden

Abstract

This thesis describes the development of a general framework for solving process and product design problems. Targeting the desired performance of the system in a systematic manner relieves the iterative nature of conventional design techniques. Furthermore, conventional component based methods are not capable of handling problems, where the process or product objectives are driven by functionalities or properties rather than chemical constituency. The framework is meant to complement existing composition based methods by being able to handle property driven problems.

By investigating the different roles a property model plays at different stages of the solution to a design problem, it is discovered that by decoupling the constitutive equations, that make up the property model, from the balance and constraint equations of the process or product model, a significant reduction in problem complexity is achieved including an added flexibility compared to existing solution methods. The decoupling of the constitutive equations allows for reformulating a conventional forward problem into two reverse problems. The first reverse problem is the reverse of a simulation problem, where the process model is solved in terms of the constitutive (synthesis/design) variables instead of the process variables, thus providing the synthesis/design targets. The second reverse problem (reverse property prediction) solves the constitutive equations to identify unit operations, operating conditions and/or products by matching the synthesis/design targets. The reverse problem formulation technique extends the application range of the numerical solvers as well as the models themselves, thus it is possible to identify alternative designs that conventional methods are not capable of finding.

A novel way of representing the constitutive variables is presented in this thesis. The framework is based on tracking functionalities or properties of the process streams rather than the chemical constituency. The motivation for developing this framework comes from a number of cases where conventional composition based methods fail to adequately solve the design problems. The methodology for tracking stream functionalities or properties is referred to as property clustering. The clusters are derived to obey the principles of intra- and interstream conservation, which allow for the development of consistent additive rules along with their ternary representation, thereby facilitating visualization on triangular diagrams. An important feature of the clustering technique is the ability to reduce a high dimensional problem into a two or three dimensional space allowing for visualization of the problem.

The developed framework provides a systematic methodology for solving process and product design problems. The reverse problem formulation techniques

allow for easier identification of optimal solutions and the property clustering techniques enable the systematic solution of problems that are driven by physical properties rather than components. It should be emphasized that the work presented here introduces the general framework of reverse problem formulation, where the links between the two formulated reverse problems are the constitutive variables. One representation of such variables is the property clustering methodology presented here, but the general framework is applicable to any representation of the properties.

Resumé på dansk

Design af kemiske produkter og processer trækker på mange forskellige discipliner indenfor den tekniske kemi samt andre naturvidenskabelige fagområder, såsom matematik og fysik. Mangfoldigheden af aspekter, der skal inddrages i løsningen af ethvert design problem har ført til en generel accept af, at design er en iterativ proces.

Traditionelt indebærer løsningen af et design problem, at der udvikles en model, som så efterfølgende evalueres igen og igen på baggrund af kriterier for succes og fiasko, der er blevet formuleret i forvejen. Et fejlslagent design er ofte ligeså informativt som en succesfuld løsning. Designprocessen gentages således med ændringer og forbedringer indtil designeren er overbevist om at det designede anlæg eller produkt opfører sig som simuleringen påstår. Det er disse overvejelser, der danner baggrunden for det første spørgsmål, som denne afhandling forsøger at besvare:

Er det muligt at formulere en generel metodik, som målrettet kan identificere den optimale løsning til et givent design problem, dvs. gøre op med tanken om iterativt design?

På grund af den iterative opbygning af traditionelle designmetoder, kan identifikationen af den optimale løsning være både besværlig og tidskrævende. Det vil være mere effektivt hvis den optimale løsning kunne identificeres fra starten, således at der findes et sammenligningsgrundlag at holde alternative løsninger op imod. Da det ikke er muligt apriori at specificere den optimale proces- eller produktkonfiguration fra et designperspektiv, vil det være fordelagtigt at specificere den ønskede opførsel og derefter identificere de proces- eller produktbetingelser, der kan realisere dette mål. Ved at anvende en sådan målrettet fremgangsmåde reduceres den iterative natur, som løsningen af designproblemer lider af, idet designmålene identificeres før de detaljerede beregninger udføres. Enhver løsning, der opfylder designmålene er således en brugbar løsning på problemet.

Efterhånden som udviklingen indenfor kemisk proces design fokuserer mere og mere på udviklingen af integrerede strategier, der tillader samtidig løsning af proces- og produktdesign problemer, øges kompleksiteten af det overordnede design problem betragteligt. Matematiske programmeringsmetoder er velkendte men viser sig ofte at være meget komplekse og tidskrævende ved anvendelse på store kemiske, biokemiske og/eller farmaceutiske processer. Modelanalyse kan give den indsigt, der tillader simplificering af det overordnede problem og samtidig udvider simplificeringen anvendelsesområderne af de oprindelige modeller. En matematisk model, som repræsenterer en kemisk proces består i princip-

pet af balanceligninger, begrænsningsligninger og konstitutive ligninger, der beskriver fænomener/egenskaber. Kompleksiteten af en procesmodel skyldes i de fleste tilfælde sammenhængen mellem de konstitutive variable og de intensive variable såsom temperatur, tryk og sammensætning. Den valgte konstitutive (egenskabs) model repræsenterer disse sammenhænge. Hvis de konstitutive ligninger afkobles fra balance- og begrænsningsligningerne kan et konventionelt proces-/produktdesign problem omformuleres til to omvendte problemer. Det første omvendte problem svarer til at lave omvendt eller baglæns simulering, hvor procesmodellen løses med hensyn til de konstitutive (syntese/design) variable istedet for procesvariablene. Herved identificeres syntese-/designmålene. Det andet omvendte problem (baglæns egenskabsforudsigelse) løser de konstitutive ligninger for at identificere enhedsoperationer, operationsbetingelser og/eller produkter ved at matche syntese-/designmålene. Det er vigtigt at understrege, at sålænge løsningen af de konstitutive ligninger matcher målene er det IKKE nødvendigt at løse balance- og begrænsningsligningerne igen. Således opnås ydermere en beregningsmæssigt gevinst.

Forbindelsesleddet mellem de to omvendte problemformuleringer er de konstitutive variable. En repræsentation af sådanne variable er de fysiske egenskaber af komponenter og strømme i systemet. Designmålene, der identificeres ved løsning af det første omvendte problem, beskriver bestemte egenskabsværdier, som skal være opfyldt, når det andet omvendte problem løses. Derfor er der brug for en metodik, der på en systematisk måde kan kortlægge hvorledes strømmenes egenskaber ændrer sig gennem anlægget. En sådan metodik skal endvidere kunne håndtere problemer, som er drevet af strømmenes egenskaber og ikke den kemiske sammensætning. Et eksempel på et sådant problem, er design af papir til et givent formål, hvor kvaliteten af papiret ikke kan beskrives med sammensætning alene. Papir består primært af cellulose og papirets egenskaber såsom reflektivitet og opacitet afhænger ikke af cellulose renheden, men derimod egenskaber som fiberlængde. Dette fører til formuleringen af det andet spørgsmål, som behandles i denne afhandling:

Er det muligt at udvikle en systematisk metode til repræsentation og løsning af problemer, som er drevet af egenskaber og ikke komponenter?

For nyligt er der blevet introduceret en metode til systematisk kortlægning af fysiske egenskaber i kemiske processer. Konceptet er baseret på egenskabsfunktioner (på engelsk: property clusters), der tillader en systematisk repræsentation af strømme og enheder fra et egenskabsperspektiv. I denne afhandling er disse teknikker blevet implementeret i den generelle metodik med omvendte problemformuleringer, hvorved der opnås en repræsentation af de konstitutitve variable. Disse egenskabsfunktioner er udviklet på en sådan måde at de overholder de fundamentale bevarelsesprincipper for intra- og interstrømme. Derved bliver det muligt at udvikle konsistente additive regler samt visualiseringsværktøjer. Kombinationen af omvendte problemformuleringer og egenskabsfunktioner resulterer i en metodik, som anvendes til målrettet løsning af en række egenskabsbetingede problemer.

Der præsenteres to primære fremskridt i denne afhandling, nemlig introduktionen af omvendte problemformuleringer, der reducerer eller fjerner det iterative islæt ved løsning af design problemer, samt muligheden for at løse problemer, der er drevet af fysiske egenskaber i stedet for sammensætning, ved anvendelse af egenskabsfunktioner. En vigtig fordel er evnen til at identificere optimale løsninger på proces- og produktdesignproblemer langt nemmere end ved brug af traditionelle løsningsmetoder. Endvidere giver egenskabsfunktioner mulighed for at løse problemer, som de konventionelle sammensætningsbaserede metoder ikke kan håndtere. Metodikken komplementerer således de eksisterende molekylære komponentbaserede metoder.

Contents

Appendices

List of Figures

List of Tables

1

INTRODUCTION

Designing chemical processes and products draws upon many disciplines in the chemical engineering field along with other natural sciences, such as chemistry, mathematics, and physics. Due to the diversity of disciplines involved and the number of aspects that need to be addressed in the solution of any type of design problem, it has become widely accepted, that design is iterative (Suh, 1990; Eppinger, 1991; Greenbaum and Kyng, 1991; Schuler and Namioka, 1993; Winograd, 1996). Furthermore in many cases the goal of the design changes as the designer begins to understand what can and can not be done.

The conventional means of addressing design problems involves the development of a model, which is then tested and evaluated over and over again for the success or failure of a process or product using key criteria, that have been established in advance. Failure is as informative, often more so, than success. The design cycles through with refinements and improvements are developed until the designers are satisfied that the process and/or product will meet the desired performance criteria. These considerations form the basis for the first question, that this work tries to answer:

> *Is it possible to formulate a general framework capable of targeting the optimum solution to a design problem, i.e. challenge the acceptance of iterative design and thus reduce the number of iterations?*

Due to the iterative nature of traditional design methods, identification of the optimum solution may be cumbersome and time consuming. It would be more efficient if the target could be identified beforehand, thus providing a comparative measure for the identified solutions. Since it is not possible to specify the optimum solution apriori from a design perspective, it would be beneficial to be able to specify the desired process/product performance targets, and then identify the process/product specifics capable of achieving this target. By employing such a targeted approach, the iterative nature of solving design problems could be relieved, provided that the design targets are identified prior to performing any design calculations. Any design that matches the targets would therefore be a viable solution to the problem.

In recent years the chemical engineering design community has moved towards the development of integrated solution strategies for simultaneous consideration of process and product design issues. By doing so, the complexity of

the design problem increases significantly. Mathematical programming methods are well known, but may prove rather complex and time consuming for application to large and complex chemical, biochemical and/or pharmaceutical processes. Model analysis can provide insights that allow for simplification of the overall problem as well as extending the application range of the original models. In principle, the model equations representing a chemical process and/or product consist of balance equations, constraint equations and constitutive equations. The nonlinearity of the model, in many cases, is attributed to the relationships between the constitutive variables and the intensive variables. The model selected for the constitutive equations usually represents these relationships. By decoupling the constitutive equations from the balance and constraint equations a conventional process/product design problem may be reformulated as two reverse problems. The first reverse problem is the reverse of a simulation problem, where the process model is solved in terms of the constitutive (synthesis/design) variables instead of the process variables, thus providing the synthesis/design targets. The second reverse problem (reverse property prediction) solves the constitutive equations to identify unit operations, operating conditions and/or products by matching the synthesis/design targets. Since the reverse problem formulation technique extends the application range of the numerical solvers as well as the models themselves, it is possible to identify alternative designs that conventional methods are not capable of finding. Thus business decision making can be facilitated, as the methodology allows for easy screening of alternatives and identification of candidate designs that should be selected for further, more rigorous investigation.

The links between the two reverse problems are the constitutive variables. One representation of such variables is the physical properties of the components and streams in the system. The design targets, which are identified by the first reverse problem describe specific property values, that need to be matched when solving the second reverse problem. Therefore a framework capable tracking properties in a systematic manner is called for. Such a framework should also be able to handle problems that are driven by properties rather than chemical constituency. An example of such a problem is the design of paper of a specified quality, where the performance of the paper machine, i.e. the quality of the paper, can not be described by composition based methods alone, since paper consists primarily of cellulose. The properties of paper, that determine whether or not the quality is acceptable, e.g. reflectivity, opacity etc., depend on the physical properties of the paper, e.g. fiber length, fibrocity etc. These considerations lead to the formulation of the second question, which is addressed in this thesis:

Is it possible to develop a systematic framework for representation and solution of problems that are driven by properties rather than components?

Recently, the concept of property clustering has been introduced for systematic tracking of properties throughout chemical processes. These clusters allow for

systematic representation of process streams and units from a property perspective. In this thesis, the property clustering techniques are incorporated into the reverse problem formulation framework, thus providing a representation of the constitutive variables. The clusters are tailored to have the attractive features of intra-stream and inter-stream conservation, thus enabling the development of consistent additive rules along with their ternary representation. The combined reverse problem formulation and property clustering framework is used to solve several property driven problems in a targeted manner.

The primary contributions of the work presented in this thesis are two-fold, i.e. first, by the introduction of the reverse problem formulation concepts, the inherently iterative nature of solving design problems is relieved for a large class of problems. Secondly, the use of property based clusters allows for solution of problems that are driven by properties rather than chemical constituency. A key advantage of the reverse problem formulation framework presented in this thesis, is the ability to identify optimum solutions to process and product design problems much easier than by solving the conventional forward problem. Furthermore, the property clustering technique represents an extension to conventional composition based methods, as it enables the solution of problems where conventional methods are not capable of describing the problem adequately.

In chapter 2, a review of the different methodologies and conceptual frameworks, that constitute the basis and the mindset of the reverse problem formulation and property clustering framework, is presented. In chapter 3, the different roles of property models in design are discussed and the motivation for the reverse problem formulation framework presented. The concept of reverse problem formulations is presented in chapter 4, where the reformulation methodology is described along with the general solution strategy. The property clustering techniques, which are used for the representation of the constitutive variables or properties, are presented in chapter 5. Along with the cluster formulation and solution strategies, different visualization tools and design rules based on lever-arm analysis are presented as well. In chapter 6 simple cluster based unit operation models are developed. Application examples of the developed framework are presented in chapters 7 and 8. In chapter 7 different process design problems are solved, while in chapter 8 the focus is on product design, particularly the synthesis and design of different formulations. Two of the problems presented in chapter 7 illustrate the interface between process and product design problems. In the metal degreasing example presented in section 7.4, the process design problem is solved by targeting the optimum process performance, thus yielding the properties of a required external process fluid, that can achieve the target. In the orange juice manufacturing example presented in section 7.7, the design of the finishing process, is determined by targeting the optimum solution to the product design problem. Finally, a summary of the thesis along with concluding remarks and a discussion of the remaining challenges is given in chapter 9. Following the appendices, the nomenclature used in this thesis as well as a list of definitions is presented.

Parts of this thesis have already been published or are currently in press. The original derivation of the property cluster formulation given in section 5.2 was presented by Shelley and El-Halwagi (2000). The further development of the clustering technique along with the identification of the feasibility region boundary in section 5.4 is being published by El-Halwagi *et al.* (2004). Finally, the reverse problem formulation framework presented in chapter 4 and the composition free modeling aspects presented in chapter 6 is being published by Eden *et al.* (2004).

2

THEORETICAL BACKGROUND

2.1 Introduction

The concept of design in chemical engineering can be illustrated through the product tree, which is presented in figure 2.1 below. At the root of the tree are a limited number (approximately 10) of basic raw materials, e.g. crude oil, natural gas, salts, air, water etc. These are then processed to form basic products (approximately 30) such as ethylene, synthesis gas, ammonia, benzene, sulfuric acid and so on. From the basic products a wide range of intermediates (approximately 300) are produced, e.g. methanol, styrene and acetic acid, which in turn form the basis for the production of the specialty chemicals and consumer products (approximately 30,000) at the top of the tree.

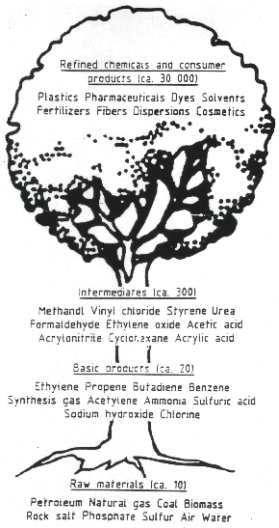

Figure 2.1: The product tree

A chemical design engineer faces many challenges when asked to design a chemical processing facility. The product tree represents a holistic view of the problems, that need to be addressed in an efficient manner:

- The important fruits (products)
- The optimal path (process) to reach them
- The feasibility of the designed process
- A control structure ensuring operational stability
- Resource conservation strategies
- Environmental, health and safety issues

Each of these problems have developed into independent research areas and although some parts of the problems are often solved in an integrated manner, it is not likely that all problems can be solved simultaneously. The focus of the research presented in this thesis has been on the development of new process and product design techniques as well as implementation of resource conservation strategies.

The property based design techniques presented in this thesis constitute a supplement to the existing design methodologies. Although the developed framework is novel, it is based on conventional methods and utilizes strategies and insights from the conventional design disciplines. In this chapter, a review of several process and product synthesis/design techniques and methods is presented in order to introduce the mindset and available methodologies that served as the foundation for the developed framework.

2.2 Process Synthesis and Design

A chemical or pharmaceutical manufacturing process can be divided into separate parts based on the types of operation to be performed in each part, e.g. reaction, separation and utility operations. This division forms the basis for the primary research areas in process synthesis and design:

- Synthesis of reactors and reactor networks
- Synthesis of separation schemes
- Synthesis of Heat Exchanger Networks (HEN)

It should be pointed out that utility operations are not restricted to heat exchange and heat recovery, but also involves, e.g. pumping and compressing. Furthermore, the issue of recycling within the processing plant needs to be addressed systematically, as such efforts may significantly change the design variables and thus also the design itself. Recycling increases the utilization of

process resources, but the increased mass flows in the recycle loops also result in larger processing units and potential nonlinear dynamics, which might need to be stabilized. The division outlined here is similar to the lower 3 levels in the decision hierarchy presented by Douglas (1985), where the synthesis decisions are divided into 5 levels of decreasing influence on the design. The preceding two levels consist of the choice between batch or continuous processing and the input-output structure of the flowsheet:

Level 1 Batch vs. continuous
Level 2 Input-output structure of the flowsheet
Level 3 Reactor and recycle structure of the flowsheet
Level 4 Separation system synthesis
Level 5 Heat recovery network

The first step in the solution of any process synthesis problem is the gathering of information, e.g. the physical and chemical properties of the components and/or mixtures present in the system. Furthermore, the implementation of alternative reaction schemes, solvents, separation techniques and equipment must be considered. Biegler *et al.* (1997) state the principal steps in process synthesis as follows:

1. Gathering information
2. Representation of alternatives
3. Assessment of preliminary designs
4. Generating and searching among alternatives

The task of identifying the optimum solution to a process synthesis/design problem can also be represented in a mathematical form as shown in equation 2.1. This set of equations is the mathematical expression of the following statement (Biegler *et al.*, 1997):

Given a system or process, identify the best possible solution to this process within a specified set of constraints.

When solving such a problem, a measure of optimality is required, thus an objective function is defined for the problem. The objective function is usually a mathematical expression relating the design parameters to an economic measure for the design, e.g. total annualized cost or annual profit. Solving the optimization problem thus yields the optimum values of the decision variables that result in the best value of the objective function.

$$\begin{aligned} & \min f(\bar{x}) \\ s.t. & \\ & g(\bar{x}) \leq 0 \\ & h(\bar{x}) = 0 \end{aligned} \tag{2.1}$$

In equation 2.1, $\bar{x}$ is a vector of continuous variables, i.e. the decision and dependent variables, $f(\bar{x})$ is the objective function, while $g(\bar{x})$ and $h(\bar{x})$ are vectors of inequality and equality constraints, respectively.

If the equations representing the system are all linear, then the resulting optimization problem described by equation 2.1 is denoted a Linear Program (LP) for which effective solvers are readily available, e.g. the simplex method. Unfortunately most problems encountered in process engineering involve non-linear equations, thus resulting in Non-Linear Programs (NLPs). Solutions to NLPs must satisfy a specific set of conditions known as the Karush-Kuhn-Tucker (KKT) conditions (Karush, 1939; Kuhn and Tucker, 1951), and one widely used solver is the Successive Quadratic Programming (SQP) method (Biegler *et al.*, 1997). Unfortunately, the methods for solving NLPs are not able to guarantee global optimality unless the objective function and the feasible region is convex.

In many cases, an integral part of a process synthesis problem is the identification of optimal unit operations and/or equipment to be used in the process. If multiple options exist for a certain process task, e.g. two reactor types, then the solution must include the optimal choice between the alternatives. By adding a vector of integer variables $\bar{y}$, which often are Boolean, to the system in equation 2.1, such decisions/selections can be included in the optimization problem as presented in equation 2.2:

$$\begin{aligned} & \min f(\bar{x}, \bar{y}) \\ s.t. & \\ & g(\bar{x}, \bar{y}) \leq 0 \\ & h(\bar{x}, \bar{y}) = 0 \end{aligned} \tag{2.2}$$

Optimization problems that include discrete decision variables, e.g. for choosing between different reactor types, are often represented graphically using superstructures (Hostrup, 2002). If the resulting optimization problem consists only of linear equations the result is a Mixed Integer Linear Program (MILP), which can be interpreted as a series of LPs, one for each feasible combination of the discrete decision variables. Common solution strategies for MILPs involve the use of branch and bound techniques, where the optimum solution is identified by evaluating the search space of integer variables and comparing the objective function from the LPs generated by fixing the discrete variables.

If the optimization problem includes non-linear equations, then the result is a Mixed Integer Non-Linear Program (MINLP). Solving MINLPs involves the sequential solution of MILPs and NLPs, and the two primary methods for this task are Generalized Benders Decomposition (GBD) and outer-approximation. A comprehensive review of the use of optimization techniques in process synthesis and design is given by Grossmann and Daichendt (1996).

In sections 2.2.1 and 2.2.2, the prevailing existing solution approaches for reactor and separation system synthesis are presented. All the methods presented try to address the problems outlined by equations 2.1 and/or 2.2, however the methods can be divided into three groups:

- **Heuristic or knowledge-based approaches** - Solves only the constraint equations of the optimization problem, thus generating feasible, but not necessarily optimal solutions.

- **Mathematical programming approaches** - Solves the entire optimization problem including the objective function. Depending on problem complexity, current solvers cannot guarantee a solution much less a globally optimum one.

- **Hybrid approaches** - Utilizes heuristics, thermodynamic insights and/or other qualitative process knowledge to impose bounds on the optimization variables, thus reducing the search space and resulting in a less complex optimization problem to be solved by mathematical programming.

The reaction section of a process plant has an enormous effect on the design of the process, as the reactor effluent dictates the downstream processing required to achieve the desired products. Furthermore the efficiency of the reactor in terms of reactant conversion also determines the potential for recycles of unreacted materials. The reactor synthesis problem can be stated as follows (Biegler *et al.*, 1997):

For a given set of reaction stoichiometry, kinetics, process objectives and system constraints, determine the optimal reactor network structure and flow characteristics along with the appropriate mixing, heating and cooling.

Similarly, Barnicki and Fair (1990, 1992), Jaksland *et al.* (1995), and Hostrup (2002) have formulated the separation system synthesis problem as follows:

Given	1. Specification of the mixture to be separated
	2. Desired products and purities
	3. List of potential separation techniques
Determine	1. Separation tasks and techniques
	2. Sequence of tasks
	3. Appropriate conditions of operation

2.2.1 Heuristic and Knowledge-Based Approaches

Knowledge-based methods for reactor design and selection are often incorporated in expert systems, e.g. Krishna and Sie (1994) and Schembecker *et al.* (1995), where the design is performed in different stages with increasing level of detail. The expert system approach usually combines heuristic rules with simple numerical calculations to aid in the decision process.

Heuristic methods for separation system synthesis provide guidelines or rules of thumb to how a process flowsheet should be composed. The hierarchical decision framework introduced by Douglas (1985), divides the separation system into two parts, i.e. vapor recovery and liquid recovery. For each part, a set of rules exist for the selection of separation tasks, i.e the separations to be performed, and the separation techniques, i.e. the techniques to be employed in order to achieve the desired separations. Most heuristic methods for separation system synthesis utilize this two-level approach, e.g. Barnicki and Fair (1990, 1992) as well as Chen and Fan (1993). Recently McCarthy *et al.* (1998) introduced an automated procedure for product separation synthesis. First, the procedure performs an in-depth tree search to locate solutions and unit operations design variable discretization to reduce the search space. Then, based on stream specifications, alternative separation methods are proposed on the basis of different splits and mixers, until the desired products have been recovered. This methodology has the benefit of avoiding mapping into an apriori generated superstructure. The efforts of Glinos and Malone (1984, 1985, 1988) involved graphical synthesis methods based on triangular diagrams with residue curve maps or distillation boundaries. The work was extended by Malone and Doherty (1995) through the development of a comprehensive synthesis framework implemented in a computer program named MAYFLOWER. This software combined triangular diagrams and residue curves with heuristic rules in a hybrid synthesis approach capable of handling not only simple distillation but also complex column configurations as well as liquid-liquid extraction processes.

A special case of separation system synthesis, that has been the focus of a lot of research effort in the past, is the knowledge-based synthesis of separation systems, where distillation is the only available separation technique, e.g. Thompson and King (1972) and King (1980). Distillation is the most widely used separation technique in the chemical processing industry, and also accounts for a majority of the energy consumption. Therefore, it has been of interest to develop methods for identification of the optimum sequence of distillation columns required for a given separation task. Nishida *et al.* (1981) and Bek-Pedersen (2003) have presented comprehensive reviews of such methods.

2.2.2 Optimization and Hybrid Solution Approaches

The most common method for analysis of reactor networks is the attainable region analysis, which was introduced by Horn (1964). The analysis is based on

calculating a region within the composition space, that is attainable, i.e. feasible to achieve, by the reactors considered in the problem. There are several examples of methods for the construction of such regions, e.g. the geometric approach by Glasser *et al.* (1987) and the algorithmic method presented by Hildebrandt and Biegler (1995). An attractive feature of the attainable region approach is that these regions are convex, thus ensuring that the optimal reactor network lies within this region. Therefore the method allows for graphical analysis and solution of simple problems and in the case of more complex problems, the method can assist in the formulation of the constraints in a mathematical optimization problem. Balakrishna and Biegler (1992) and Kokossis and Floudas (1994) among others, have presented mathematical programming and superstructure approaches for reactor synthesis based on attainable region analysis.

Most purely optimization based approaches to separation system synthesis is limited to the design and sequencing of distillation columns only. Although distillation is a widely used separation technique, this is a serious limitation from a general process synthesis point of view. Thompson and King (1972) state, that the number of possible sequences NS that may separate NC components by NT potential separation techniques can be calculated using equation 2.3.

$$NS = \frac{(2 \cdot (NC-1))!}{NC! \cdot (NC-1)!} \cdot NT^{NC-1} \tag{2.3}$$

If only one separation technique, e.g. distillation is considered, then equation 2.3 is reduced to equation 2.4.

$$NS = \frac{(2 \cdot (NC-1))!}{NC! \cdot (NC-1)!} \tag{2.4}$$

The reduction in the size of the problem is significant, e.g. for 10 components to be separated and 5 available separation techniques including distillation, the number of possible sequences is reduced by a factor of 10,000 if only distillation is considered. However, the use of mathematical programming for solving even such reduced size problems requires that the distillation column and thermodynamic models are rigorous and accurate. Most implementations of purely optimization based approaches are in the form of structural optimization of superstructures. Andrecovich and Westerberg (1985) presented a framework based on a MILP model for the synthesis of heat integrated distillation sequences. Unfortunately the approach is limited to the separation of ideal mixtures as the component split calculations are linear. Aggarwal and Floudas (1990) presented a MINLP formulation that included nonlinear models of the separations, thus allowing the handling of more complex mixtures. An in depth review of the methods for distillation column sequencing has been presented by Bek-Pedersen (2003).

Since purely mathematical optimization based approaches are computationally intensive and require very detailed models for all the considered unit operations, it is desirable to reduce the search space prior to invoking the optimization solver. As an alternative to using heuristics in order to bound the optimization variables, Jaksland *et al.* (1995) suggested the use of thermodynamic insights for the sequencing and selection of separation techniques. The method consists of two main levels, where the first level calculates the difference in component properties as ratios. These ratios are subsequently used as screening criteria to identify the feasible separation techniques that might exploit those property differences. At the second level, more detailed mixture analysis is applied to achieve further screening and if needed mass separating agents for solvent-based separations are identified using the molecular design methods developed by Gani *et al.* (1991).

Recent efforts by Gani and Bek-Pedersen (2000) and Bek-Pedersen (2003) have introduced the concept of driving force based separation design. The method is based on the notion that by using the thermodynamic driving force as the selection and sequencing criterion, a near optimal design is obtained without having to resort to computationally intensive calculations. A design utilizing the maximum driving force at all stages should result in the most efficient separation system, and the methodology enables fast and easy identification of such designs.

Hostrup (2002) presented an integrated approach to the solution of synthesis, design and analysis problems. This integrated approach combines the thermodynamic insights of Jaksland (1996) and Gani and Bek-Pedersen (2000) with the formulation of structural optimization problems, thus allowing for efficient screening among the alternative routes. The integrated framework consists of a problem formulation step, an optimization step and a validation/analysis step. For the evaluation of candidate solvents, the methodology of Harper (2000) is utilized, which is an extension of the original method by Gani *et al.* (1991).

2.3 Process Integration

Over the past two decades, process integration has witnessed significant progress in the development of systematic methodologies, tools, and applications. In particular, two main branches of process integration have been developed: Energy integration and mass integration. Energy integration focuses on the system-level optimization of heat, power, fuel, and utilities (Linnhoff *et al.*, 1982; Linnhoff and Hindmarsh, 1983). On the other hand, mass integration is a holistic approach to the allocation, separation, and generation of streams and species throughout the process (El-Halwagi and Maniousiouthakis, 1989; El-Halwagi, 1997).

2.3.1 Heat Integration

Traditionally process integration has been synonymous with energy or heat integration. The design of utility systems is a very important part of designing a chemical processing facility. The task of designing a single heater or cooler to facilitate a temperature change within a process stream is fairly straight forward, however if each heating or cooling task is solved unit by unit, the overall utility consumption may end up being very high. It may be very beneficial to integrate several heating and cooling tasks to reduce the overall energy consumption. Consider the case where a stream must be heated prior to being fed to a reactor and the reactor effluent must be cooled to facilitate separation using a PT-flash. For such a simple scenario, the integration strategy is straightforward, i.e. the reactor effluent is used to preheat the feed, thereby reducing the overall heating and cooling load. Unfortunately it is not possible to readily identify the optimum integration strategy for more complex systems. As the process becomes more complex, the number of heat sources and sinks throughout the flowsheet increase dramatically and therefore numerous possible combinations for exchanging heat exist. The same common sense approach used in integrating the simple reactor system can no longer guarantee even a near optimal solution.

Synthesis of heat exchanger networks (HENs) deals with systematic approaches for identifying the optimal allocation of energy throughout the process. In order to address the energy allocation problem it is necessary to classify the process streams in a systematic way. In heat integration studies, a process cold stream is a stream that requires heating, while a process hot stream is a stream that needs to be cooled. El-Halwagi (1997) presents the solution of synthesizing an optimal HEN as identification of the answers to the following questions:

- Which heating and/or cooling utilities should be used?
- What is the optimal energy load to be removed or added by each utility?
- How should the process hot and cold streams be matched, i.e. stream pairings?
- What is the optimal system configuration, i.e. how should the heat exchangers be arranged? Do any streams require mixing or splitting?

The first method for identification of the minimum external utility load for a HEN was presented by Hohmann (1971). The work was further extended primarily by Linnhoff and coworkers (Linnhoff *et al.*, 1982; Linnhoff and Hindmarsh, 1983) by introducing systematic methods for visualizing and decomposing the synthesis problem.

Thermal pinch analysis is a tool for determination of the potential for internal heat exchange between the process streams, thereby reducing the need for external energy utilities. Using the hot and cold stream classification described

earlier, the streams are characterized using three steady state parameters, i.e. the supply temperature of the stream T_s, the target temperature T_t, and the heat capacity flowrate HC. The heat capacity flowrate is calculated using equation 2.5 for streams not encompassing a phase change, while phase changes are evaluated using the approximation in equation 2.6:

$$HC = m \cdot C_P \tag{2.5}$$

$$HC = \frac{m \cdot \Delta H}{1\mathrm{K}} \tag{2.6}$$

When these parameters have been calculated for all streams in the plant, a composite curve is constructed for all the process hot streams and analogously a composite curve is constructed for all the cold streams. In figure 2.2, the construction of a hot composite curve for a system of three hot streams with no phase changes is presented. The cold composite curve is constructed in the same manner and when the two composite curves are plotted together, the thermal pinch diagram is obtained. An example of a thermal pinch diagram is given in figure 2.3. From a thermal pinch diagram several quantities can be inferred, i.e. the area where both composite curves exist over the same enthalpy range represents the potential for internal process to process heat exchange. The areas at each end of the diagram represents the minimum requirements for external heating and cooling utilities. The final placement of the composite curves depends on an arbitrary choice of the minimum allowable temperature driving force ΔT_{min}. Reducing the minimum allowable temperature driving force reduces the overall consumption of external energy resources, because the two composite curves are moved closer to each other. Unfortunately by reducing the value of ΔT_{min}, the heat transfer area of the process heat exchangers also increase. Therefore the optimal choice of ΔT_{min} becomes a trade-off between the cost of external utilities and the cost of the heat exchangers. Once the value of ΔT_{min} has been fixed, the final placement of the composite curves can be performed. There are several strategies for obtaining the final pinch diagram, however two methods are the most common. One approach fixes the location of the hot composite and then moves the cold composite horisontally until the smallest distance between the two composite curves is exactly ΔT_{min}. This is then referred to as the thermal pinch point. The other approach utilizes a shift in the temperature scale for the cold composite by ΔT_{min} as shown in equation 2.7. By plotting the cold composite on the diagram using the hot composite temperature scale, it is ensured that the cold composite is always at least ΔT_{min} from the hot composite curve. Thus, by moving the cold composite horisontally until it touches the hot compsite curve, the pinch point is identified.

$$T_{HotScale} = T_{ColdScale} + \Delta T_{\min} \tag{2.7}$$

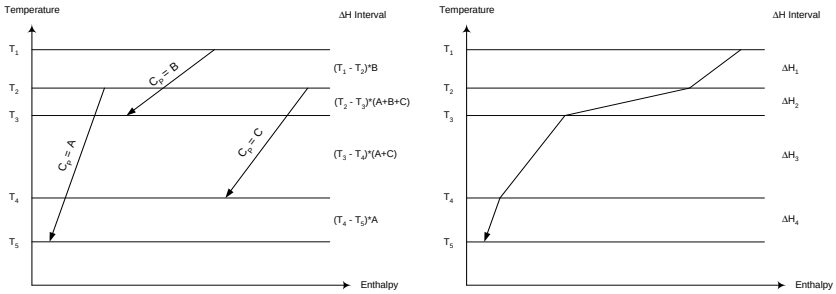

Figure 2.2: Construction of the hot composite curve from three streams

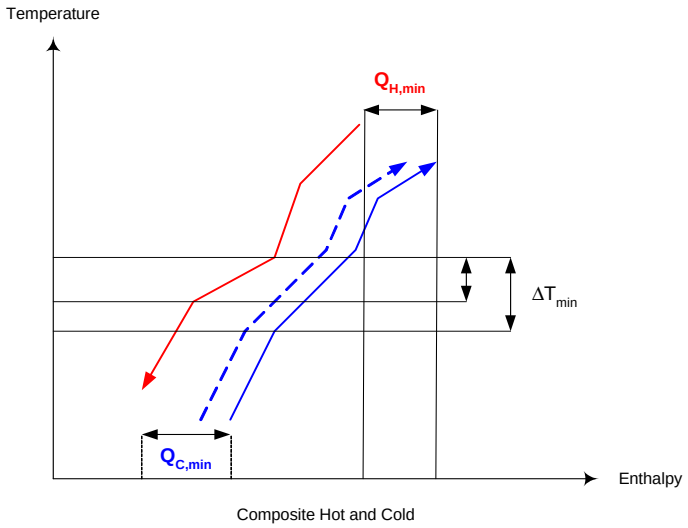

Figure 2.3: Thermal pinch diagram

The thermal pinch point represents a thermodynamic bottleneck, that allows for decomposition of the energy allocation problem into two sub-problems, i.e. one above and one below the pinch. It is important to emphasize, that thermal pinch analysis targets the optimal allocation of the energy resources, thus identifying the minimum utility requirements. This is a powerful result as it provides a target to which any HEN design can be compared, i.e. the optimum performance is identified first and then the designs capable of achieving the target are determined.

Once the thermal pinch point has been identified, Linnhoff *et al.* (1982) stated three rules for designing a HEN that achieves the minimum utility requirements identified by the pinch diagram:

Rule 2.1 *Do not transfer heat across the pinch.*

Rule 2.2 *Do not use external cooling utilities above the pinch.*

Rule 2.3 *Do not use external heating utilities below the pinch.*

If an amount of energy E is transferred across the pinch, then an overall enthalpy balance states that an additional equivalent amount E of external heating or cooling will be necessary to match the process heating and cooling demands. Therefore the use of external cooling above the pinch results in additional external heating requirement by an equal amount, and vice versa for the use of external heating below the pinch.

Along with the graphical solution approaches, systematic mathematic solution strategies were also developed (Linnhoff *et al.*, 1982). By introducing temperature intervals and evaluating the enthalpy balance around each interval, a cascading calculation sequence is obtained. This cascade approach has been implemented in several commercial software packages and subsequently used for achieving substantial energy savings. Gundersen and Naess (1988) presented a very useful industrial review of the current state of the art.

Thermal pinch analysis results in HEN designs that require the minimum amount of external utility, however when taking the equipment cost into account, the pinch analysis solution may prove not to be optimal. The first attempts at trying to identifying the true optimum solution were presented by Papoulias and Grossmann (1983), in the form of LP and MILP transshipment problems for sequential determination of minimum utility loads and minimum number of exchanger units. The minimum investment cost of a given network may be determined by formulation of a NLP problem (Floudas *et al.*, 1986). Since minimum utility cost does not necessarily result in minimum investment cost, a trade-off between the two arises. In order to avoid this trade-off, numerous examples of simultaneous solution approaches have been presented, e.g. Yee and Grossmann (1990), Ciric and Floudas (1991) and Briones and Kokossis (1999). Although significant advances have been made with respect to optimizing utility consumption and equipment cost, most of these approaches are limited by the assumptions of constant heat capacities, ideal thermodynamics

and so on. Nielsen *et al.* (1996) employed simulated annealing methods to solve problems, where detailed thermodynamics as well as property correlations had been included.

Conventional heat exchanger networks are examples of process integration measures, where energy is transferred from one stream to another. However, thermal integration can also be achieved by transferring energy from one part of a separation sequence to another. An example of such an implementation is an energy-integrated distillation column (Koggersbøl, 1995; Eden *et al.*, 2000), where the energy released in the condenser is used in the reboiler by employing an indirect heat pump. The heat pump raises the quality, i.e. the temperature level, of the energy released in the condenser, thus making it useable in the reboiler. Such integration strategies have been reported to reduce the primary energy consumption by up to 80%.

2.3.2 Mass Integration

In many process plants large amounts of energy are consumed for removal of small impurities in the terminal streams. Often these problems are solved by implementing an end-of-pipe solution, e.g. waste treatment facilities and/or additional separation units. Both approaches solve the problem, however by only evaluating the terminal streams, a wide variety of solutions are omitted, which might be significantly better. The mass integration framework (El-Halwagi and Maniousiouthakis, 1989; El-Halwagi, 1997) is a holistic methodology based on tracking the overall flow of mass throughout the process, thereby enabling the determination of optimal recovery and allocation of the species. Mass integration has proven to be a valuable tool when designing new processes or trying to optimize existing processes. It has been widely used in the pulp and paper industry as well as for VOC recovery from gaseous waste streams. Since it is based on equilibrium relationships as well as mass and energy balances, it is applicable to a wide range of processes.

In most cases the component to be removed from one stream requires a Mass Separating Agent (MSA) to assist in the separation, e.g. an adsorbent or a solvent. Similar to thermal pinch analysis, the streams are characterized by the flowrate, a supply composition and a target composition. Streams, that need to have the content of the targeted component reduced, are designated rich, while streams that can accept the targeted species are designated lean. Analogously to the thermal pinch method, El-Halwagi and Maniousiouthakis (1989) developed the mass pinch method for the design of Mass Exchange Networks (MENs). Instead of plotting temperature versus the amount of energy to be exchanged between the process hot and cold streams, composition and the amount of mass to be exchanged between the process rich and lean streams is plotted. The rich and lean composite curves are generated similarly to the hot and cold composite curves in conventional pinch analysis. An example of the resulting mass pinch diagram is given in figure 2.4.

The area where the two composite curves co-exist over the same composition

interval, represents the potential for internal process to process mass exchange, thus corresponding to the maximum amount of the targeted species that can be accepted by internal process streams. The area at the top of the mass pinch diagram represents the excess capacity of the internal process streams, which may be eliminated by reducing the flowrate of those streams. At the bottom of the diagram, the amount of the targeted species that needs to be removed by external MSAs can be identified. Just as in thermal pinch analysis, the pinch point is identified at the point where the minimum driving force is obtained. In mass integration studies, the driving force is the composition difference between the rich and lean process streams. The choice of driving force is a trade-off between the cost of the MSA and the cost of regeneration and pumping. It should be noted that it is not necessary to specify an overall minimum allowable composition difference, in most cases there will be a specific value of the minimum driving force for each MSA. Furthermore since the equilibrium expression also can be different for all the MSAs, El-Halwagi (1997) introduced the concept of corresponding composition scales. Employing this concept allows representation of all the process streams in the same composition diagram, similar to using hot and cold temperature scales in thermal pinch analysis, however in mass pinch analysis each MSA has it's own composition scale and is then mapped to the rich composition scale. The pinch point is identified by moving the lean composite curve vertically until it exists completely above the rich composite. The point closest to the rich composite curve, when the lean composite is completely above the rich composite, is designated the mass pinch point.

Along with the graphical solution approaches, systematic mathematic solution strategies were also developed (El-Halwagi, 1997). Analogously to thermal pinch analysis, by introducing composition intervals and evaluating the mass balance around each interval, a cascading calculation sequence is obtained. Availability of commercial software packages for mass integration studies is limited, however since the problems are linear (because of the equilibrium assumption), it is fairly straightforward to solve them using mathematical programming.

An important visualization tool developed within the mass integration framework is the source-sink mapping diagram (El-Halwagi and Spriggs, 1998). This tool allows for easy identification of direct recycle opportunities and is constructed by plotting the pollutant load, i.e. flowrate times composition, or flowrate versus composition. In figure 2.5 an example of a source-sink mapping diagram is presented, where the sources (streams carrying the targeted species) are designated by shaded circles while sinks (units capable of processing the sources) are represented by hollow circles.

Process and/or equipment constraints limit the range of pollutant composition as well as load or flowrate that each sink can accept. In figure 2.5, these constraints are represented by two bands, and the intersection of the bands constitutes the locus of acceptable composition and load for direct recycle. If a source, e.g. source A, lies within the acceptable zone, it can be directly

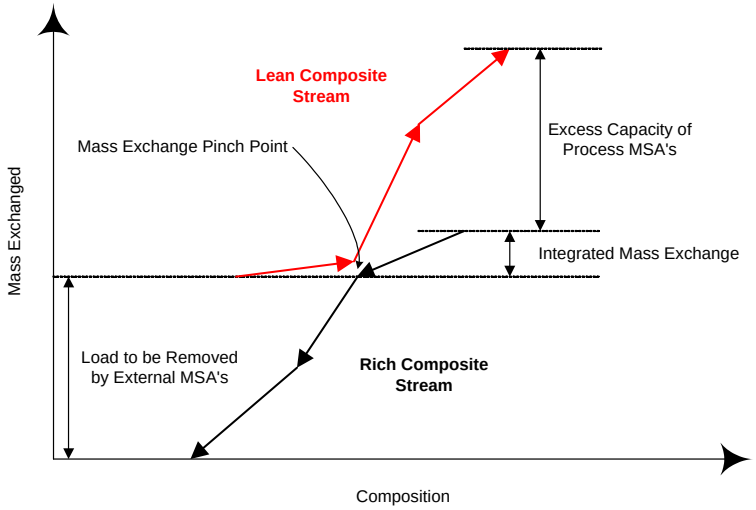

Figure 2.4: Mass pinch diagram

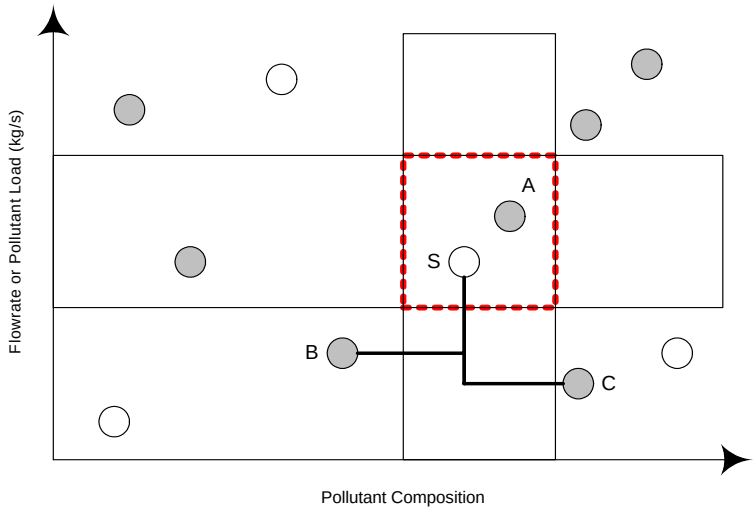

Figure 2.5: Identification of recycle opportunities using source-sink mapping

recycled to the sink, e.g. sink S. Furthermore, applying lever-arm principles allows for mixing sources B and C to create a mixture that may be recycled to sink S. Finally, the source-sink mapping diagram allows for determination of the required interception that would allow a given source to be recycled to the sink. El-Halwagi (1997) gives numerous examples of the use of source-sink mapping diagrams in the synthesis of MENs.

The main challenges encompassed when solving a mass integration problem is the identification of suitable MSAs, minimization of the MSA usage, and finally the total number of mass exchangers must be minimized. The framework for mass integration presented by El-Halwagi (1997) is based on a series of assumptions:

- All mass exchangers operate isothermally.
- The mass transfer equilibrium between the process rich and lean streams can be approximated by a linear expression.
- All suitable MSAs and their cost are known.

The assumption of isothermal operation is reasonable as long as no reactive MSAs are used. It is more questionable whether the assumption of linear equilibrium between the two phases is valid, especially if mass integration is to be used as a design methodology for other than very dilute systems. Furthermore since the optimal solution to the mass integration problem depends highly on the choice of MSAs, it is not desirable to use qualitative process knowledge to determine the candidate MSAs. In addition, the equilibrium between the two phases depends on variables such as temperature and composition, therefore the operating conditions will effect the overall cost of using the MSA.

These reservations aside, mass integration has proven to be a very powerful tool in e.g. waste recovery and pollution prevention, and has also been successfully applied to water usage minimization studies.

2.4 Product Synthesis and Design

In the chemical processing industry, the terms product synthesis and design designate problems involving identification and selection of compounds or mixtures that are capable of performing certain tasks or possess certain physical properties. Unlike in mechanical engineering, where the physical shape also is a consideration along with the capabilities, in chemical engineering the primary considerations are often limited to the physical and/or chemical properties of the product. Therefore, this review is also focused on the identification and design of compounds or mixtures that satisfy a given set of property constraints.

2.4.1 Computer Aided Molecular Design (CAMD)

In many cases, the solution to a separation synthesis problem requires the use of separation techniques, that utilize a solvent. Therefore a systematic

methodology for identifying or designing candidate solvents is required. Two primary methods exist for solvent identification, i.e. Computer Aided Molecular Search (CAMS) and Computer Aided Molecular Design (CAMD). The difference between the two approaches can be described by the type of problem they are solving. In CAMS, the problem is limited to selecting, i.e. identifying candidates from a database of known compounds, while in CAMD, the problem involves selection as well as design of the candidate compounds. Since CAMS does not involve any design, the topic is not explored further, however additional information on algorithms for CAMS can be found in, e.g. Joback (1994), Hermansen (1994), and Modi *et al.* (1996).

A CAMD problem can be described as a reverse property prediction problem. In property prediction, the physical properties of a compound are estimated based on structural information about the compound, while in CAMD molecular structures, that result in certain properties, are identified. Therefore the principal problem to be solved by CAMD algorithms can be stated as follows:

Given the desired properties of the molecules to be designed and a selection of molecular fragments (building blocks), determine feasible molecular structures with the desired properties

Several different solution approaches have been reported in the open literature and they can be classified using the following three categories:

- **Mathematical optimization techniques** - Odele and Macchietto (1993), Vaidyanathan and El-Halwagi (1994), Duvedi and Achenie (1996), and Pistikopoulos and Stefanis (1998).
- **Stochastic optimization techniques** - Marcoulaki and Kokossis (1998) and Venkatasubramanian *et al.* (1995).
- **Enumeration (generate and test) techniques** - Gani *et al.* (1991), Pretel *et al.* (1994), Friedler *et al.* (1998), and Harper (2000).

Harper and Gani (2001) presented a review of the different molecular design techniques, where it was pointed out that although the optimization based approaches could provide an optimum solution, they did suffer from computational intensiveness, if the problem was not well-defined or highly nonlinear. The enumeration based techniques also have their drawbacks, as they tend to suffer from combinatorial explosion. Harper (2000) presented a multi level generate and test approach, where molecules are generated from molecular building blocks using a rule based combinatorial approach, that guarantees the generation of only feasible molecules. At each level of method, the level of detail increases from simple group vectors on the lowest level to detailed 3-dimensional structures at the highest level. Before moving from one level to the next, the generated structures are evaluated against the design constraints, thus eliminating infeasible candidates at each level and reducing the risk of

combinatorial explosion. The general procedure for the formulation and solution of a CAMD problem is given in figure 2.6. The first step is to formulate the objective or target for the investigation, which is then converted to a set of numerical constraints along with some molecular fragments or building blocks to be included in the generation steps. This constitutes the pre-design phase, the design phase involves the generation of molecular structures and testing against the property constraints at each level. The post-design phase involves ranking of the candidates and evaluation of properties that can not be estimated from structural information, e.g. environmental impact, health and safety aspects.

The post-design evaluation of certain properties implies that the solution of CAMD problems is often iterative, since not all problems can be solved and some solutions are unacceptable. Figure 2.7 illustrates this iterative nature of CAMD algorithms. The methodology developed by Harper (2000) has been implemented in a software tool called ProCAMD, which is a part of the Integrated Computer Aided System (ICAS) (CAPEC, 2003). It should be noted that although CAMD is a very powerful tool for solvent design and selection, the method is currently limited to identification of hydrocarbon structures, which are the compound types covered by the group contribution based prediction models. Aqueous solutions as well as ionic and polymeric fluids can not be identified using current CAMD methods. Inclusion of such compounds requires updating the prediction methods with the appropriate group parameters and group descriptions.

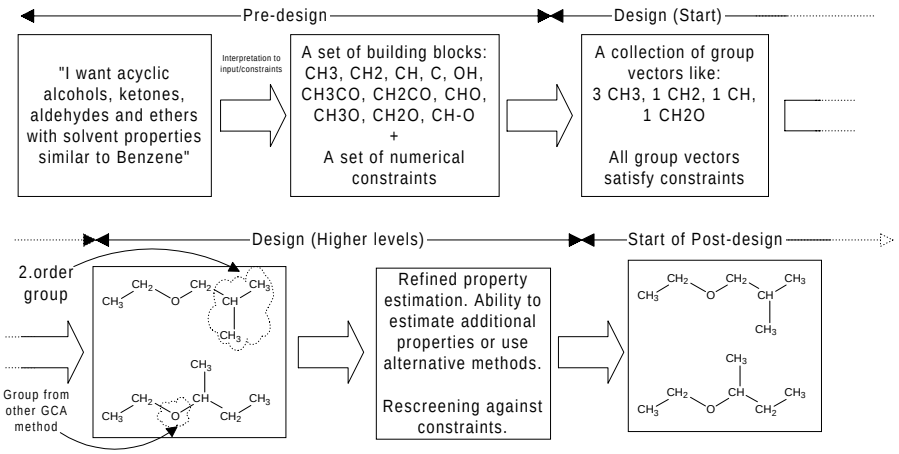

Figure 2.6: Schematic representation of the formulation and solution of a CAMD problem

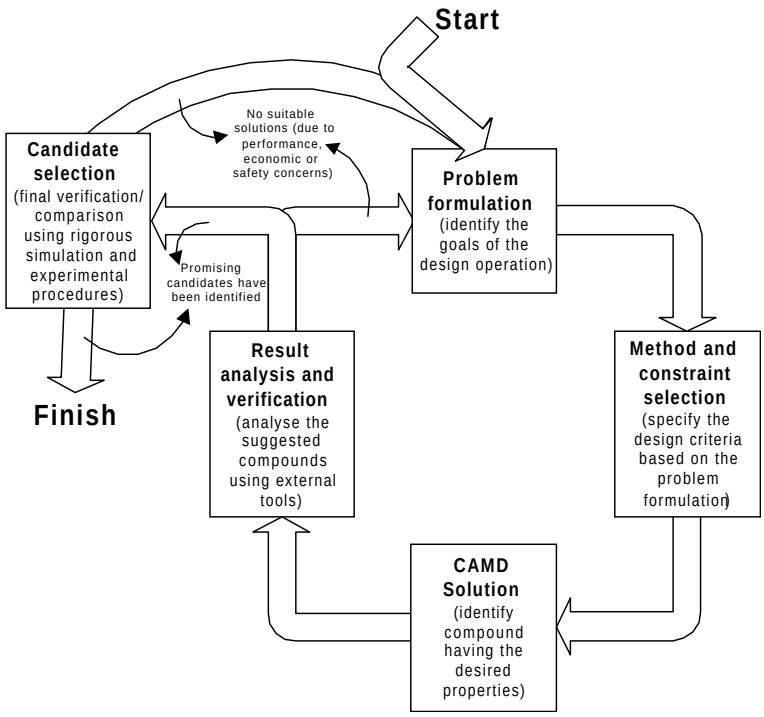

Figure 2.7: Conceptual flow diagram of the multi level CAMD framework developed by Harper (2000)

2.4.2 Design of Experiments (DOE)

Many product design problems involve the identification of an optimal mixture of a given set of constituents, which achieves a specified performance objective. In some cases, mixture data is available, but most likely it will be necessary to perform experiments to obtain the required mixture data. Now important questions arise, i.e. which parameters should be investigated for their effect on the mixture? How can the experiments be performed most efficiently, i.e. provide the most information using the minimum number of trials?

The traditional approach to systematic experimentation requires changing one factor, e.g. content of one constituent, at a time (OFAT). Unfortunately, the OFAT strategy does not provide any data on the interactions of different factors, which are very likely to occur in chemical processes. Therefore, an alternative strategy called two-level factorial design (TLFD) has been developed. The TLFD strategy provides a parallel testing scheme, which is much more efficient than the serial OFAT approach. The TLFD approach is based on statistical methods, where all factors are adjusted simultaneously at two levels, i.e. their low and high values. By initially limiting the tests to only two levels, the number of experiments is minimized. In most cases, this evaluation should give enough information to identify improvements, however additional levels can be added if necessary, e.g. the midpoint value of a given factor can be included to increase the resolution. The fundamentals of TLFD has been presented by several authors, e.g. Box *et al.* (1978), Cornell (1990), and Montgomery (1991), however the application to chemical engineering was pioneered by Anderson and Whitcomb (1996).

Factorial design methods have proven very effective in reducing the number of experiments required to produce adequate estimates of main effects and simple interactions. Unfortunately, there are applications where factorial methods are not the ideal choice. In some cases, the response to a change in a factor depends on the proportions of the ingredients such as in the chemical and food industry. As an example of misleading results obtained from a factorial design method, the production of lemonade is presented in table 2.1 (Anderson and Whitcomb, 1998).

Run	Lemons	Water (cups)	Lemon to Water Ratio	Taste
1	1	1	1.0	Good
2	2	1	2.0	Sour
3	1	2	0.5	Weak
4	2	2	1.0	Good

Table 2.1: Erroneous factorial design for lemonade

The results of the analysis show that the taste of the lemonade in runs 1 and 4 is the same, despite the contents being different, i.e. in run 1 both factors are low, while in run 4 both factors are high. Therefore, it would

make more sense to evaluate taste as a function of the proportion of lemons to water, not the amount. Statistical mixture design accounts for the dependence of the response on proportionality of the ingredients. Since most chemical engineering applications, especially product design problems will depend only on proportions, mixture design methods are called for.

As the availability of high performance computers has increased dramatically over the last two decades, it has become practical to apply advanced statistical matrix-based algorithms to the identification of optimal sets of experiments. Software implementations of such statistical methods, e.g. Design Expert (Stat-Ease, 1999), provides user friendly assistance in many of the initial steps necessary prior to invoking the advanced algorithms. The general procedure for achieving an optimal design of experiments is as follows (Anderson and Whitcomb, 1999):

1. Specify the polynomial order, i.e. first, second, third or beyond, that is needed to model the response
2. Generate a "candidate set" with more than enough points to fit the specified model
3. From the candidate set, select the minimum number of points needed to fit the model

Each of the investigated factors will have constraints that need to be taken into account, and these restrictions along with the choice of polynomial order constitutes the initial input to the algorithm. After the statistical analysis, the result is a sub-set of experiments that yields the maximum information using a minimum of experiments. Once the experiments have been conducted and the response values measured, a model is generated for each response, which now can be used for generation of response plots encompassing the entire factorial space, thus allowing for the identification of truly optimal mixtures.

In order to illustrate the application of mixture design methods, a simple example was presented by Anderson and Whitcomb (1999), where the producers of an automotive coating explored the effects of three components (monomer, crosslinker and resin) on two key responses, i.e. the hardness and the solids content. After application of the optimal design of experiments algorithm and the subsequent data acquisition, response plots for the hardness and solids content are generated. These plots are shown in figures 2.8 and 2.9, respectively. In figure 2.8 the contour labeled 10 represents a crucial boundary, i.e. to the right of this contour the hardness drops below specification. Similarly, in figure 2.9, the crucial boundary lies on the contour labeled 50. By combining the two plots and shading out the areas, that lie outside the specifications, i.e. hardness less than 10 and solids content less than 50, an overlay plot can be generated, which reveals a "sweet spot" for the formulation of the clearcoat. In figure 2.10, the overlay plot is presented with an arbitrarily placed flag in the middle of the operating window. Numerical search techniques may employed to identify the true optimum solution by including additional ranking criteria such as

cost. It should be emphasized however, that any recommended formulations should be verified by confirmation runs, since the identified optimum is based on predictive models, which in turn are based on a very select, but optimal set of sample blends.

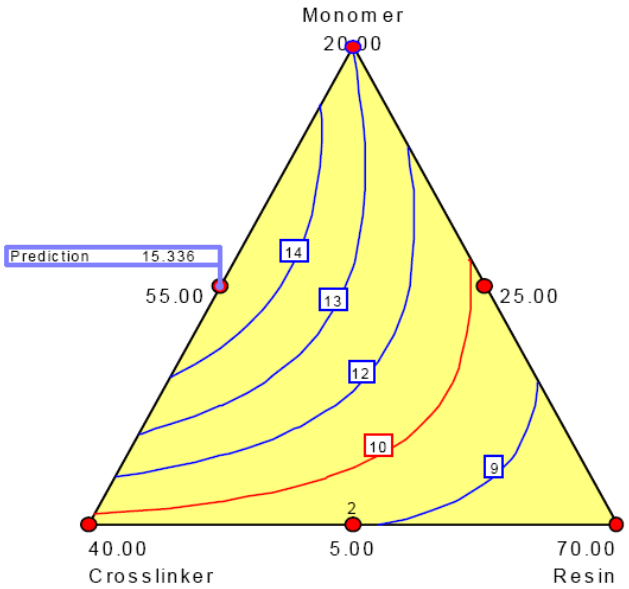

Figure 2.8: Response plot of hardness for coating mixture design

2.5 Summary

In this chapter an overview of some of the prevailing methods for process and product design has been presented. These methods and the mindset they represent form the basis for the property based design techniques presented in this thesis. The principal concepts and benefits that can be ascertained from the different methods are as follows:

- Identification of design targets without performing detailed calculations
- Visualization of problem and solution can provide valuable information
- Using process insights to formulate well defined optimization problems

Being able to target the optimum solution to a problem from the beginning is a very efficient strategy. Examples of such targeted approaches include the driving force based separation design techniques, where the specification of achieving maximum driving force results in designs with the easiest separation

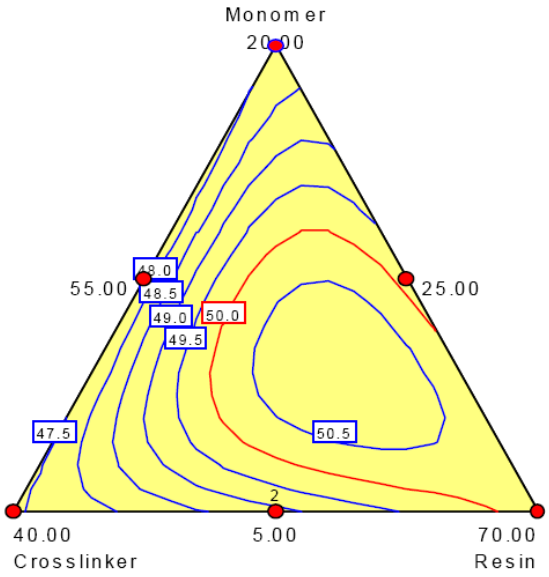

Figure 2.9: Response plot of solids content for coating mixture design

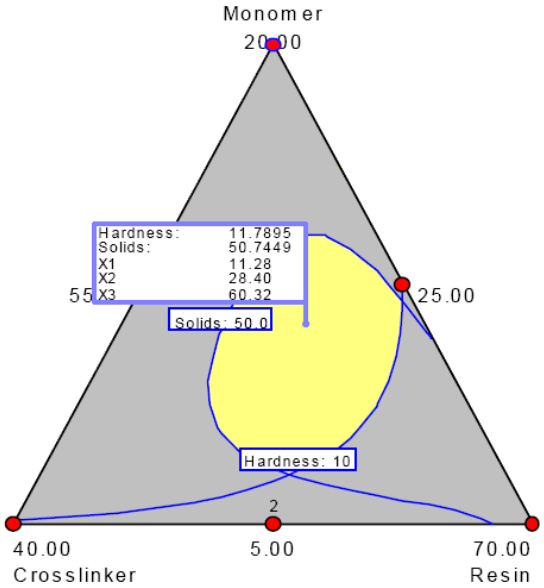

Figure 2.10: Overlay plot of responses for coating mixture design

and therefore minimum energy requirement. The pinch methodologies used in thermal and mass pinch analysis also target optimum performance by identification of thermodynamic bottlenecks, thereby assisting in the determination of optimum resource conservation strategies. Similarly the CAMD algorithms also establish the targeted performance prior to invoking any detailed calculations. The main benefit of establishing the optimum target values beforehand is enabling the designer to evaluate the quality of the generated solutions right away. For example when designing a heat exchanger network, the maximum level of internal heat exchange and thus the minimum utility requirements can be identified by means of pinch analysis. This solution may not be practical for several reasons, however it provides a target to which more practical designs may be compared, thereby giving a qualitative as well as quantitative measure of the divergence from optimality since no solution can be found that performs better.

Visualization is a powerful tool to obtain some initial insights to the solution of a given problem and may also be used for quick screening of alternative solutions. An inherent benefit of visualization is that the dimensionality of the problem has to be reduced to two or three dimensions in order to produce useful illustrations. This reduction in complexity alone may provide a useful simplification of the problem. In the driving force based design method, plots of the achievable driving force or relative separability of the targeted species can provide a visual representation of the proper selection and sequence of unit operations as well as operating conditions. Similarly the pinch diagrams and source-sink mapping diagrams used in process integration studies can provide insights on feasible recycle opportunities, thereby reducing raw material consumption and minimizing waste generation. Finally the use of response plots in the design of experiments methodology allows for visualization of certain key properties and the identification of optimal regions within the search space.

In some cases the complexity of the problem does not allow for visual solution, thus mathematical optimization is called for. However if the problem is highly nonlinear then available solvers cannot guarantee an optimal solution, thus the correct formulation of the problem is imperative. For example, the integrated approach to process synthesis described earlier, utilizes thermodynamic insights to formulate a well defined structural optimization problem. Similarly the design HENs and MENs relies on pinch analysis to identify the thermodynamic bottlenecks prior to formulating an optimization problem to determine a network design, that can achieve the targeted performance.

The development of any novel methodologies should incorporate the benefits of the principal concepts outlined in this section. Such methods should be capable of determining the optimum target values with only a minimal level of detailed calculations, thereby defining the desired performance of the solution apriori. If possible, visualization tools should be developed along with systematic strategies for performing design calculations visually. If the problem can not be solved graphically, then the visualization of the problem should assist in the formulation of a well defined optimization problem.

3

PROPERTY MODELS

3.1 Introduction

In most calculations involving simulation and/or design of a chemical product and/or process, an essential part is the mathematical model representing the product or process. Russel *et al.* (2002) describe a mathematical model for a product or process as consisting of three types of equations, i.e. balance equations, constitutive equations and constraint equations. In this representation, the constitutive equations consist of a set of selected property models and the constraint equations represent the conditions of equilibrium. Therefore, the property models affect the simulation and design results and should be chosen appropriately. Furthermore the same property models may play different roles in the simulation and design calculations (Gani and Pistikopoulos, 2002).

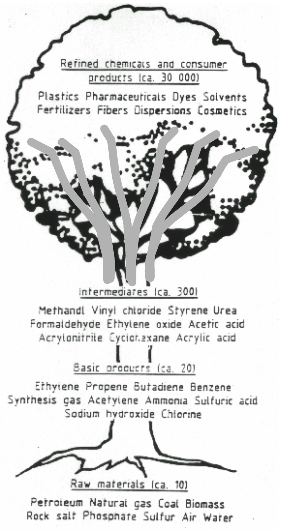

Figure 3.1: The product tree - revisited

The insights outlined by Gani and O'Connell (2001) can be illustrated by revisiting the product tree as presented in figure 3.1, where the three sets of grey

lines denote different constitutive or property models. The choice of a constitutive model implicitly defines the search space, thus determining the fruits that may be obtained from the product tree. This observation gives rise to two important questions:

- Do we know enough to derive a single property model?
- Can we solve simulation and/or optimization problems with multiple constitutive models representing the same variable?

In order to address these questions it is important to investigate the different roles that the constitutive or property models play at different stages in solving simulation and design problems.

3.2 Roles of Property Models

Recent efforts by Gani and O'Connell (2001) have highlighted the roles of property models in computer aided process/product engineering (CAPE) applications. According to Gani and O'Connell (2001), the property models play three distinctive roles in CAPE applications - a service role (provides values for a specified set of properties when requested), a service plus advice role (in addition to property values, the property models are used to advice on the feasibility of operation/design), and a service/advice/solve role (in addition to the above, the property models contribute directly in the method of solution). Therefore, the roles of the property models need to be considered during their selection as well as during the development of design (simulation) algorithms. Typically, properties play a service role in simulation and a service plus advice role in design. This section highlights the use of property models in product/process simulation and design in terms of the roles property models are expected to play and how to make the selected models play their expected roles more efficiently.

The most common role of a constitutive or property model is to provide property values for a component or a mixture. Figure 3.2 shows how the property model, in the service role, fits in the overall solution of a design problem. The property model requires input such as temperature, pressure, compositions and component identities, and based on these data a set of properties is calculated. For example, during simulation of a distillation operation, the property models provide values for fugacity coefficients, enthalpies, etc., when requested. The property model can then be connected to the process model, which in case of a chemical production facility represents the conversion of raw material into products. This model now requests the property values during solution and also provides the input to the property model. The process model forms the basis of a process simulator, which allows for evaluating the effect of changing process parameters. If the process simulator is combined with a process synthesis/design algorithm, the results of each simulation can be used to update

the process parameters, thus resulting in operating conditions for the process that gives satisfactory performance.

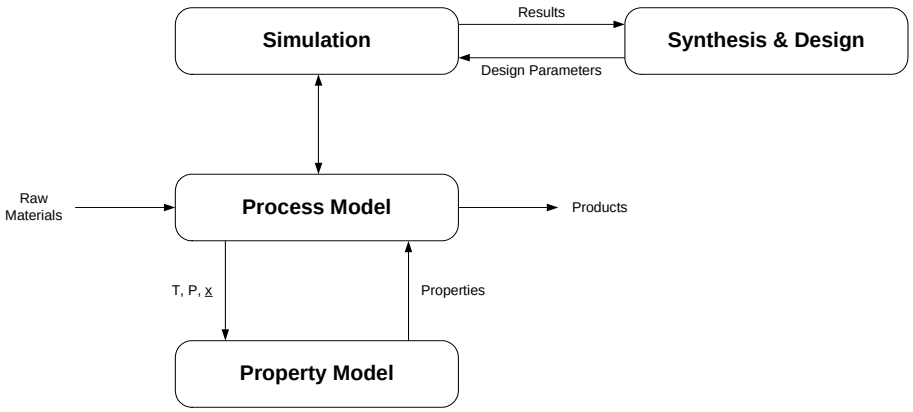

Figure 3.2: Property model providing a service to the process model

The property model can also be used to identify constraints on feasible conditions of operation and optimum values of process conditions, thus providing advice to the synthesis and design algorithms in terms of reducing the solution space. For design of distillation columns (operations), however, calculated relative volatility values indicate the feasibility of separation while their derivatives with respect to the intensive variables indirectly influences the choice (determination) of the optimal condition of operation. Figure 3.3 shows the connection between the different modules, when the property model is working in this role.

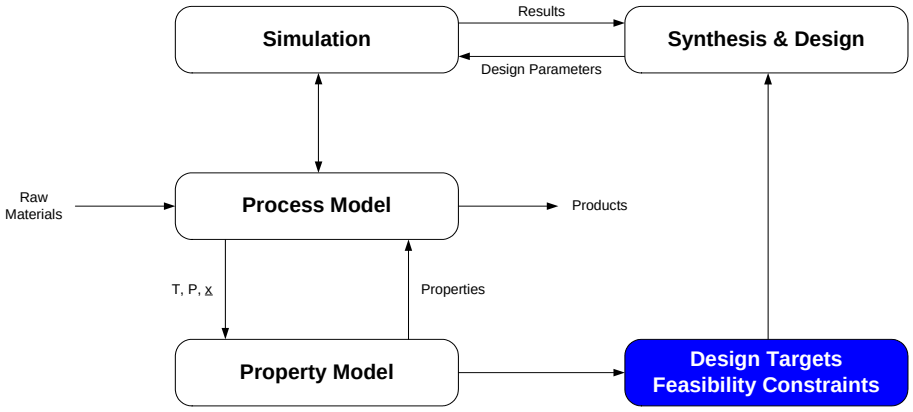

Figure 3.3: Property model providing service and advice to the process model

Since the property model can provide design targets along with constraints on feasible property values, it should be possible to include the property model

as a part of the solution routine, thus adding a solve role to the service and advice roles. Figure 3.4 illustrates this implementation. It should be noted that by using the property model in the solve role, the constitutive equations are decoupled from the process model and solved separately. Furthermore it must be emphasized that by performing this decoupling, the information flow to and from the property model is also reversed, i.e. the process model is solved for the values of the constitutive variables (properties), and then the property model is solved to yield the corresponding intensive variables, e.g. temperature, pressure, compositions and/or component identities.

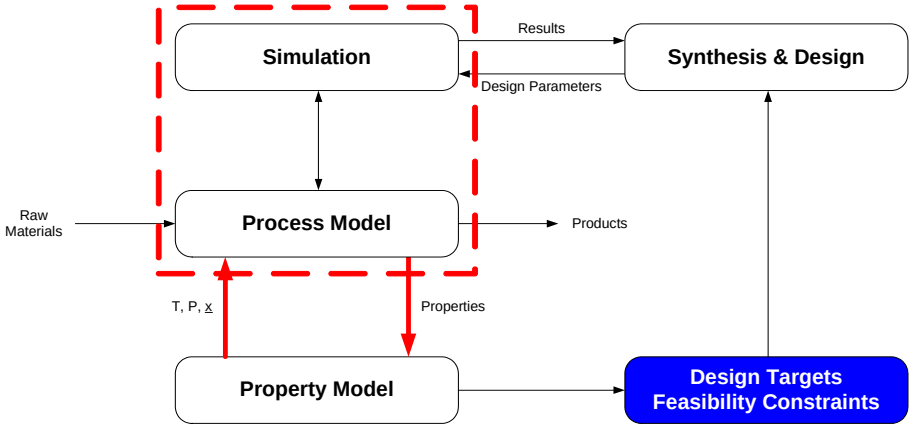

Figure 3.4: Property model providing service, advice and solutions to the process model

Now it is possible to answer the two questions stated in the beginning of this chapter. It is doubtful whether it is possible to identify one single property model capable of handling all phenomena encompassed in a given system. Therefore in order to adequately describe such complex processing systems, a solution methodology capable of handling multiple constitutive models is called for. By setting up the problem to use the property model in the solve mode, it is possible to use different property models at different stages of the solution.

4

REVERSE PROBLEM FORMULATIONS

4.1 Introduction

Traditionally process design and molecular design have been treated as two separate problems, with little or no feedback between the two approaches. Each problem has been conveniently isolated or decoupled from the other. Figure 4.1 shows a schematic representation of the two problems, e.g. the required inputs and solution objectives of the different design algorithms. Both approaches have some inherent limitations due to the amount of information that is required prior to invoking the design algorithm. When considering conventional process design methodologies, the selected species are chosen from among a list of pre-defined candidate components, therefore, limiting performance to the listed components. On the other hand, with molecular design techniques, the desired target properties are required input to the solution algorithm. Once again these decisions are made ahead of design and are usually based on qualitative process knowledge and/or experience and thus possibly yield a sub-optimal design.

To overcome the limitations encompassed by decoupling the process and molecular design problems, a simultaneous approach as outlined in figure 4.2 is proposed. Using this approach the necessary input to the methodology is the molecular building blocks and the desired process performance, for the molecular and process design algorithms respectively. The final outputs of the algorithm are the design variables, which facilitate the desired process performance target and the molecules that satisfy the property targets identified by solution of the process design problem. The strength of this approach is the capability to identify the property values that correspond to the optimum process performance without committing to any components at this stage. These property values are then used for the molecular design, which returns the corresponding components. One inherent problem with this approach is the need to solve the process design problem in terms of properties and not components. The conventional decoupled solution methodology presented in figure 4.1 can be described as a "forward" problem formulation, whereas the simultaneous solution approach given in figure 4.2 consists of solving two "reverse" prob-

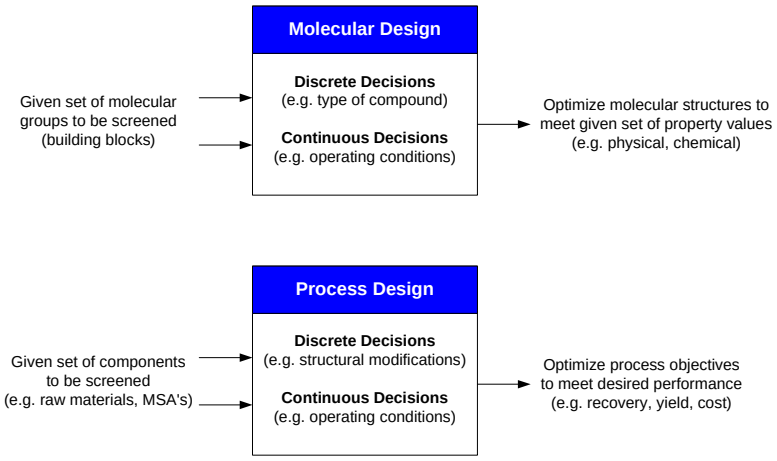

Figure 4.1: Conventional solution approach for process and molecular design problems

lems. Solving the process design problem in terms of properties corresponding to the desired process performance identifies the property design targets. In principle this part is the reverse of a simulation problem. Similarly solution of the molecular design problem to identify candidates that match the optimal design targets is the reverse of a property prediction problem.

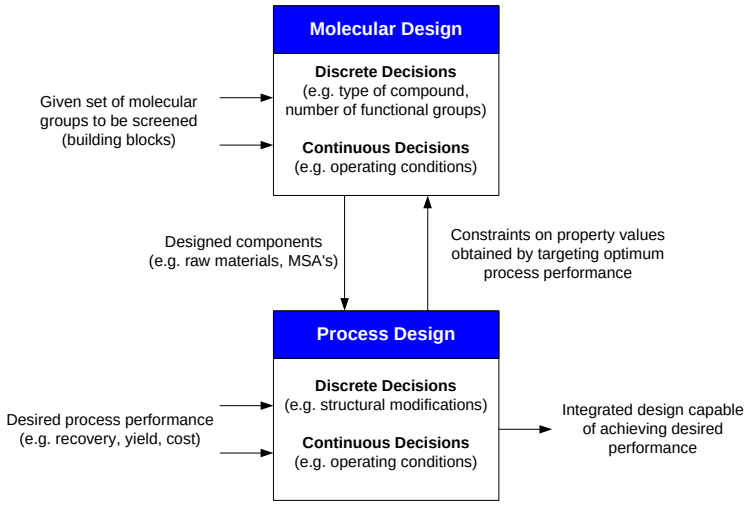

Figure 4.2: New approach for simultaneous solution of process and molecular design problems

Thus the design targets are described by constitutive variables like physical

properties and/or phenomena, such as reaction rates, equilibrium constants and mass/energy transfer rates. The constitutive variables are related to the process (intensive) variables such as temperature, pressure and composition through a constitutive model.

4.2 Problem Definition

A general process/product synthesis and/or design problem can be represented in generic terms by equations 4.1-4.6, where F_{obj} is the objective function that needs to be minimized or maximized in order to satisfy the desired performance criteria; $\mathbf{x}$ and $\mathbf{y}$ are the optimization real and integer variables respectively; h_1 represents the process model equations including the transport model; h_2 represents process equality constraints; g_1 and g_2 represent other process/product related inequality constraints, while equation 4.6 represents structural constraints related to process as well as products.

$$F_{obj} = \min \left\{ \mathbf{A}^T y + f(\mathbf{x}) \right\} \tag{4.1}$$

s.t.

$$h_1 \left(\frac{\partial \mathbf{x}}{\partial \mathbf{z}}, \mathbf{x}, \mathbf{y} \right) = 0 \tag{4.2}$$

$$h_2(\mathbf{x}, \mathbf{y}) = 0 \tag{4.3}$$

$$g_1(\mathbf{x}) > 0 \tag{4.4}$$

$$g_2(\mathbf{x}, \mathbf{y}) > 0 \tag{4.5}$$

$$\mathbf{B} \cdot \mathbf{y} + \mathbf{C} \cdot \mathbf{x} > d \tag{4.6}$$

Although the model equations themselves may be static, the transport model will most likely consist of ordinary differential equations. Furthermore, the model structure given by equation 4.2 is limited to simple transport models, where there is no coupling between the constitutive equations and the flow field. If spatial gradients are negligible, then equation 4.2 reduces to:

$$h_1(\mathbf{x}, \mathbf{y}) = 0 \tag{4.7}$$

It is important to point out, that all synthesis/design problems may be described using this generalized set of equations. Depending on the specific problem some terms and equations may be omitted, e.g. determination of

only feasible solutions will not require equation 4.1. It must be emphasized however, that regardless of the problem type a process model represented by equation 4.2 is needed and it is the model type and validity ranges that defines the application range of the solution. Hence heuristic and graphical methodologies resulting in a feasible but not necessarily optimal solution, as well as mathematical programming techniques, that determine optimal solutions can be employed by defining and solving equations 4.1-4.6.

The ease of solution of the overall problem as well as the attainable solutions are governed primarily by the complexity of the process model. Therefore it is reasonable to investigate if the solution of the process model can be simplified.

4.3 Reformulation Methodology

In principle the process model equations consist of balance equations, constraint equations and constitutive equations (Russel *et al.*, 2002). The model type and complexity is implicitly related to the constitutive equations, hence decoupling the constitutive equations from the balance and constraint equations will in many cases reduce the model complexity considerably. The concept is shown in figure 4.3, where the result of decoupling the constitutive equations is illustrated, and also provides the foundation for two reverse problem formulations:

1. Given input stream(s) variables, equipment parameters and known output stream(s) variables, determine the constitutive variables.

2. Given values of the constitutive variables, determine the unknown intensive variables (from the set of temperature, pressure and composition) and/or compound identity and/or molecular structure.

The first problem is the reverse of a simulation problem, i.e. it determines the property design targets for a given set of specified inputs and outputs. The second problem matches the calculated targets, for the process conditions, process flowsheets or products (including molecular structure). As long as the targets are matched, the process model equations (minus the constitutive equations) do not need to be solved again. It should be emphasized that optimization problems based on reverse simulation problems, are not limited by the application range or complexity of the constitutive equations. Therefore the solution is easy and can be visualized. Another advantage is that for the second reverse problem, any number of independent models may be used, as long as they match the target constitutive variable values. This implies that more than one process and/or product can be identified by matching the design targets, thus it is possible to determine all feasible solutions. Once the feasible solutions have been identified, the optimal solution may be found by ranking the solutions according to a performance index.

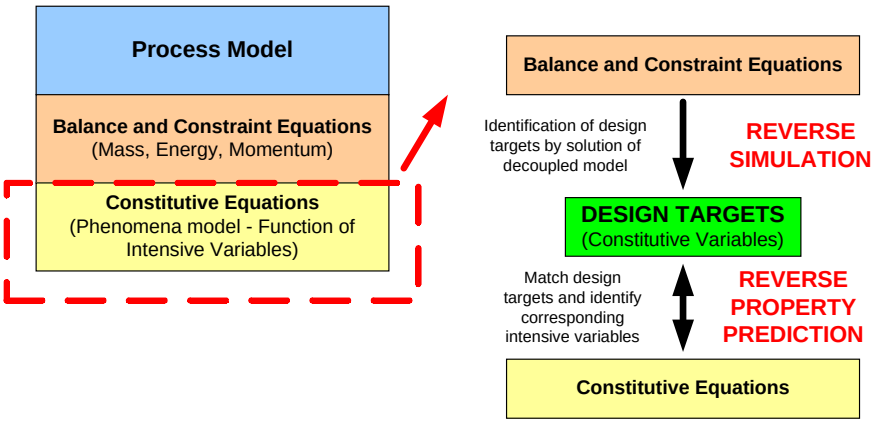

Figure 4.3: Reverse problem formulation methodology

4.3.1 Example of Reverse Problem Formulations

A simple, yet illustrative example of the application of reverse problem formulations is the solution of a solvent-based mass exchange problem. Consider a wastewater stream with a flowrate of F_w and a molefraction of phenol of X_1. Due to environmental regulations the phenol content must be reduced to a molefraction of X_2. Using the reverse problem formulation methodology, solution of the balance equation provides the necessary maximum solubility S (calculated as amount of phenol per unit mass of solvent):

$$S = \frac{(X_1 - X_2) \cdot F_w}{F_s} \tag{4.8}$$

Since the solvent flowrate F_s is unknown, the design target (constitutive variable) is determined as the solvent capacity, i.e. the solubility multiplied by the solvent flowrate:

$$S \cdot F_s = (X_1 - X_2) \cdot F_w \tag{4.9}$$

Solution of this problem is the reverse of a simulation problem. A conventional simulation problem requires the specification of the input conditions, i.e. F_w, X_1, S and F_s and then the resulting outlet composition X_2 can be calculated. By specifying the desired outlet composition and solving for the solvent capacity instead, the simulation problem is reversed.

The second problem, i.e. the identification of candidate solvents, which match the design target, is a reverse property prediction problem. Conventional property prediction calculates the physical properties of a compound

based on molecular structure. A reverse property prediction problem identifies molecular structures possessing a given set of properties. Reverse property prediction is often referred to as Computer Aided Molecular Design (CAMD) (Harper, 2000). In this reverse property prediction problem, it is possible to use a database, liquid activity coefficient models as well as correlations such as solubility parameters or any other suitable model in order to determine S and then select various values of F_s that matches the target solvent capacity. Using a ranking approach based on e.g. total solvent cost the optimal solution is identified. It is important to point out that solution of the reverse property prediction problem identifies all feasible solutions, thereby ensuring that the truly optimal solution can be found.

4.4 General Solution Strategy

The methodology for solving the integrated process and product design problems is divided in three parts. Below the different steps of the method are presented, steps 1 and 2 constitute the input specification and model generation steps, step 3 formulates and solves the reverse simulation problem, while step 4 solves the reverse property prediction problem. Finally step 5 identifies the optimal solution by employing a ranking approach.

1. Specify the synthesis/design problem in terms of known inputs and known outputs (for new processes and products) and/or equipment parameters (for retrofit problems).

2. Select the unit operations to be considered and generate the corresponding individual process models (without the constitutive equations).

3. Formulate and solve the reverse simulation problem with the constitutive variables as the unknown (design) variables that match a specified design target (can be solved as optimization problem or simply as a reverse simulation problem).

4. Formulate and solve the reverse property prediction problem in order to determine the conditions of operation, flowsheet structure and/or product that match the target values identified in step 3.

5. Compute the performance index for all feasible solutions from step 4 and order them to determine the optimal solution.

The benefit of employing the reverse problem formulation technique is that by solving for the constitutive variables directly and thereby identifying the design targets, the solution becomes simpler since the (often complex) constitutive equations have been decoupled from the process model.

It is important to point out that the solution strategy presented above is valid for new process synthesis/design problems as well as for retrofit problems. The

different problem types define the choice of equations and variables, but the overall strategy remains the same.

5

PROPERTY CLUSTERING

5.1 Introduction

Standard techniques for process design are based on individual chemical species. Therefore, tracking, manipulation, and allocation of species are key design tools starting with component materials balances to process modeling equations that are based on targeted species. But, should components always constitute the basis for process design and optimization? Interestingly enough, the answer is no! Many process units are designed to accept or yield certain properties of the streams regardless of the chemical constituents. For instance, the design and performance of a papermaking machine is based on properties (e.g., reflectivity, opacity, and density to name a few). A heat exchangers performance is based on the heat capacities and heat transfer coefficients of the matched streams. The chemical identity of the components is only useful to the extent of determining the values of heat capacities and heat transfer coefficients. Similar examples can be given for many other units (e.g., vapor pressure in condensers, specific gravity in decantation, relative volatility in distillation, Henry's coefficient in absorption, density and head in pumps, density, pressure ratio, and heat capacity ratio in compressors, etc.).

Since properties (or functionalities) form the basis of performance of many units, it will be very insightful to develop design procedures based on key properties instead of key compounds. The challenge, however, is that while chemical components are conserved, properties are not. Therefore, the question is, whether or not it is possible to track these functionalities instead of compositions? The answer is yes!

5.2 Definition of Property Clusters

The essence of this novel approach is to develop conserved quantities called clusters that are related to the non-conserved properties. The clusters can be described as functions of the raw physical properties themselves. The clusters are obtained by mapping property relationships into a low dimensional domain,

thus allowing for visualization of the problem. The clusters are tailored to possess the two fundamental properties of inter- and intra-stream conservation, thus enabling the development of consistent additive rules along with their ternary representation (Shelley and El-Halwagi, 2000).

5.2.1 Property Operator Functions

The clustering approach utilizes property operators, which are functions of the original raw physical properties. For each process stream s, the properties P_{js} may be evaluated using one of the following options:

- Experimental measurements.
- Empirical correlations.
- Process simulation.
- First moments, i.e. integral or algebraic averaging.

The property operator functions describe the class of properties that can be described by the following generalized mixing rule:

$$\psi_j(P_{jM}) = \sum_{s=1}^{N_s} \frac{F_s}{\sum_{s=1}^{N_s} F_s} \cdot \psi_j(P_{js}) = \sum_{s=1}^{N_s} x_s \cdot \psi_j(P_{js}) \tag{5.1}$$

In equation 5.1, $\psi_j(P_{js})$ is an operator on the j'th property P_{js} of stream s. The property operator formulation must be such, that it allows for simple linear mixing rules, i.e. the operators correspond to the actual properties as given in equation 5.2 or the operators may describe functional relationships of the properties, e.g. for density, where the resulting property of mixing two streams is given as the inverse of the summation over the reciprocal property values multiplied by their fractional contribution x_s as shown in equation 5.3.

$$P_{jM} = \sum_{s=1}^{N_s} x_s \cdot P_{js} \qquad , \psi_j(P_{jM}) = P_{jM} \qquad \psi_j(P_{js}) = P_{js} \tag{5.2}$$

$$\frac{1}{\rho_M} = \sum_{s=1}^{N_s} x_s \cdot \frac{1}{\rho_s} \qquad , \psi_j(P_{jM}) = \frac{1}{\rho_M} \qquad \psi_j(P_{js}) = \frac{1}{\rho_s} \tag{5.3}$$

Since the properties may have various functional forms and units, the operators are normalized into a dimensionless form by dividing by a reference operator. This reference is appropriately chosen such that the resulting dimensionless

properties are of the same order of magnitude. The normalized property operator is given as:

$$\Omega_{js} = \frac{\psi_j(P_{js})}{\psi_j(P_j^{ref})} \tag{5.4}$$

An Augmented Property index AUP for each stream s is defined as the summation of all the NP dimensionless property operators:

$$AUP_s = \sum_{j=1}^{NP} \Omega_{js} \tag{5.5}$$

The property cluster C_{js} for property j of stream s is defined as:

$$C_{js} = \frac{\Omega_{js}}{AUP_s} \tag{5.6}$$

Incorporating these clusters into the mass integration framework (El-Halwagi and Maniousiouthakis, 1989; El-Halwagi, 1997) enables the identification of optimal strategies for recovery and allocation of plant utilities. Process insights are obtained through visualization tools based on optimization concepts. Since the clusters are tailored to maintain the two fundamental rules for intra- and inter-stream conservation, lever-arm analysis may be employed extensively to identify recycle potentials (El-Halwagi and Spriggs, 1998; Parthasarathy and El-Halwagi, 2000). For visualization purposes the number of clusters is limited to three, however when using mathematical programming this limitation is relieved.

It should be noted that even though for visualization purposes the number of clusters is limited to three, this does not imply that the number of properties describing each stream is also limited to three. Assuming that the necessary number of properties describing the process streams is five, then a one-to-one mapping to property clusters would yield five clusters. If it is desired to visualize this problem, then it is necessary to reformulate the property operator descriptions in such a way that the total number of operators is three, i.e. property operators can be functions of several properties.

5.2.2 Conservation Rules

To ensure that the clusters are conserved and can be represented on a ternary diagram, it is imperative that they possess two fundamental properties, i.e. intra-stream conservation and inter-stream conservation. Intra-stream conservation requires that for each stream s the individual clusters must sum to unity

as described by equation 5.7, and a graphical representation of the conservation rule is presented in figure 5.1. Inter-stream conservation requires that mixing of two streams should be performed so that the resulting individual clusters are conserved. Therefore consistent additive rules, e.g. lever-arm rules as shown in equation 5.8, are needed to ensure that the mixture property cluster of two streams with different individual property clusters can be easily determined.

$$\sum_{j=1}^{N_C} C_{js} = 1 \tag{5.7}$$

$$C_{jM} = \sum_{s=1}^{N_s} \beta_s \cdot C_{js} \tag{5.8}$$

From a visualization standpoint, intra-stream conservation means that once two clusters are known, the third one is automatically determined. The validity of the intra-stream conservation rule given by equation 5.7 is easily verified by combining the cluster definition in equation 5.6 and the AUP definition in equation 5.5. Summation of the cluster values for all j properties in stream s yields:

$$\sum_{j=1}^{N_C} C_{js} = \frac{\sum_{j=1}^{N_C} \Omega_{js}}{AUP_s} = \frac{AUP_s}{AUP_s} = 1 \tag{5.9}$$

Although, mixing of the original properties may be based on nonlinear rules, the clusters are tailored to exhibit linear mixing rules as shown graphically by figure 5.2. When two sources S_1 and S_2 are mixed, the locus of all mixtures on the ternary cluster diagram is given by the straight line connecting sources S_1 and S_2. Depending on the fractional contributions of the streams, the resulting mixture splits the mixing line in ratios β_1 and β_2. In order to validate the inter-stream conservation rule given by equation 5.8, it is important to note that the cluster definition presented in equation 5.6 applies to any cluster, therefore the resulting cluster from the mixing process can be formulated as:

$$C_{jM} = \frac{\Omega_{jM}}{AUP_M} \tag{5.10}$$

By dividing the generalized mixing rule in equation 5.1 by the property operator reference values, equation 5.11 is obtained. Combining equation 5.11 with the dimensionless property operator expression given in equation 5.4 yields the

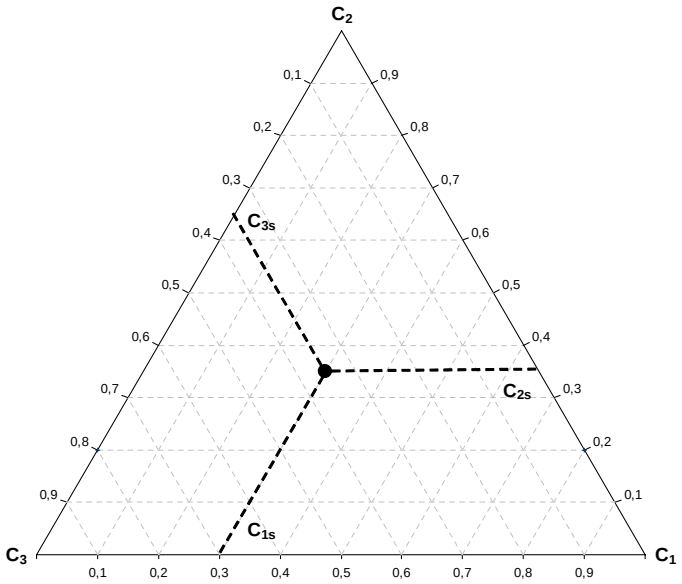

Figure 5.1: Representation of intra-stream conservation of clusters

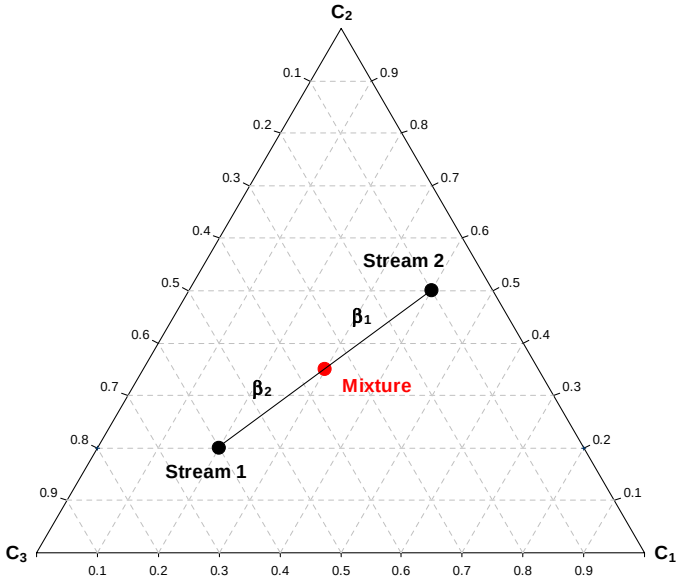

Figure 5.2: Representation of inter-stream conservation of clusters

mixing rule given by equation 5.12. Inserting this expression in equation 5.10, while rearranging the cluster definition given by equation 5.6, yields the proof for inter-stream conservation of the clusters given in equations 5.13 and 5.14.

$$\frac{\psi_j(P_{jM})}{\psi_j(P_j^{ref})} = \sum_{s=1}^{N_s} x_s \cdot \frac{\psi_j(P_{js})}{\psi_j(P_j^{ref})} \tag{5.11}$$

$$\Omega_{jM} = \sum_{s=1}^{N_s} x_s \cdot \Omega_{js} \tag{5.12}$$

$$C_{jM} = \frac{\sum_{s=1}^{N_s} x_s \cdot \Omega_{js}}{AUP_M} = \frac{\sum_{s=1}^{N_s} x_s \cdot AUP_s \cdot C_{js}}{AUP_M} = \sum_{s=1}^{N_s} \beta_s \cdot C_{js} \tag{5.13}$$

$$\beta_s = \frac{x_s \cdot AUP_s}{AUP_M} \tag{5.14}$$

Finally an expression for the AUgmented Property index of a mixture, AUP_M, must be derived. Combining equations 5.10 and 5.13 yields equation 5.15, and subsequent substitution of equation 5.7 provides the conservation rule for the relative cluster arms β_s as shown in equation 5.16.

$$\sum_{j=1}^{N_C} \sum_{s=1}^{N_s} \beta_s \cdot C_{js} = 1 \qquad \Leftrightarrow \qquad \sum_{s=1}^{N_s} \beta_s \sum_{j=1}^{N_C} C_{js} = 1 \tag{5.15}$$

$$\sum_{s=1}^{N_s} \beta_s = 1 \tag{5.16}$$

By combining equations 5.16 and 5.14, an expression for AUP_M is obtained:

$$\sum_{s=1}^{N_s} \frac{x_s \cdot AUP_s}{AUP_M} = 1 \qquad \Leftrightarrow \qquad AUP_M = \sum_{s=1}^{N_s} x_s \cdot AUP_s \tag{5.17}$$

The validated intra- and inter-stream conservation characteristics of the property clusters, enable the visual tracking of clusters and can provide unique insights into process and product design from the perspective of properties.

5.3 Visualization and Analysis

The clusters are tailored to maintain the two fundamental rules for intra- and inter-stream conservation. Incorporating these clusters into the mass integration framework enables the identification of optimal strategies for recovery and allocation of plant utilities. Process insights are obtained through visualization tools based on optimization concepts. The conversion of property data to cluster values is performed as outlined in table 5.1.

Step	Description	Equation
1	Calculate dimensionless stream property values	5.4
2	Calculate stream AUP indices	5.5
3	Calculate ternary cluster values for each stream	5.6
4	Plot the the points on the ternary cluster diagram	—

Table 5.1: Calculation of cluster values from property data

5.3.1 Ternary Source-Sink Mapping Diagram

The first visualization tool from the mass integration framework is the source-sink mapping diagram (El-Halwagi, 1997). This tool enables the identification of recycle and mixing potentials. Consider the case of two sources that do not meet the constraints of the process sink. Since the mixing of sources in cluster space follows lever-arm rules due to the derivation of the clusters, the mixing line will be a straight line between the two clusters. If the line passes through the sink region, direct mixing of the two sources results in a feasible feed to the sink. The source sink mapping for this case is given in figure 5.3.

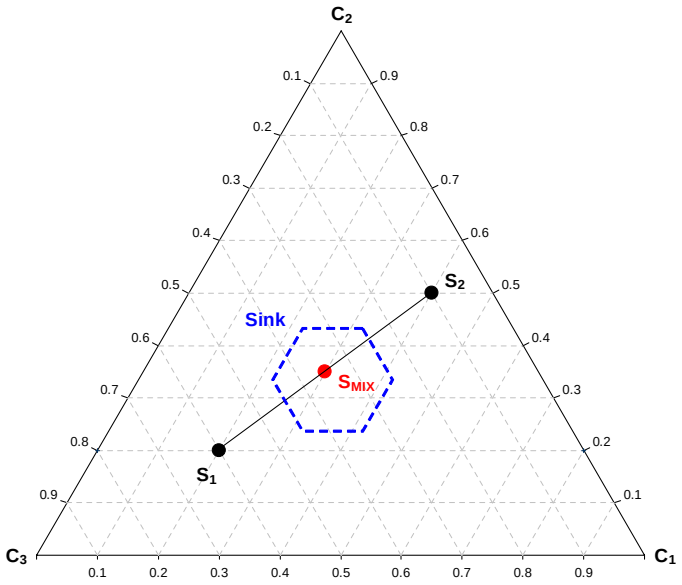

Figure 5.3: Source-sink mapping using clusters

The sink region in figure 5.3 is depicted as a hexagon. It should be emphasized that is is not always the case. The sink region is determined by the constraints on the feed conditions for the particular process unit and thus may have different geometric shapes. A systematic method for determining the sink region

is presented in section 5.4.

Rule 5.1 *Three conditions must be satisfied in order to ensure the feasibility of feeding sources or a mixture of sources into a sink:*

1. *The cluster value of the source (or mixture of sources) must be contained within the feasibility region of the sink on the cluster ternary diagram.*
2. *The values of the augmented property index (AUP) for the source (or mixture of sources) and the sink must match.*
3. *The flowrate of the source (or mixture of sources) must lie within the acceptable feed flowrate range for the sink.*

If one of these conditions (primarily the first two conditions) is not satisfied, interception devices may be used to adjust the values of the clusters or augmented property index for the sources. If the flowrate resulting from mixing streams S_1 and S_2 is higher than the maximum value accepted by the sink, a fraction of the mixture has to be removed before it is fed to the sink. The question is how much the flowrate should be reduced, since all flowrates between the lower and upper limits are feasible. However in some cases the optimum flowrate is easily calculated, e.g. if the process sink is a condensation system, the optimum flowrate corresponds to the minimum feasible flowrate, i.e. the minimum heat duty. If the process sink is e.g. a mixer or a reactor the optimum flowrate will be equal to the maximum feasible flowrate corresponding to the maximum recycle.

In some cases the flowrate of the mixture obtained by mixing streams S_1 and S_2 is below the minimum value accepted by the process sink. In such cases it is necessary to add an external source to obtain a total flowrate, which lies within the feasible region. However by doing so, not only does the total flowrate change, the property values of the mixture also change. Since the property constraints still need to be satisfied this puts a limitation on the feasible degree of mixing. In figure 5.4 a source-sink mapping diagram is presented for the case where mixing of the two sources satisfies the property constraints, but an external stream needs to be added to satisfy the flowrate constraints. The design problem depicted in figure 5.4 is for visualization purposes only, as no optimization is performed with respect to location of the mixing region or flowrate of the external stream. The solution presented is feasible but not necessarily the optimum.

5.3.2 Lever-arm Analysis

Once the available mixing potential for satisfying the sink constraints has been identified, optimization must be performed. A process is usually driven to obtain minimum cost or maximum profit. Consider two sources S_1 and S_2, that are mixed to satisfy the property constraints of a certain sink. According to equation 5.1, x_1 and x_2 denote the fractional contributions of the sources S_1

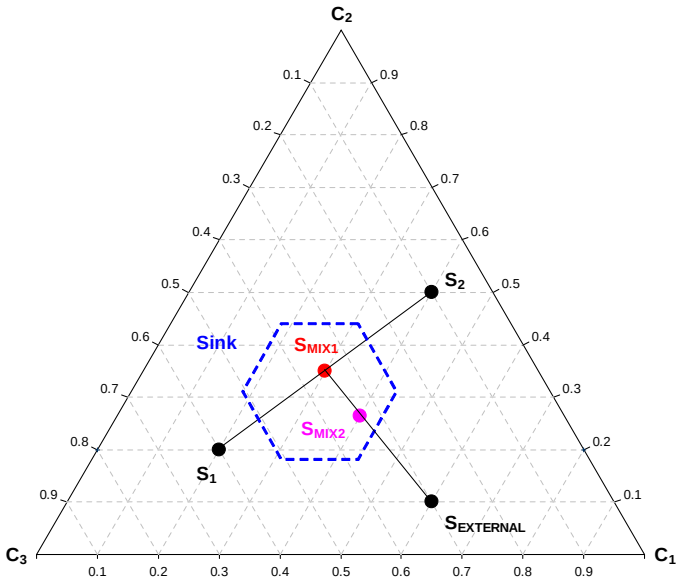

Figure 5.4: Property constraints are satisfied, but an external source must be added to satisfy flowrate constraints

and S_2 into the total flowrate of the mixture. Let source S_1 be more expensive than source S_2, i.e. $Cost_1 > Cost_2$. Thus, the cost of the mixture can be calculated as:

$$Cost_M = x_1 \cdot Cost_1 + x_2 \cdot Cost_2 \tag{5.18}$$

Utilizing that the fractional contributions x_s sum to unity, equation 5.18 can be rewritten as:

$$Cost_M = x_1 \cdot Cost_1 + (1 - x_1) \cdot Cost_2 \tag{5.19}$$

$$Cost_M = (Cost_1 - Cost_2) \cdot x_1 + Cost_2 \tag{5.20}$$

Noting that ($Cost_1$ - $Cost_2$) is a positive term, therefore, the cost of the mixture is linearly proportional to x_1. Hence, the lower the value of x_1, the lower the cost of the mixture and the following can now be deduced:

Rule 5.2 *When two sources, S_1 and S_2, are mixed to satisfy the property constraints of a sink with source S_1 being more expensive than S_2, minimizing the cost of the mixture is achieved by the minimum feasible value of x_1.*

Unfortunately, x_s cannot be directly visualized on the ternary cluster diagram. Instead, the lever arms visualized on figure 5.2, represent another quantity, β_s. While x_s represents the fractional contribution of the stream's flowrate to the total flow, β_s has a more subtle meaning. It is necessary to relate both since β_s is the visualization arm on the ternary diagram, but x_s is directly related to the cost of the mixture. The two terms are related through the augmented property index AUP, as described by equations 5.14 and 5.17. Combining equations 5.14 and 5.17 yields the following expression for the case of mixing two sources, S_1 and S_2:

$$\beta_1 = \frac{x_1 \cdot AUP_1}{x_1 \cdot AUP_1 + (1 - x_1) \cdot AUP_2} \tag{5.21}$$

Rearranging yields an expression for x_1 as a function of β_1:

$$x_1 = \frac{\beta_1 \cdot AUP_2}{\beta_1 \cdot (AUP_2 - AUP_1) + AUP_1} \tag{5.22}$$

Taking the first derivative of x_1 with respect to β_1 gives:

$$\frac{\partial x_1}{\partial \beta_1} = \frac{AUP_1 \cdot AUP_2}{(\beta_1 \cdot (AUP_2 - AUP_1) + AUP_1)^2} \tag{5.23}$$

With both AUP_1 and AUP_2 being non-negative, the right-hand side of equation 5.23 is also non-negative. Therefore, the flow fraction x_1 as a function of the relative cluster arm β_1 is monotonically increasing. From this the following rule can be stated:

Rule 5.3 *On a ternary cluster diagram, minimization of the cluster arm of a source corresponds to minimization of the flow contribution of that source. In other words, minimum β_s corresponds to minimum x_s.*

The two derived rules are important findings as they allow for developing cluster lever-arm minimization rules (visualized by β_s on the ternary cluster diagram) to correspond to minimum usage of the more expensive source and, consequently, the minimum cost of the mixture. The development of design and optimization rules based on lever-arm analysis is discussed in further detail in section 5.5.

5.3.3 Conversion from Ternary to Cartesian Coordinates

Although triangular diagrams are used in many applications such as, e.g. phase diagrams for crystallization systems and ternary mixtures, the number of commercially available software packages that support plotting and more importantly customizing such diagrams is limited. Fortunately there is a possibility for constructing a ternary diagram within a conventional rectangular diagram,

however in order to accomplish this, it is necessary to have a method for converting a set of ternary coordinates into the corresponding set of Cartesian coordinates.

In order to describe the conversion methodology figure 5.5 is used. All axes on the triangular diagram have a length of 1, thus for a given point in the ternary diagram, the X-value in Cartesian coordinates is directly found on the C_3 - C_1 axis. On this axis two values are known from the ternary coordinates, i.e. C_{1s} and $(1 - C_{3s})$. Since the triangular diagram has sides of equal length, the X-value in Cartesian coordinates will always be the mean of the two values C_{1s} and $(1 - C_{3s})$, as given by equation 5.24. As for the Y-value in Cartesian coordinates, it is obvious from figure 5.5 that it is related to the value of C_{2s}. Since the length of the axes are all equal to 1, the height of the diagram, i.e. from the C_3 - C_1 axis to the C_{2s} vertex is not equal to 1. Thus the ternary coordinate for C_{2s} has to be scaled by a constant to obtain the Cartesian Y-coordinate. This scaling factor can be obtained from the Pythagorian theorem as shown in equation 5.25 and the resulting expression for the Cartesian Y-value as a function of the ternary coordinates is given in equation 5.26.

$$X_{CC,s} = \frac{C_{1s} + (1 - C_{3s})}{2} = C_{1s} + 0.5 \cdot C_{2s} \tag{5.24}$$

$$\left(\frac{1}{2}\right)^2 + Y_{Scaling}^2 = (1)^2 \qquad \Leftrightarrow \qquad Y_{Scaling} = \frac{\sqrt{3}}{2} \tag{5.25}$$

$$Y_{CC,s} = \frac{\sqrt{3}}{2} \cdot C_{2s} \tag{5.26}$$

The coordinate transformations given in equations 5.24 and 5.26 can also be described in terms of dimensionless property operators:

$$X_{CC,s} = \frac{\Omega_{1s} + 0.5 \cdot \Omega_{2s}}{\Omega_{1s} + \Omega_{2s} + \Omega_{3s}} \tag{5.27}$$

$$Y_{CC,s} = \frac{\Omega_{2s}\sqrt{3}}{2 \cdot (\Omega_{1s} + \Omega_{2s} + \Omega_{3s})} \tag{5.28}$$

Furthermore, the coordinate transformations can also be used for calculating cluster lever-arm values from the diagram directly. As an example, the inter-stream conservation illustration in figure 5.2 is used. By utilizing the Pythagorian theorem, the relative arms can be calculated according to equations 5.29 and 5.30.

$$\beta_1 = \sqrt{\frac{(X_{CC,2} - X_{CC,mix})^2 + (Y_{CC,2} - Y_{CC,mix})^2}{(X_{CC,2} - X_{CC,1})^2 + (Y_{CC,2} - Y_{CC,1})^2}} \tag{5.29}$$

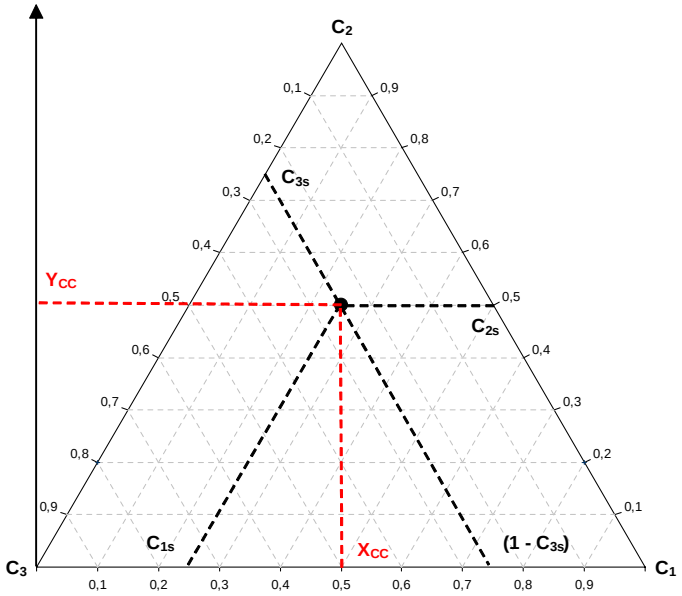

Figure 5.5: Illustration of conversion from ternary to Cartesian coordinates

$$\beta_2 = \sqrt{\frac{(X_{CC,1} - X_{CC,mix})^2 + (Y_{CC,1} - Y_{CC,mix})^2}{(X_{CC,2} - X_{CC,1})^2 + (Y_{CC,2} - Y_{CC,1})^2}} \tag{5.30}$$

Equations 5.29 and 5.30 can also be described in terms of the ternary coordinates by substitution of equations 5.24 and 5.26:

$$\beta_1 = \sqrt{\frac{(C_{12} - C_{1mix} + C_{3mix} - C_{32})^2 + 3 \cdot (C_{22} - C_{2mix})^2}{(C_{12} - C_{11} + C_{31} - C_{32})^2 + 3 \cdot (C_{22} - C_{21})^2}} \tag{5.31}$$

$$\beta_2 = \sqrt{\frac{(C_{11} - C_{1mix} + C_{3mix} - C_{31})^2 + 3 \cdot (C_{21} - C_{2mix})^2}{(C_{12} - C_{11} + C_{31} - C_{32})^2 + 3 \cdot (C_{22} - C_{21})^2}} \tag{5.32}$$

Analogous to the Cartesian coordinates, the cluster lever-arms could be described in terms of dimensionless property operators by subsequent substitution of equations 5.27 and 5.28.

5.4 Feasibility Region Boundaries

Obtaining an exact mapping of the sink region from the property domain to the cluster domain requires the conversion of an infinite number of feasible points. Therefore, it is highly desirable to identify the exact shape of the feasibility region without extensive enumeration. Achieving an accurate representation of the feasibility region, gives rise to two questions:

- What is the minimum number of boundary lines defining the feasibility region?
- What are the mathematical expressions for the vertices and lines constituting the feasibility region?

The following mathematical analysis answers these questions and establishes the exact expressions for the feasibility region a priori and without enumeration (El-Halwagi *et al.*, 2004). Consider a sink with three targeted properties, for which the property constraints for a feasible feed s, are given by equation 5.33. Assuming that the property operators describing the mixing rules for each of the targeted properties are monotonically increasing as a function of the raw property P_{js}, the expressions in equations 5.34 and 5.35 are obtained.

$$P_{j,\text{sink}}^{\min} \leq P_{js} \leq P_{j,\text{sink}}^{\max} \qquad j = 1, 2, 3 \tag{5.33}$$

$$\Omega_{j,\text{sink}}^{\min} = \frac{\psi_j(P_{j,\text{sink}}^{\min})}{\psi_j(P_j^{ref})} \tag{5.34}$$

$$\Omega_{j,\text{sink}}^{\max} = \frac{\psi_j(P_{j,\text{sink}}^{\max})}{\psi_j(P_j^{ref})} \tag{5.35}$$

According to the cluster definition given in equation 5.6 and the definition of the augmented property index in equation 5.5, the expressions for the three cluster boundary values for the sink constraints can be written as equations 5.36-5.41.

$$C_{1,\text{sink}}^{\min} = \frac{\Omega_{1,\text{sink}}^{\min}}{\Omega_{1,\text{sink}}^{\min} + \Omega_{2,\text{sink}}^{\max} + \Omega_{3,\text{sink}}^{\max}} \tag{5.36}$$

$$C_{1,\text{sink}}^{\max} = \frac{\Omega_{1,\text{sink}}^{\max}}{\Omega_{1,\text{sink}}^{\max} + \Omega_{2,\text{sink}}^{\min} + \Omega_{3,\text{sink}}^{\min}} \tag{5.37}$$

$$C_{2,\text{sink}}^{\min} = \frac{\Omega_{2,\text{sink}}^{\min}}{\Omega_{1,\text{sink}}^{\max} + \Omega_{2,\text{sink}}^{\min} + \Omega_{3,\text{sink}}^{\max}} \tag{5.38}$$

$$C_{2,\text{sink}}^{\max} = \frac{\Omega_{2,\text{sink}}^{\max}}{\Omega_{1,\text{sink}}^{\min} + \Omega_{2,\text{sink}}^{\max} + \Omega_{3,\text{sink}}^{\min}} \tag{5.39}$$

$$C_{3,\text{sink}}^{\min} = \frac{\Omega_{3,\text{sink}}^{\min}}{\Omega_{1,\text{sink}}^{\max} + \Omega_{2,\text{sink}}^{\max} + \Omega_{3,\text{sink}}^{\min}} \tag{5.40}$$

$$C_{3,\text{sink}}^{\max} = \frac{\Omega_{3,\text{sink}}^{\max}}{\Omega_{1,\text{sink}}^{\min} + \Omega_{2,\text{sink}}^{\min} + \Omega_{3,\text{sink}}^{\max}} \tag{5.41}$$

The strategy for identification of the feasibility region will start by determining an overestimation followed by an underestimation of the feasibility region, and then identifying the true feasibility region in between the two estimations. The overestimation is illustrated in figure 5.6, while the underestimation is presented in figure 5.7. The overestimation of the feasibility region is determined by simply bounding the region within the minimum and maximum values of the clusters, thereby describing the overestimation by six line segments. It is not given that this overestimation corresponds to the true feasibility region, but it is guarenteed that no feasible points lie outside the overestimation. Furthermore, equations 5.36-5.41 provide a point on each of the line segments of the overestimation. Since all these points are part of the true feasibility region it is given that any mixtures of those points also lie within the true feasibility region. Therefore an underestimation of the true region can be obtained by connecting these six points.

Consider the two points Q_i and Q_j characterized by the values $Q_i(\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\min})$ and $Q_j(\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\max})$ respectively. According to equations 5.27 and 5.28, the Cartesian coordinates for Q_i and Q_j are given as:

$$\left(X_{CC}^{Q_i}, Y_{CC}^{Q_i}\right) = \left(\frac{\Omega_1^{\min} + 0.5 \cdot \Omega_2^{\max}}{\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\min}}, \frac{\Omega_2^{\max} \cdot \sqrt{3}}{2 \cdot \left(\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\min}\right)}\right) \tag{5.42}$$

$$\left(X_{CC}^{Q_j}, Y_{CC}^{Q_j}\right) = \left(\frac{\Omega_1^{\min} + 0.5 \cdot \Omega_2^{\max}}{\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\max}}, \frac{\Omega_2^{\max} \cdot \sqrt{3}}{2 \cdot \left(\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\max}\right)}\right) \tag{5.43}$$

In Cartesian coordinates, the slope of the line connecting two points Q_i and Q_j can be calculated as shown in equation 5.44 and by inserting the Cartesian coordinates given in equations 5.42 and 5.43, the slope of the line can be calculated as shown in equation 5.45.

$$Slope = \frac{Y_{CC}^{Q_i} - Y_{CC}^{Q_j}}{X_{CC}^{Q_i} - X_{CC}^{Q_j}} \tag{5.44}$$

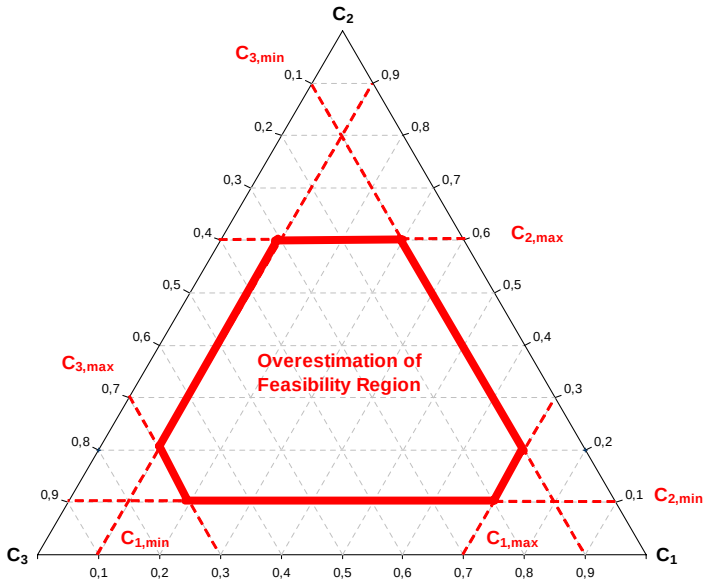

Figure 5.6: Overestimation of the feasibility region of a sink

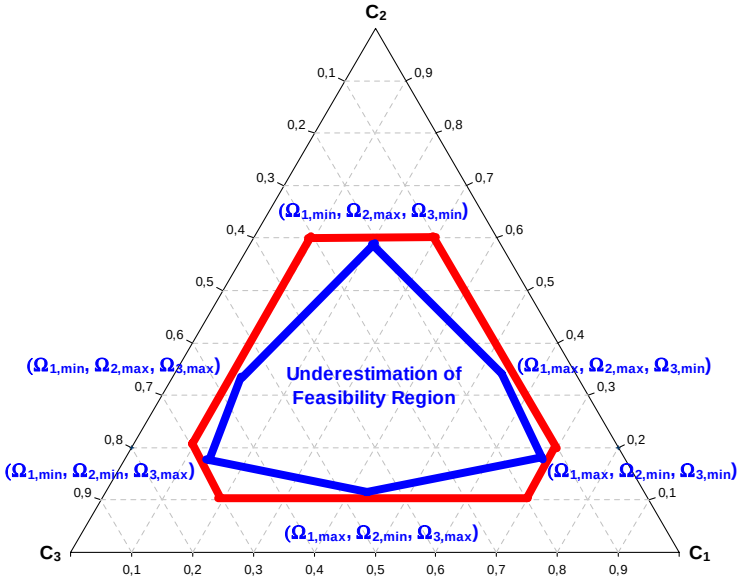

Figure 5.7: Underestimation of the feasibility region of a sink

$$Slope = \left(\frac{\frac{\Omega_2^{\max} \cdot \sqrt{3}}{2 \cdot \left(\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\min}\right)} - \frac{\Omega_2^{\max} \cdot \sqrt{3}}{2 \cdot \left(\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\max}\right)}}{\frac{\Omega_1^{\min} + 0.5 \cdot \Omega_2^{\max}}{\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\min}} - \frac{\Omega_1^{\min} + 0.5 \cdot \Omega_2^{\max}}{\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\max}}} \right) \tag{5.45}$$

By rearranging equation 5.45, the simplified expression for the slope of the line given by equation 5.46 is obtained. The intercept of the line connecting the two points is obtained using the slope and one point on the line as described by equation 5.47.

$$Slope = \frac{\sqrt{3}}{1 + 2 \cdot \frac{\Omega_1^{\min}}{\Omega_2^{\max}}} \tag{5.46}$$

$$Y_{CC}^{Q_i} = Slope \cdot X_{CC}^{Q_i} + Int \tag{5.47}$$

By inserting the Cartesian coordinates given in equation 5.42 along with the expression for the slope calculated in equation 5.46, the intercept of the line can be calculated as shown in equation 5.48.

$$Int = \frac{\Omega_2^{\max} \cdot \frac{\sqrt{3}}{2}}{\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\min}} - \frac{\sqrt{3}}{1 + 2\frac{\Omega_1^{\min}}{\Omega_2^{\max}}} \cdot \frac{\Omega_1^{\min} + 0.5 \cdot \Omega_2^{\max}}{\Omega_1^{\min} + \Omega_2^{\max} + \Omega_3^{\min}} = 0 \tag{5.48}$$

Since the origin coordinates are located at ($C_1 = 0$, $C_2 = 0$, $C_3 = 1$), which correspond to ($X_{CC} = 0$, $Y_{CC} = 0$), then the zero intercept indicates that this straight line connecting the two feasible points passes through the origin. The same result can be obtained by deriving the slope equation for any line emanating from a vertex point, whose Cartesian coordinates are designated as ($\tilde{X}_v$,$\tilde{Y}_v$), with the origin located at ($C_1 = 0$, $C_2 = 0$, $C_3 = 1$). The slope of any line emanating from the v^{th} vertex is described by:

$$Slope_{(\tilde{X}_v, \tilde{Y}_v)} = \frac{Y_{CC} - \tilde{Y}_v}{X_{CC} - \tilde{Y}_v} \tag{5.49}$$

Substituting equations 5.24 and 5.26 provides an expression in terms of ternary cluster values as given in equation 5.50, while substitution of equations 5.27 and 5.28 provides the corresponding expression in terms of dimensionless property operators given by equation 5.51.

$$Slope_{(\tilde{X}_v, \tilde{Y}_v)} = \frac{\frac{\sqrt{3}}{2} C_{2s} - \tilde{Y}_v}{C_{1s} + 0.5 C_{2s} - \tilde{X}_v} \tag{5.50}$$

$$Slope_{(\tilde{X}_v, \tilde{Y}_v)} = \frac{\frac{\sqrt{3}}{2} \Omega_{2s} - \tilde{Y}_v \left(\Omega_{1s} + \Omega_{2s} + \Omega_{3s}\right)}{\Omega_{1s} + 0.5 \Omega_{2s} - \tilde{X}_v \left(\Omega_{1s} + \Omega_{2s} + \Omega_{3s}\right)} \tag{5.51}$$

Equation 5.51 can be rearranged to obtain equation 5.52 using the coefficients given by equations 5.53-5.55.

$$\lambda_1 \cdot \Omega_{1s} + \lambda_2 \cdot \Omega_{2s} + \lambda_3 \cdot \Omega_{3s} = 0 \tag{5.52}$$

$$\lambda_1 = (1 - \tilde{X}_v) \cdot Slope_{(\tilde{X}_v, \tilde{Y}_v)} + \tilde{Y}_v \tag{5.53}$$

$$\lambda_2 = (0.5 - \tilde{X}_v) \cdot Slope_{(\tilde{X}_v, \tilde{Y}_v)} + \frac{\sqrt{3}}{2} + \tilde{Y}_v \tag{5.54}$$

$$\lambda_3 = \tilde{Y}_v - \tilde{X}_v \cdot Slope_{(\tilde{X}_v, \tilde{Y}_v)} \tag{5.55}$$

The slope of any line emanating from vertex coordinates ($\tilde{X}_v$,$\tilde{Y}_v$)=(0,0) can be calculated from equation 5.51:

$$Slope_{(0,0)} = \frac{\frac{\sqrt{3}}{2} \cdot \Omega_{2s}}{\Omega_{1s} + 0.5 \cdot \Omega_{2s}} = \frac{\sqrt{3} \cdot \Omega_{2s}}{2 \cdot \Omega_{1s} + \Omega_{2s}} = \frac{\sqrt{3}}{1 + 2 \cdot \frac{\Omega_{1s}}{\Omega_{2s}}} \tag{5.56}$$

Evaluation of the minimum and maximum slopes of lines emanating from vertex coordinates ($\tilde{X}_v$,$\tilde{Y}_v$)=(0,0) is achieved using minimum and maximum values of the dimensionless property operator values as shown in equations 5.57 and 5.58. The maximum slope is exactly the slope of the line connecting $Q_i(\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\min})$ and $Q_j(\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\max})$ as presented in equation 5.46. This result means that no feasible points exist between this line and the overestimator, thus making this line part of the boundary of the true feasibility region.

$$Slope_{(0,0)}^{\min} = \frac{\sqrt{3}}{1 + 2 \cdot \frac{\Omega_1^{\max}}{\Omega_2^{\min}}} \tag{5.57}$$

$$Slope_{(0,0)}^{\max} = \frac{\sqrt{3}}{1 + 2 \cdot \frac{\Omega_1^{\min}}{\Omega_2^{\max}}} \tag{5.58}$$

It should be noted, that for the two points $Q_i(\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\min})$ and $Q_j(\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\max})$, two values are the same, i.e. $\Omega_1^{\min}$ and $\Omega_2^{\max}$, thus the slope of the line connecting the two points is independent of the value of Ω_{2s}. Therefore the same result as given by equation 5.56 is obtained by setting $\lambda_2 = 0$ and solving equation 5.52 for the slope.

This procedure can be repeated for all the lines connecting the six points lying on the intersection of the overestimating and underestimating regions. The mathematical expressions are shown in table 5.2 and the graphical results are shown in figure 5.8. It should be noted that, any point on the boundary of

the true feasibility region can be uniquely mapped to the property domain. The reason for this exception is that each of the six vertices of the feasibility region is unique, thus any lines (representing mixtures) connecting these vertices will also have a unique mapping. The findings of the foregoing analysis can be summarized by the following important results:

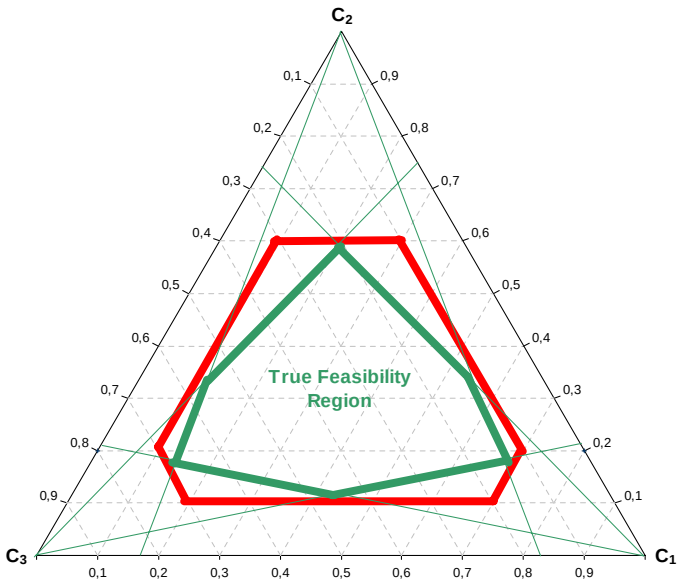

Figure 5.8: True feasibility region of a sink

Rule 5.4 *The boundary of the true feasibility region can be accurately represented by no more than six linear segments.*

Rule 5.5 *When extended, the linear segments of the boundary of the true feasibility region constitute three convex hulls (cones) with their heads lying on the three vertices of the ternary cluster diagram.*

Rule 5.6 *The six points defining the boundary of the true feasibility region are determined apriori and are characterized by the following values of dimensionless operators:*

$$(\Omega_1^{\min}, \Omega_2^{\min}, \Omega_3^{\max}) \quad (\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\max}) \quad (\Omega_1^{\min}, \Omega_2^{\max}, \Omega_3^{\min})$$
$$(\Omega_1^{\max}, \Omega_2^{\max}, \Omega_3^{\min}) \quad (\Omega_1^{\max}, \Omega_2^{\min}, \Omega_3^{\min}) \quad (\Omega_1^{\max}, \Omega_2^{\min}, \Omega_3^{\max})$$

Now that the boundary of the true feasibility region has been uniquely identified apriori and without requiring extensive enumeration, visual optimization strategies based on lever-arm analysis can be developed.

Vertex (Ternary) (C_1s,C_2s,C_3s)	Vertex (Cartesian) $(\tilde{X}_v,\tilde{Y}_v)$	Min. Slope $f_v^{\min}(\Omega_{js})$	Points on Region Ω_{js}	Max. Slope $f_v^{\max}(\Omega_{js})$	Points on Region Ω_{js}
(0,0,1)	(0,0)	$\frac{\sqrt{3}}{1+2\cdot\frac{\Omega_1^{\max}}{\Omega_2^{\min}}}$	$\left[\begin{array}{l}(\Omega_1^{\min},\Omega_2^{\max},\Omega_3^{\min})\\(\Omega_1^{\min},\Omega_2^{\max},\Omega_3^{\max})\end{array}\right]$	$\frac{\sqrt{3}}{1+2\cdot\frac{\Omega_1^{\min}}{\Omega_2^{\max}}}$	$\left[\begin{array}{l}(\Omega_1^{\max},\Omega_2^{\min},\Omega_3^{\min})\\(\Omega_1^{\max},\Omega_2^{\min},\Omega_3^{\max})\end{array}\right]$
(0,1,0)	$(0.5,\frac{\sqrt{3}}{2})$	$\frac{\sqrt{3}\left(1+\frac{\Omega_1^{\min}}{\Omega_3^{\max}}\right)}{1-\frac{\Omega_1^{\min}}{\Omega_3^{\max}}}$	$\left[\begin{array}{l}(\Omega_1^{\max},\Omega_2^{\min},\Omega_3^{\min})\\(\Omega_1^{\max},\Omega_2^{\max},\Omega_3^{\min})\end{array}\right]$	$\frac{\sqrt{3}\left(1+\frac{\Omega_1^{\max}}{\Omega_3^{\min}}\right)}{1-\frac{\Omega_1^{\max}}{\Omega_3^{\min}}}$	$\left[\begin{array}{l}(\Omega_1^{\min},\Omega_2^{\min},\Omega_3^{\max})\\(\Omega_1^{\min},\Omega_2^{\max},\Omega_3^{\max})\end{array}\right]$
(1,0,0)	(1,0)	$\frac{-\sqrt{3}}{1+2\cdot\frac{\Omega_3^{\max}}{\Omega_2^{\min}}}$	$\left[\begin{array}{l}(\Omega_1^{\min},\Omega_2^{\max},\Omega_3^{\min})\\(\Omega_1^{\max},\Omega_2^{\max},\Omega_3^{\min})\end{array}\right]$	$\frac{-\sqrt{3}}{1+2\cdot\frac{\Omega_3^{\min}}{\Omega_2^{\max}}}$	$\left[\begin{array}{l}(\Omega_1^{\min},\Omega_2^{\min},\Omega_3^{\max})\\(\Omega_1^{\max},\Omega_2^{\min},\Omega_3^{\max})\end{array}\right]$

Table 5.2: Coordinates and slopes of the boundary of the true feasibility region

5.5 Design and Optimization Rules

Based on the visualization tools and lever-arm rules presented in section 5.3, it is now possible to derive rules for visual design and optimization within the ternary cluster diagram. Consider the case, where two external sources are available for use along with one internal process resource, which can be used to reduce the consumption of the more expensive external sources. Examples of external sources include raw materials, virgin fibers, fresh solvent, etc., while internal process streams denote unreacted raw materials, waste streams, spent solvents, etc. The cost of the internal source is negligible, as it is an available process resource. In figure 5.9 a generic source-sink mapping diagram is presented for two competing external sources. A number of different scenarios may be envisioned, e.g. only one external source may be used, the externals may be mixed, etc. In the following four such cases are investigated:

1. Only one external source is used and the total flowrate is constant.
2. Either external source may be used, but no mixing and the total flowrate is constant.
3. Mixing of the two external sources is allowed, but the total flowrate is constant.
4. Mixing of external sources is allowed and the total flowrate is no longer constant.

For each case the design problem is described as a series of equations, which may be solved using mathematical optimization. The equations are derived in a generic form using lever-arm analysis of the process sink constraints. It is assumed that the cost of the external sources is known and that there is a constraint on the available flowrate of the internal process resource. Finally the objective function in all optimization problems derived is formulated to minimize the total cost $Cost_{Total}$ for use of external sources. It is assumed that the cost of external sources is constant but dependent on flowrate.

It should be emphasized that for all cases presented here, the identified solutions have to be validated by checking that the AUP conditions of rule 5.1 are satisfied.

5.5.1 Case No. 1 - Using Only One External Source and Total Flowrate Constant

The derivations presented here are based on the use of External 1, however analogous results can be derived for External 2. Lever-arm analysis of the sink constraints provides the constraints for the internal and external sources:

$$\frac{\overline{a_1 i}}{\overline{e_1 i}} \cdot F_{Total} \leq F_1 \leq \frac{\overline{b_1 i}}{\overline{e_1 i}} \cdot F_{Total} \tag{5.59}$$

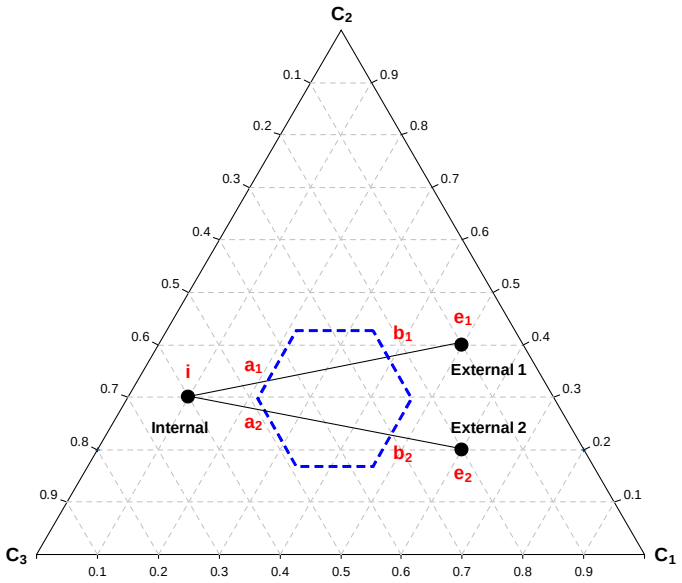

Figure 5.9: Allocation of internal and competing external sources (no mixing of externals allowed)

$$\frac{\overline{b_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \leq F_{Internal} \leq \frac{\overline{a_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \tag{5.60}$$

Assuming that there is a limit on the available flowrate of the internal process resource yields equation 5.61, while the total mass balance yields equation 5.62:

$$F_{Internal} \leq F_{Internal}^{Available} \tag{5.61}$$

$$F_{Total} = F_1 + F_{Internal} \tag{5.62}$$

Assuming that the total flowrate is constant, and that the flowrate of the internal source is also constant:

$$F_{Total} = \text{Constant} \tag{5.63}$$

$$F_{Internal} = \text{Constant} \tag{5.64}$$

Equations 5.59-5.64 constitute the system of equations that describe the constraints that the objective function is subject to. For negligible cost of the

internal source, the total cost may be calculated as given in equation 5.65 (assuming constant, but flowrate dependent, cost of the external source).

$$Cost_{Total} = F_1 \cdot Cost_1 + F_{Internal} \cdot Cost_{Internal} = F_1 \cdot Cost_1 \tag{5.65}$$

Equation 5.65 shows that minimum total cost is obtained by minimizing the flowrate of the external source. This is not a surprising result and the minimum value of the external source flowrate is given in equation 5.59, however the constraints on the internal resource flowrate given by equations 5.60 and 5.61, must also be satisfied, thus resulting in:

$$F_{Internal}^{Available} \geq \frac{\overline{a_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \qquad \Rightarrow \begin{cases} \min(F_1) = \frac{\overline{a_1 i}}{\overline{e_1 i}} \cdot F_{Total} \\ \max(F_{Internal}) = \frac{\overline{a_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \end{cases} \tag{5.66}$$

$$F_{Internal}^{Available} < \frac{\overline{a_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \qquad \Rightarrow \begin{cases} \min(F_1) = F_{Total} - F_{Internal}^{Available} \\ \max(F_{Internal}) = F_{Internal}^{Available} \end{cases} \tag{5.67}$$

It should be noted that the results presented in equations 5.66 and 5.67 have been obtained by using visualization tools only. Subsequent mathematical optimization will yield the same results as the optimum values are given explicitly by the constraint and balance equations.

5.5.2 Case No. 2 - Either External Source, No Mixing of External Sources and Total Flowrate Constant

The derivations presented here are analogous to those in section 5.5.1. Lever-arm analysis of the sink constraints provides the flowrate constraints for the internal and external sources, depending on which external source is used:

$$F_1 = 0 : \begin{cases} \frac{\overline{a_2 i}}{\overline{e_2 i}} \cdot F_{Total} \leq F_2 \leq \frac{\overline{b_2 i}}{\overline{e_2 i}} \cdot F_{Total} \\ \frac{\overline{b_2 e_2}}{\overline{e_2 i}} \cdot F_{Total} \leq F_{Internal} \leq \frac{\overline{a_2 e_2}}{\overline{e_2 i}} \cdot F_{Total} \end{cases} \tag{5.68}$$

$$F_2 = 0 : \begin{cases} \frac{\overline{a_1 i}}{\overline{e_1 i}} \cdot F_{Total} \leq F_1 \leq \frac{\overline{b_1 i}}{\overline{e_1 i}} \cdot F_{Total} \\ \frac{\overline{b_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \leq F_{Internal} \leq \frac{\overline{a_1 e_1}}{\overline{e_1 i}} \cdot F_{Total} \end{cases} \tag{5.69}$$

Assuming that there is a limit on the available flowrate of the internal process

resource yields equation 5.70, while the overall mass balance yields equation 5.71.

$$F_{Internal} \leq F_{Internal}^{Available} \tag{5.70}$$

$$F_{Total} = F_1 + F_2 + F_{Internal} \tag{5.71}$$

Constant flowrates yield:

$$F_{Total} = \text{Constant} \tag{5.72}$$

$$F_{Internal} = \text{Constant} \tag{5.73}$$

For negligible cost of the internal source, the total cost may be written as (assuming constant but different cost of the external sources):

$$Cost_{Total} = F_1 \cdot Cost_1 + F_2 \cdot Cost_2 \tag{5.74}$$

Thus the minimization of the total cost will depend on the cost difference between the two external sources and the relative usage of them. The relative cost, α_{Cost} of the two external sources is defined as follows:

$$\alpha_{Cost} = \frac{Cost_1}{Cost_2} \qquad , \alpha_{Cost} \neq 1 \tag{5.75}$$

Furthermore, introducing the ratio γ_{Flow} of the minimum feasible flowrates of the two external sources yields:

$$\gamma_{Flow} = \frac{F_1^{Minimum}}{F_2^{Minimum}} = \frac{\frac{\overline{a_1 i}}{\overline{e_1 i}} \cdot F_{Total}}{\frac{\overline{a_2 i}}{\overline{e_2 i}} \cdot F_{Total}} = \frac{\overline{a_1 i} \cdot \overline{e_2 i}}{\overline{e_1 i} \cdot \overline{a_2 i}} \tag{5.76}$$

Thus the minimum total cost of using either external source may be evaluated:

$$\alpha_{Cost} \cdot \gamma_{Flow} = \left(\frac{Cost_1}{Cost_2}\right) \cdot \left(\frac{F_1^{Minimum}}{F_2^{Minimum}}\right) \tag{5.77}$$

Equation 5.77 shows that for $\alpha_{Cost} \cdot \gamma_{Flow} < 1$ the total cost of using external 1 is less than for using external 2, thus minimizing the total cost requires minimizing the flowrate of external 1 and not using external 2. Similarly for $\alpha_{Cost} \cdot \gamma_{Flow} > 1$, minimizing the total cost involves minimizing the flowrate of external 2 and not using external 1.

Both scenarios must still satisfy the constraints on the internal resource flowrate given by equations 5.70 and 5.71. For $\alpha_{Cost} \cdot \gamma_{Flow} < 1$, the solution yields the same result as case 1, i.e. equations 5.66 and 5.67, which was expected since this result implies the use of only external 1. An analogous result is obtained for $\alpha_{Cost} \cdot \gamma_{Flow} > 1$ and given in equations 5.78 and 5.79.

$$F_{Internal}^{Available} \geq \frac{\overline{a_2 e_2}}{\overline{e_2 i}} \cdot F_{Total} \quad \Rightarrow \begin{cases} \min(F_2) = \frac{\overline{a_2 i}}{\overline{e_2 i}} \cdot F_{Total} \\ \max(F_{Internal}) = \frac{\overline{a_2 e_2}}{\overline{e_2 i}} \cdot F_{Total} \end{cases} \tag{5.78}$$

$$F_{Internal}^{Available} < \frac{\overline{a_2 e_2}}{\overline{e_2 i}} \cdot F_{Total} \quad \Rightarrow \begin{cases} \min(F_2) = F_{Total} - F_{Internal}^{Available} \\ \max(F_{Internal}) = F_{Internal}^{Available} \end{cases} \tag{5.79}$$

5.5.3 Case No. 3 - Allowing Mixing of External Sources and Total Flowrate Constant

The generic source-sink mapping presented in figure 5.9 is no longer applicable as it does not allow the possibility of mixing the two external sources. For this case the source-sink diagram given in figure 5.10 will be used instead, yielding the flowrate constraints for the internal and external sources:

$$\frac{\overline{e_{mix} e_2}}{\overline{e_1 e_2}} \cdot \frac{\overline{a_{mix} i}}{\overline{e_{mix} i}} \cdot F_{Total} \leq F_1 \leq \frac{\overline{b_1 i}}{\overline{e_1 i}} \cdot F_{Total} \tag{5.80}$$

$$\frac{\overline{e_1 e_{mix}}}{\overline{e_1 e_2}} \cdot \frac{\overline{a_{mix} i}}{\overline{e_{mix} i}} \cdot F_{Total} \leq F_2 \leq \frac{\overline{b_2 i}}{\overline{e_2 i}} \cdot F_{Total} \tag{5.81}$$

$$\frac{\overline{b_{mix} e_{mix}}}{\overline{e_{mix} i}} \cdot F_{Total} \leq F_{Internal} \leq \frac{\overline{a_{mix} e_{mix}}}{\overline{e_{mix} i}} \cdot F_{Total} \tag{5.82}$$

Assuming that there is a limit on the available flowrate of the internal process resource yields equation 5.70, while the assumption of constant total flow yields the balance equation 5.71 and the constant flowrates yield equations 5.73 and 5.72. For negligible cost of the internal source, the total cost may be written as equation 5.74 (still assuming constant but different cost of the external sources).

Equation 5.77 shows that for $\alpha_{Cost} \cdot \gamma_{Flow} < 1$ the total cost of using external 1 is less than for using external 2, thus minimizing the total cost requires maximizing the flowrate of the internal source and minimizing the flowrate of external 2. Similarly for $\alpha_{Cost} \cdot \gamma_{Flow} > 1$, minimizing the total cost involves maximizing the flowrate of the internal source and minimizing the flowrate of

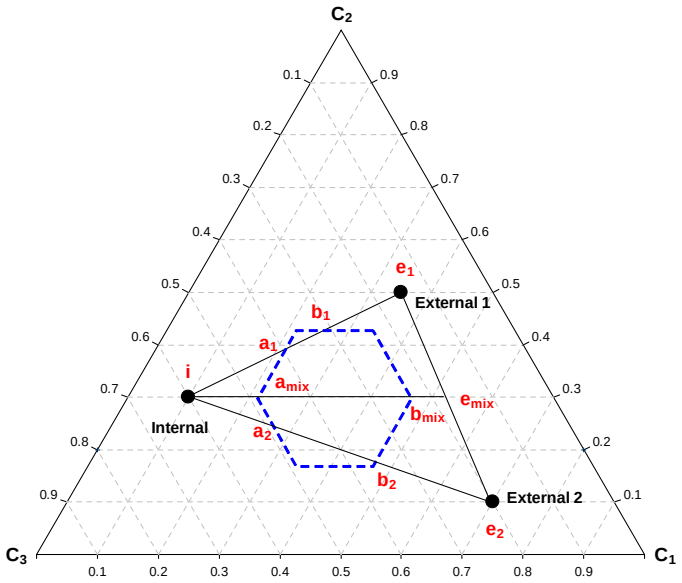

Figure 5.10: Allocation of internal and competing external sources (mixing of externals allowed)

external 1. Both scenarios must still satisfy the constraints on the internal resource flowrate given by equations 5.70 and 5.71.

$\alpha_{Cost} \cdot \gamma_{Flow} < 1$, $F_{Internal}^{Available} \geq \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} \cdot F_{Total}$:

$$\min(Cost_{Total}) \Rightarrow \begin{cases} \max(F_{Internal}) = \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} \cdot F_{Total} \\ \min(F_2) = \frac{\overline{e_1e_{mix}}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}} \cdot F_{Total} \\ F_1 = F_{Total} \cdot \left(1 - \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} - \frac{\overline{e_1e_{mix}}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}}\right) \end{cases} \tag{5.83}$$

$\alpha_{Cost} \cdot \gamma_{Flow} < 1$, $F_{Internal}^{Available} < \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} \cdot F_{Total}$:

$$\min(Cost_{Total}) \Rightarrow \begin{cases} \max(F_{Internal}) = F_{Internal}^{Available} \\ \min(F_2) = \frac{\overline{e_1e_{mix}}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}} \cdot F_{Total} \\ F_1 = F_{Total} \cdot \left(1 - \frac{\overline{e_1e_{mix}}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}}\right) - F_{Internal}^{Available} \end{cases} \tag{5.84}$$

$\alpha_{Cost} \cdot \gamma_{Flow} > 1$, $F_{Internal}^{Available} \geq \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} \cdot F_{Total}$:

$$\min(Cost_{Total}) \Rightarrow \begin{cases} \max(F_{Internal}) = \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} \cdot F_{Total} \\ \min(F_1) = \frac{\overline{e_{mix}e_2}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}} \cdot F_{Total} \\ F_2 = F_{Total} \cdot \left(1 - \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} - \frac{\overline{e_{mix}e_2}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}}\right) \end{cases} \tag{5.85}$$

$\alpha_{Cost} \cdot \gamma_{Flow} > 1$, $F_{Internal}^{Available} < \frac{\overline{a_{mix}e_{mix}}}{\overline{e_{mix}i}} \cdot F_{Total}$:

$$\min(Cost_{Total}) \Rightarrow \begin{cases} \max(F_{Internal}) = F_{Internal}^{Available} \\ \min(F_1) = \frac{\overline{e_{mix}e_2}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}} \cdot F_{Total} \\ F_2 = F_{Total} \cdot \left(1 - \frac{\overline{e_{mix}e_2}}{\overline{e_1e_2}} \cdot \frac{\overline{a_{mix}i}}{\overline{e_{mix}i}}\right) - F_{Internal}^{Available} \end{cases} \tag{5.86}$$

5.5.4 Case No. 4 - Allowing Mixing of External Sources and No Constant Total Flowrate

For this case the source-sink diagram given in figure 5.10 will be used again. The constraint equations derived in case 3, i.e. equations 5.80-5.82, are still valid, regardless of the total flow not being constant. Furthermore the internal flowrate constraint in equation 5.70, the overall balance equation 5.71 and the assumed constant internal flowrate in equation 5.73 are also still valid. However the assumption of total constant flow (equation 5.72) is no longer applicable. Fortunately the overall cost function does not change as result of the total flowrate not being constant, i.e. equation 5.74 is still valid. Since the objective function is unchanged as well as the constraints, the solution must also be the same, thus equations 5.83-5.86 are valid also in this case.

The only difference between cases 3 and 4 is that in case 3 the total flowrate is a known constant. Since this is not true for case 4, the solution given by equations 5.83-5.86 represents the fractional contributions of the internal and external sources. In many cases the total flowrate of the system will correspond to the minimum flowrate, which satisfies the sink constraints, i.e. yielding the minimum cost solution.

It should be noted that the solution algorithms presented in this chapter are "building blocks", which may be used for larger problems, e.g. the solution

presented above may not be the best design in some cases. There may be other considerations than just minimum cost for that system, another objective may be superimposed on this problem, e.g. there may be disposal problems with one of the two external sources, thus the objective could be to reuse as much of that solvent as possible. In this case the solution presented by equations 5.83-5.86 provides the inner loop to an optimization procedure, where for different values of the recycle flowrate, the minimum cost solution is identified until the outer loop objective function has been solved.

Once again, it must be emphasized, that the design rules presented in this chapter, are based on visual analysis only, i.e. they are matching only the clustering targets. All the solutions have to be validated to satisfy rule 5.1, since matching the clustering targets is a necessary, but NOT sufficient criteria for matching the original property targets. The AUP values must also match, as pointed out in rule 5.1.

5.6 Summary

In this chapter, a systematic framework for representation of constitutive variables, e.g. properties or functionalities, has been presented. The framework is based on the original property clustering concepts presented by Shelley and El-Halwagi (2000), which have been analyzed and extended in this thesis. The clustering technique utilizes property operators, which are functional relationships describing the physical properties. Although the operator mixing rules are linear, the operators themselves can be nonlinear, e.g. for density, where the density of a mixture is the inverse of the weighted average of the reciprocal densities of the constituents. The property clusters are tailored to satisfy intra- and interstream conservation, thus enabling the development of consistent additive rules along with a ternary visualization.

If the number of property clusters can be limited to three, then the problem and solution can be visualized using triangular diagrams, however the methodology is applicable to systems requiring a higher number of clusters as well. The solution of these problems requires mathematical programming, but since all the mixing rules are linear, the solution is relatively simple.

A systematic procedure for converting ternary coordinates to Cartesian coordinates has been developed, thus allowing for the creation of triangular diagrams using a simple rectangular coordinate system. Furthermore, stepwise procedures have been presented for converting physical property data to cluster points for individual process streams. Converting the property constraints of a process unit to property clusters directly would require extensive enumeration in order to cover the entire acceptable property range. Therefore, a systematic procedure has been developed for identification of the feasibility region boundary, which can be described by just six unique points on the ternary cluster diagram. The necessary conditions, that need to be satisfied in order to feed a given stream or mixture of streams to a unit, have also been pre-

sented. These feasibility conditions ensure that the solution actually satisfies the original property constraints by comparing the AUP values of the clustering solution to the AUP value of the corresponding point on the sink region. Hence, matching the clustering constraints is a necessary but NOT sufficient requirement for satisfying the original property constraints, the AUP values must also match.

Once the problem has been visualized on the ternary cluster diagram, optimization can be performed based on lever-arm analysis. Design rules for different scenarios have been developed and presented.

Property clustering enables systematic tracking of properties or functionalities throughout a process, thus providing a framework for representation and solution of problems that are driven by properties rather than chemical constituency. In addition, the clustering approach provides a representation of the constitutive variables of a system, which can then be utilized in the reverse problem formulation framework.

6

COMPOSITION FREE MODELING

6.1 Introduction

In order to utilize the possibilities of visualizing process synthesis/design problems by means of property clusters it is necessary to have models for different unit operations reformulated in terms of such clusters.

6.2 Fundamental Process Models

In the following the fundamental composition based balance models are derived and reformulated to obtain cluster based models, which satisfy the original mass balance equations. The models are derived for mixing, splitting and stoichiometric conversion reactors, thus covering most reaction, separation and recycle problems.

6.2.1 Mixer

Any mixing operation can be described by a series of binary mixing processes, i.e. where two feed streams are mixed to obtain one product stream.

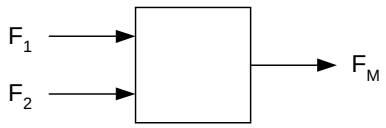

Figure 6.1: Mixer schematic

The overall material balance and the individual component balances for component i may be written as:

$$F_1 + F_2 = F_M \tag{6.1}$$

$$F_1 \cdot y_{i1} + F_2 \cdot y_{i2} = F_M \cdot y_{iM} \tag{6.2}$$

Introducing flowrate fractions and rearranging to find the mixture compositions yields:

$$y_{iM} = x_1 \cdot y_{i1} + x_2 \cdot y_{i2} \qquad , x_1 = \frac{F_1}{F_M} \qquad x_2 = \frac{F_2}{F_M} \tag{6.3}$$

The foundation of the cluster-based description of any system is the ability to adequately describe all the streams by a finite number of j properties. For visualization purposes only 3 properties are used. Equation 6.3 can be rewritten in terms of dimensionless property operators as follows:

$$\Omega_{jM} = x_1 \cdot \Omega_{j1} + x_2 \cdot \Omega_{j2} \tag{6.4}$$

The AUgmented Property index, AUP for the mixture can be calculated by summation of the dimensionless property operators, as defined by equation 5.5:

$$AUP_M = x_1 \cdot AUP_1 + x_2 \cdot AUP_2 \tag{6.5}$$

Equation 6.5 shows that a lever-arm rule exists for calculating the AUP index for a mixture of two streams using only the feed stream properties. Employing the cluster definition given in equation 5.6 the mixture clusters may be calculated.

$$C_{jM} = x_1 \cdot \frac{AUP_1}{AUP_M} \cdot C_{j1} + x_2 \cdot \frac{AUP_2}{AUP_M} \cdot C_{j2} \tag{6.6}$$

Combining equations 5.14 and 6.6 yields:

$$C_{jM} = \beta_1 \cdot C_{j1} + \beta_2 \cdot C_{j2} \tag{6.7}$$

Once again a lever-arm expression is obtained to determine the cluster values of the mixture using only the feed stream information. This was a desired feature of the clusters (inter-stream conservation), as it provides the option of consistent additive rules to be used within a ternary representation of the problem. A cluster composition corresponding to the fractional (relative arm) contributions of the two individual clusters to the mixture cluster is given in equation 6.7 by the β_s parameter. It must be emphasized at this point that the cluster based mixing model represented by equation 6.7 originates from the original mass balance equation, thus any design calculations carried out using equation 6.7 will satisfy the mass balance.

The calculation sequence employed for identification of mixture clusters is presented in table 6.1. It should be emphasized that this calculation sequence is not dependent on whether the feed stream property values are obtained from pure component property values and compositions or experimental property

Step	Description	Equation
1	Calculate dimensionless feed stream property values	5.4
2	Calculate feed stream AUP indices	5.5
3	Calculate ternary cluster values for each feed stream	5.6
4	Calculate flowrate distribution	6.3
5	Calculate AUP index for the mixture	6.5
6	Calculate ternary cluster values for the mixture	6.6

Table 6.1: Identification of mixture clusters

data. Once the mixing operation has been solved in the cluster domain, the result may be converted back to the property domain and if possible the corresponding compositions may be identified. However, it must be noted that if the initial feed stream data is not based on compositions, then the compositions cannot be calculated without inclusion of a property model capable of predicting compositions from the property values alone. In the case where compositions can indeed be calculated, the conversion to properties yields the following equation, where each property j of the mixture is a function of composition:

$$\Omega_{jM} = f_j(y_{iM}) \qquad , i \in [1, NC] \tag{6.8}$$

Furthermore the compositions of the mixture must sum to unity:

$$\sum_{i=1}^{NC} y_{iM} = 1 \tag{6.9}$$

A degree of freedom analysis of the system shows that the number of variables (unknowns) is NC, while the number of equations is $j+1$. Thus the degrees of freedom are $NC-(j+1)$. This means that for $NC > j+1$ the system cannot be uniquely solved. The reason for this result is that a high dimensional system is mapped to a system of only j dimensions. When trying to return to the composition space for $NC > j+1$, the solution is not unique, since infinitely many parameter combinations exist that obey the above equations. However only ONE solution exists that also satisfies the mass balance equations. Therefore by fixing $NC-(j+1)$ compositions from the mass balance equations, this unique solution of the original $NC * NC$ system is guaranteed. Any set of components may be chosen for which to fix the compositions, however in order to have a common rule base, the components $i \in [j+2, NC]$ are chosen. It should be noted that by including the mass balance equations as constraints, the above described problem could also be solved uniquely by mathematical optimization.

6.2.2 Component Splitter

A procedure analogous to the one performed on the fundamental mixer model, can be performed for a component splitter unit. Any splitting operation can be described by a series of binary splitting processes, i.e. where one feed stream is split to obtain two product streams.

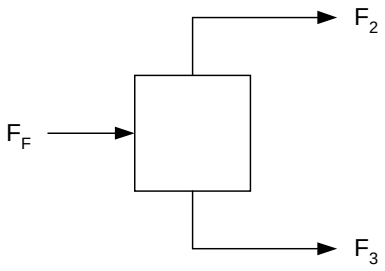

Figure 6.2: Splitter schematic

The individual component balances for component i may be written as:

$$F_F \cdot y_{iF} = F_2 \cdot y_{i2} + F_3 \cdot y_{i3} \tag{6.10}$$

Introducing a product flowrate fraction and rearranging yields:

$$y_{iF} = x_2 \cdot y_{i2} + x_3 \cdot y_{i3} \qquad , x_2 = \frac{F_2}{F_F} \qquad x_3 = \frac{F_3}{F_F} \tag{6.11}$$

Introduction of the component split factors yields:

$$S_i = \frac{F_2 \cdot y_{i2}}{F_F \cdot y_{iF}} \qquad \Leftrightarrow \qquad S_i \cdot y_{iF} = x_2 \cdot y_{i2} \tag{6.12}$$

Summation over all components yields the product flowrate fraction:

$$\sum_{i=1}^{NC} S_i \cdot y_{iF} = x_2 \cdot \sum_{i=1}^{NC} y_{i2} \qquad \Leftrightarrow \qquad x_2 = \sum_{i=1}^{NC} S_i \cdot y_{iF} \tag{6.13}$$

Combining equations 6.11 and 6.12 yields:

$$y_{iF} = S_i \cdot y_{iF} + x_3 \cdot y_{i3} \tag{6.14}$$

For visualization purposes only 3 properties are used. Equation 6.11 can be rewritten in terms of dimensionless property operators as shown in equation 6.15.

$$\Omega_{jF} = x_2 \cdot \Omega_{j2} + x_3 \cdot \Omega_{j3} \tag{6.15}$$

A similar expression can be obtained by reformulating equation 6.14 in terms of dimensionless property operators:

$$\Omega_{jF} = \Omega_{jFSplit} + x_3 \cdot \Omega_{j3} \tag{6.16}$$

In equation 6.16 a pseudo dimensionless property operator $\Omega_{jFSplit}$ is introduced. This parameter describes the relationships between the properties of product stream 2 as a function of the properties of the feed stream. It should be noted that $\Omega_{jFSplit}$ is a function of known variables only, i.e. the split factors, feed stream composition and the pure component property values, and in principle $\Omega_{jFSplit}$ can be described as a property split factor. This new parameter is easily calculated (the annotation for pure component properties, which are marked by a *, uses two indices, i.e. j is the property ID, while i denotes the component ID):

$$\Omega_{jFSplit} = \sum_{i=1}^{NC} \Omega_{ji}^* \cdot S_i \cdot y_{iF} \tag{6.17}$$

Equation 6.17 can be rearranged to provide an expression for the properties of product stream 3:

$$\Omega_{j3} = \frac{\Omega_{jF} - \Omega_{jFSplit}}{x_3} \tag{6.18}$$

Inserting this expression in equation 6.15 provides the corresponding expression for the properties of product stream 2:

$$\Omega_{j2} = \frac{\Omega_{jFSplit}}{x_2} \tag{6.19}$$

The AUgmented Property index for the two product streams can be calculated by summation of the dimensionless property operators, as defined by equation 5.5:

$$AUP_2 = \sum_j \Omega_{j2} = \sum_j \frac{\Omega_{jFSplit}}{x_2} = \frac{1}{x_2} \sum_j \Omega_{jFSplit} \tag{6.20}$$

$$AUP_3 = \sum_j \Omega_{j3} = \sum_j \frac{\Omega_{jF} - \Omega_{jFSplit}}{x_3} = \frac{1}{x_3} \sum_j (\Omega_{jF} - \Omega_{jFSplit}) \tag{6.21}$$

Employing the cluster definition in equation 5.6 the product clusters are calculated:

$$C_{j2} = \frac{\Omega_{j2}}{AUP_2} = \frac{\Omega_{jFSplit}}{\sum_j \Omega_{jFSplit}} \tag{6.22}$$

$$C_{j3} = \frac{\Omega_{j3}}{AUP_3} = \frac{\Omega_{jF} - \Omega_{jFSplit}}{\sum_j (\Omega_{jF} - \Omega_{jFSplit})} \tag{6.23}$$

An interesting feature of the splitter model is that the resulting product clusters are independent of the flowrate distribution even though the stream properties are functions of x_s. Furthermore the conservation of the clusters is achieved, since the two product clusters described by equations 6.22 and 6.23, add up to the property cluster for the original feed stream.

For a given set of component split factors S_i, the calculation sequence given in table 6.2 yields the ternary cluster values for the two product streams. The sequence in table 6.2 includes the calculation of the product stream properties and AUP indices as these are necessary for converting the solution back to composition space. It should be noted however that the product cluster values could have been calculated using only the feed stream information and the component split factors. This means that steps 6 and 7 in table 6.2 are not required for the cluster based solution, but generate the necessary data for the composition based solution.

Step	Description	Equation
1	Calculate dimensionless feed stream property values	5.4
2	Calculate feed stream AUP index	5.5
3	Calculate ternary cluster values for feed stream	5.6
4	Calculate flowrate distribution	6.13
5	Calculate the property split factors	6.17
6	Calculate dimensionless property values for products	6.18,6.19
7	Calculate AUP indices for product streams	6.20,6.21
8	Calculate ternary cluster values for product streams	6.22,6.23

Table 6.2: Identification of product clusters from splitting operation

By repeating the calculation sequence for all parameter combinations of S_i ranging from 0 to 1 in suitable intervals, e.g. with a step size of 0.1, the feasibility region for the splitting operation is obtained. It should be noted that any separation technique and conditions of operation will result in ternary clusters within this region, thus it can be used for identifying the design targets, i.e. the set of separation factors. Furthermore this discretization procedure only serves as a simple means of determining the graphical location of the

feasibility region. It is also possible to derive an analytical expression for the boundaries of this region. Once the splitting operation has been solved in the cluster domain, the results must be converted back to the property domain and finally the corresponding compositions may be identified. After the conversion to properties the following equations, where each property j of the products is a function of composition, are given. Furthermore the compositions in each product stream must sum to unity:

$$\Omega_{j2} = f_j(y_{i2}) \qquad , i \in [1, NC] \tag{6.24}$$

$$\Omega_{j3} = f_j(y_{i3}) \qquad , i \in [1, NC] \tag{6.25}$$

$$\sum_{i=1}^{NC} y_{i2} = 1 \tag{6.26}$$

$$\sum_{i=1}^{NC} y_{i3} = 1 \tag{6.27}$$

A degree of freedom analysis of the system shows that the number of variables (unknowns) is $2NC$, while the number of equations is $2(j+1)$. Thus the degrees of freedom are $2NC - 2(j+1)$. It could be argued that since all the component split factors are known, the compositions of one product stream is also known. However to obtain square matrices, only the compositions for the components $i \in [j+2, NC]$ are calculated by using the split factors. The corresponding compositions in the other product stream are fixed from the mass balance equations, thus yielding 2 $NC * NC$ systems, which can be uniquely solved to obtain the product compositions.

6.2.3 Stoichiometric Conversion Reactor

In order to advocate the use of cluster-based models for design purposes, it is also necessary to be able to handle reactive systems. A procedure analogous to the ones performed on the fundamental mixer and splitter models, can be performed for a stoichiometric conversion reactor. A generalized stoichiometric reaction is given in equation 6.28, where A is considered to be the key component.

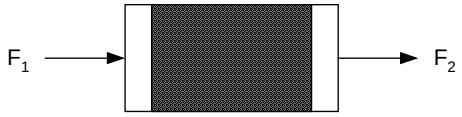

Figure 6.3: Reactor schematic

$$aA + bB \rightleftarrows cC + dD \tag{6.28}$$

Normalizing with respect to the key component and introducing stoichiometric coefficients yields:

$$A + \frac{b}{a}B \rightarrow \frac{c}{a}C + \frac{d}{a}D \tag{6.29}$$

$$\theta_{\mathrm{I}} = \frac{i}{a} \tag{6.30}$$

$$A + \theta_B B \rightarrow \theta_C C + \theta_D D \tag{6.31}$$

The reaction conversion X is defined as follows:

$$X = \frac{F_2 \cdot y_{Key2} - F_1 \cdot y_{Key1}}{F_1 \cdot y_{Key1}} \tag{6.32}$$

The overall component balance equations can be described as:

$$F_2 \cdot y_{i2} = F_1 \cdot y_{i1} + \theta_i \cdot X \cdot F_1 \cdot y_{Key1} \tag{6.33}$$

From equation 6.33 an expression for the effluent flowrate can be obtained by summation over all i components:

$$F_2 = F_1 + X \cdot F_1 \cdot y_{Key1} \cdot \sum_{i=1}^{NC} \theta_i = F_1 \cdot (1 + X \cdot y_{Key1} \cdot \sum_{i=1}^{NC} \theta_i) \tag{6.34}$$

It should be noted, that in equations 6.33 and 6.34, the stoichiometric coefficients obey the following sign rules:

$$Sign(\theta_i) = \begin{cases} Negative & \text{for reactants} \\ Positive & \text{for reaction products} \\ Zero & \text{for others} \end{cases} \tag{6.35}$$

The foundation of the cluster-based description of any system is the ability to adequately describe all the streams by a finite number of properties. For visualization purposes only 3 properties are used. The general annotation for the pure component properties uses two indices, i.e. j is the property ID, while i denotes the component ID. Equation 6.33 can be rewritten in terms of dimensionless property operators as shown in equation 6.36.

$$\Omega_{j2} = \frac{F_1}{F_2} \cdot \left(\Omega_{j1} + X \cdot y_{Key1} \cdot \sum_{i=1}^{NC} \theta_i \cdot \Omega^*_{ji} \right) \tag{6.36}$$

In equation 6.37, Ω_{jREAC}, a pseudo dimensionless property operator reaction term, is introduced. This parameter describes the relationships between the properties of effluent stream as a function of the properties of the feed stream. It should be noted that Ω_{jREAC} is a function of known variables only, i.e. the stoichiometric coefficients, feed stream composition and the pure component property values.

$$\Omega_{jREAC} = y_{Key1} \cdot \sum_{i=1}^{NC} \theta_i \cdot \Omega^*_{ji} \tag{6.37}$$

$$\Omega_{j2} = \frac{F_1}{F_2} \cdot (\Omega_{j1} + X \cdot \Omega_{jREAC}) \tag{6.38}$$

Now the AUP index can be calculated for the reactor effluent stream:

$$AUP_2 = \frac{F_1}{F_2} \sum_j (\Omega_{j1} + X \cdot \Omega_{jREAC}) \tag{6.39}$$

Employing the cluster definition given in equation 5.6 yields the ternary cluster values for the effluent stream:

$$C_{j2} = \frac{\Omega_{j2}}{AUP_2} = \frac{\Omega_{j1} + X \cdot \Omega_{jREAC}}{\sum\limits_j (\Omega_{j1} + X \cdot \Omega_{jREAC})} \tag{6.40}$$

The sequence in table 6.3 includes the calculation of the product stream properties and AUP index as these are necessary for converting the solution back to composition space. It should be noted however that the product cluster values could have been calculated using only the feed stream information, the stoichiometric coefficients and the conversion factor. This means that steps 5 and 6 in table 6.3 are not required for the cluster based solution but generate the necessary data for the composition based solution. By repeating the calculation sequence for all values of X ranging from 0 to 1 in suitable intervals, e.g. with a step size of 0.01, the feasibility region for the reaction is obtained.

Once the reactor model has been solved in the cluster domain, the results may be converted back to the property domain and finally the corresponding compositions may be identified.

The conversion to properties yields equation 6.41, where each property j of the products is a function of composition. Furthermore the compositions in the effluent stream must sum to unity as shown in equation 6.42.

Step	Description	Equation
1	Calculate dimensionless feed stream property values	5.4
2	Calculate feed stream AUP index	5.5
3	Calculate ternary cluster values for feed stream	5.6
4	Calculate effluent flowrate from feed stream data	6.34
5	Calculate property reaction term from feed stream data	6.37
6	Calculate dimensionless property values for effluent stream	6.38
7	Calculate AUP index for the effluent stream	6.39
8	Calculate ternary cluster values for effluent stream	6.40

Table 6.3: Identification of product clusters from reaction

$$\Omega_{j2} = f_j(y_{i2}) \qquad , i \in [1, NC] \tag{6.41}$$

$$\sum_{i=1}^{NC} y_{i2} = 1 \tag{6.42}$$

A degree of freedom analysis of the system shows that the number of variables (unknowns) is NC, while the number of equations is $j+1$. Thus the degrees of freedom are $NC-(j+1)$. This means that for $NC > j+1$ the system cannot be uniquely solved. However, the reaction stoichiometry reduces the degrees of freedom, i.e. the reactor effluent is explicitly given, if the conversion factor and the stoichiometric coefficients are known, hence the system is solvable without having to fix additional parameters.

6.3 Proof of Concept Example

The problem to be solved involves choosing the correct sequence of mixers and splitters for matching a set of target values. It should be noted that the example is based on purely theoretical values, which are not related to any specific components or properties. The purpose of this example is solely to illustrate and validate the use of mixing and splitting operations for solving design problems in cluster space.

6.3.1 Problem Formulation

The objective is to match a set property values for a product stream by mixing and splitting two feed streams accordingly. Three properties P_1, P_2 and P_3 have been found to be able to characterize the streams. The initial inputs, i.e. pure component property values, property references, stream summaries as well as the desired property targets are given in tables 6.4-6.6.

Component i	P^*_{1i}	P^*_{2i}	P^*_{3i}
1	50	1.2	10
2	65	0.5	8
3	38	0.8	12
4	100	2.0	4
5	118	0.2	3
6	75	0.5	12

Table 6.4: Pure component property data

Data	Feed stream 1	Feed stream 2
y_{1s}	0.1889	0.0000
y_{2s}	0.4667	0.0000
y_{3s}	0.3444	0.0000
y_{4s}	0.0000	0.1000
y_{5s}	0.0000	0.7500
y_{6s}	0.0000	0.1500
Flowrate	10.0	150.0

Table 6.5: Feed stream summaries

Property	Target value	Reference value
P_1	78.840	50
P_2	1.247	1
P_3	7.688	7

Table 6.6: Property targets and reference values

6.3.2 Visualization of Problem

Converting the feed stream and property target information to clusters is achieved by employing equations 5.4-5.6. The resulting cluster mapping is given in figure 6.4. Since all mixing operations within the cluster diagram are described by a straight line, it is evident from figure 6.4 that it is NOT possible to mix the two feed streams in any ratio to match the property targets. Therefore it is necessary to split at least one of the streams. The feasibility regions shown in figure 6.5 are identified by employing the calculation sequence outlined in table 6.2 using a parameter step size of 0.1.

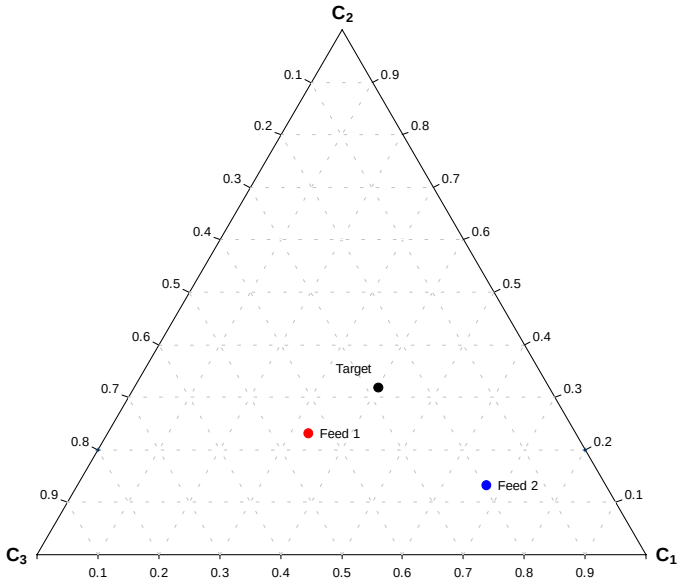

Figure 6.4: Cluster diagram of feed and product streams

6.3.3 Identification of Operating Sequence

It is desired to use a minimum number of processing units, thus it is decided to split feed stream 2 and mix one of the products with feed stream 1. Drawing a straight line between the cluster points for feed stream 1 and the desired product provides the operating line for the mixing operation. The stream to be mixed with feed stream 1 to match the target MUST lie on this line in such a location that the target cluster is between the two streams to be mixed. Furthermore the stream must also be within the split feasibility region of feed stream 2. In figure 6.6 the solid black line represents the operating line for feasible mixing agents that match the property target. However the

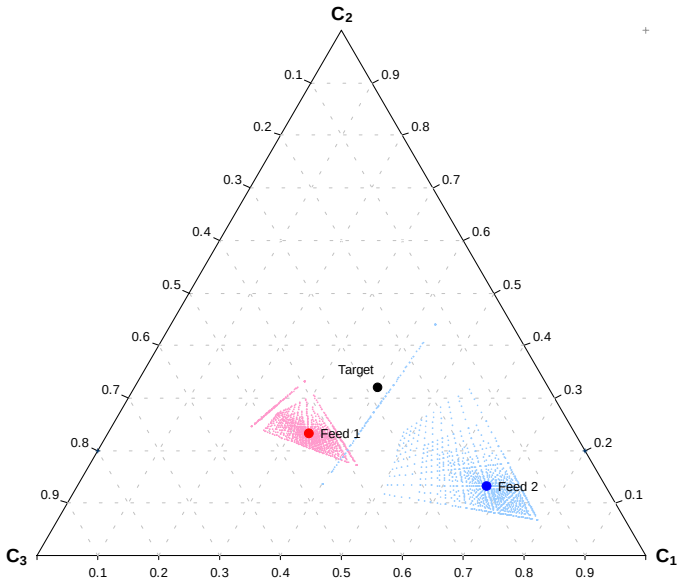

Figure 6.5: Cluster diagram including feasible splitting regions

only cluster points that also satisfy the feasibility constraints are designated by the grey line segment. In this example, it is decided to split feed stream 2 in such a way that the minimum arm of feed stream 1 is used. This means that when employing lever-arm analysis at the mixing point (target point), the arm representing feed stream 1 should be minimized. Using this objective, the optimal point to be mixed with feed stream 1 is the point located just on the border of the feasibility region in figure 6.6. When one of the products of a splitting operation is defined the other product will be located on a straight line from the first product and extended through the feed point. How far away from the feed point the second product is located depends on the choice of split factors. In this example any set of component split factors resulting in the first product are valid solutions, e.g. as shown in figure 6.7.

A powerful feature of the cluster-based mapping diagram is the ability to directly obtain the corresponding process flowsheet. This is possible because the formulation of the clusters and the unit operation models satisfy the overall balance equations. It should be emphasized that once the problem was reformulated in terms of clusters all the design calculations were performed graphically and composition free. Along with the process flowsheet (shown in figure 6.8), the stream summaries in terms of ternary cluster values and flowrates along with flowrate distributions for the two units are available. Employing the calculation sequence outlined earlier, the compositions in each stream from the cluster solution can be calculated and are given in table 6.7.

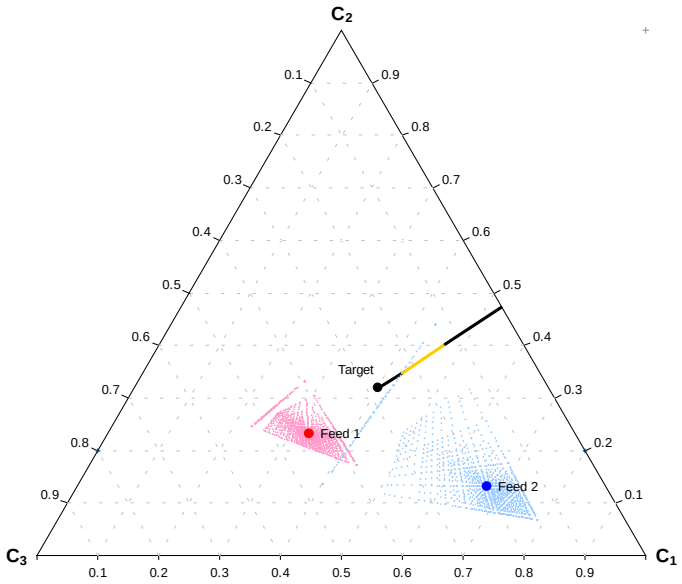

Figure 6.6: Feasible mixing points to achieve target

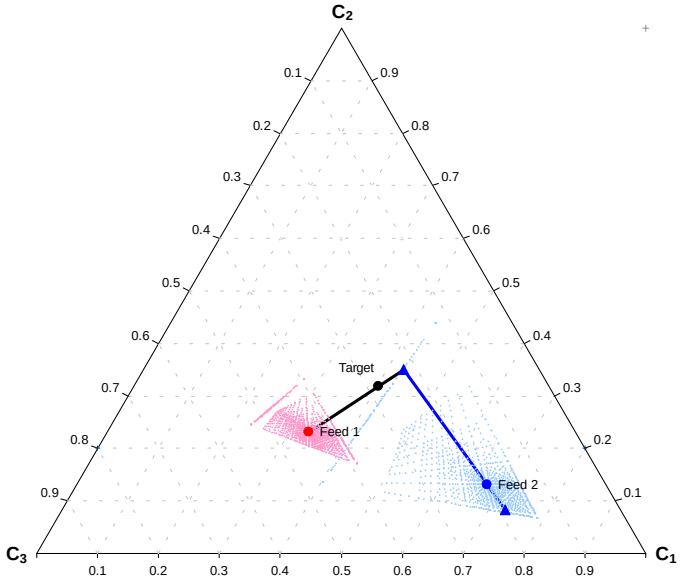

Figure 6.7: Feasible operational route

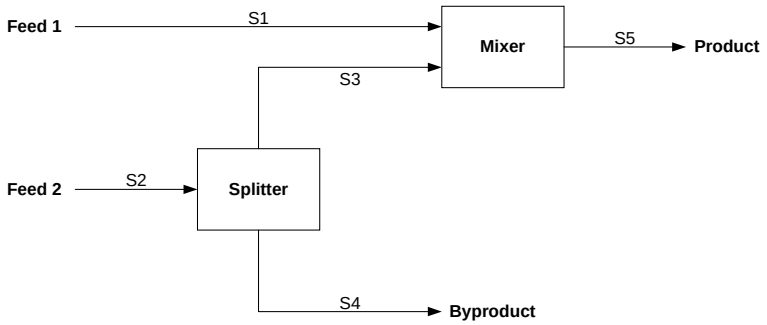

Figure 6.8: Feasible flowsheet obtained from cluster diagram

Stream ID	$S1$	$S2$	$S3$	$S4$	$S5$
y_1	0.1889	0.0000	0.0000	0.0000	0.0624
y_2	0.4667	0.0000	0.0000	0.0000	0.1543
y_3	0.3444	0.0000	0.0000	0.0000	0.1139
y_4	0.0000	0.1000	0.6667	0.0116	0.4463
y_5	0.0000	0.7500	0.0000	0.8670	0.0000
y_6	0.0000	0.1500	0.3333	0.1214	0.2231
Flowrate	10.0	150.0	20.25	129.75	30.25

Table 6.7: Composition based stream summary

Once again it should be pointed out that all the design calculations were performed on a composition free basis. The algorithm solves the process design problem in terms of property values providing design targets for the constitutive variables. In this particular example the design targets obtained by the reverse problem formulation are the component split factors. The second reverse problem consists of identifying the separation technique capable of matching these targets. In this chapter, solution of the constitutive equations to find the matching splitting operation will not be investigated further.

6.4 Dynamic Considerations

Most methods for chemical process design are limited to steady-state operation, where the process variables do not change in time. However, a lot of processes can not be considered as time invariant, since certain variables vary significantly as a function of time, e.g. in batch and periodic processes, where the conditions can be very different at the end of the batch or the middle of the period, compared to the beginning.

In this section a brief introduction, to how the property clustering techniques can be utilized in a dynamic environment, is presented. The primary requirement for establishing a framework capable of handling dynamic operation is the

ability to track the property values as a function of time. This can be achieved through online measurements or by estimation methods, although it must be emphasized that it may be cumbersome to develop or identify dynamic property estimation methods as the model complexity can be expected to increase significantly to take the dynamic interactions into account. However, the property operator description does not change regardless of how the property values are obtained, therefore the dynamic version of the original property operator mixing rule presented in equation 5.1 can be written as:

$$\psi_j(P_{jM}(t)) = \sum_{s=1}^{N_s} \frac{F_s(t)}{\sum_{s=1}^{N_s} F_s(t)} \cdot \psi_j(P_{js}(t)) = \sum_{s=1}^{N_s} x_s(t) \cdot \psi_j(P_{js}(t)) \tag{6.43}$$

Normalizing the property operators by a reference value yields the dimensionless property operator expression in equation 6.44.

$$\Omega_{js}(t) = \frac{\psi_j(P_{js}(t))}{\psi_j(P_j^{ref})} \tag{6.44}$$

The dynamic AUgmented Property index AUP_s for each stream s can be calculated as the sum of the dynamic dimensionless property operators:

$$AUP_s(t) = \sum_{j=1}^{NP} \Omega_{js}(t) \tag{6.45}$$

The dynamic property cluster C_{js} for property j of stream s can now be defined analogously to equation 5.6:

$$C_{js}(t) = \frac{\Omega_{js}(t)}{AUP_s(t)} \tag{6.46}$$

In most cases the feed constraints for the different units in the process do not change in time, therefore the procedure outlined in section 5.4 is still valid for the identification of the feasibility region boundary. So for each process sink, the feasibility region boundary is uniquely described by the six points presented in rule 5.6.

To illustrate how the clustering concepts may be applied in a dynamic environment a simple example is used. In figure 6.9 a process stream, e.g. a reactor effluent is tracked as a function of time. As the reaction progresses, the properties of the stream change. The reactor effluent is to be sent to another processing unit with specified constraints on the feed conditions. The feasibility region boundary is depicted as a dashed line forming a hexagon. At time t_3 the reactor effluent enters the feasibility region of the sink, so at this point the AUP value of the stream must be compared to the AUP value of

corresponding point on the sink region in order to validate the match according to rule 5.1. If the values match, then the stream may be fed to the sink. At time t_4, the reactor effluent leaves the feasibility region and does not re-enter. This means that the reaction must run until at least time t_3 and no later than t_4, and then the effluent can be sent to the next unit, if the AUP validation was successful.

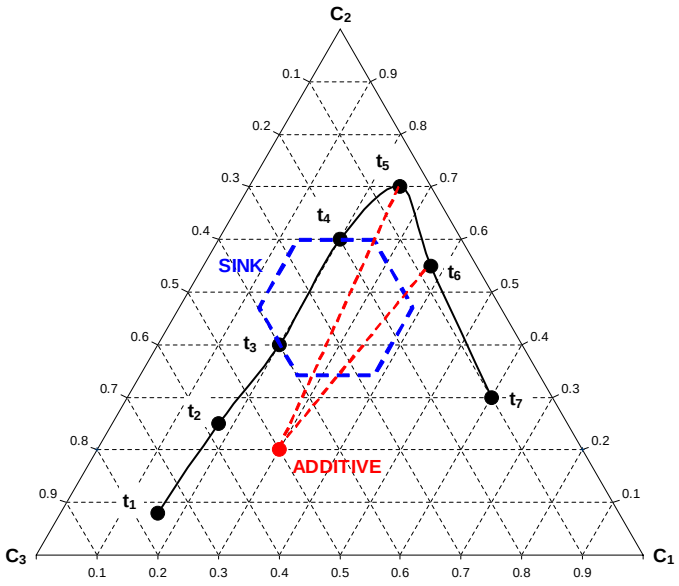

Figure 6.9: Example of dynamic source-sink mapping diagram

The dynamic tracking of the stream properties also allows for identification of mixing ratios for additives. Consider, the case where, due to some processing or equipment constraint, it is not feasible to end the reaction after t_4. The reactor effluent may be mixed with an additive stream to satisfy the sink constraints at times t_5 and t_6. At each time instant the appropriate amount can be identified graphically using lever-arm analysis and then subsequently the mixture is verified by comparing the AUP values of the sink region to that of the designed mixture.

It should be emphasized that the general framework outlined in this section can be applied to a wide range of systems, provided that property operator descriptions can be identified for the targeted properties and furthermore that property data as a function of time is also available either through experimental data or estimation methods. If such dynamic property operator descriptions can be identified, a significant model order reduction can be expected from the use of property clusters and thus the clustering approach would also yield simplified estimators and controllers from the low dimensional system.

7

PROPERTY BASED PROCESS DESIGN

7.1 Introduction

Standard techniques for process design are based on tracking individual chemical species. Component material balances are at the heart of any design approach. Nonetheless, many design problems are not component dependent, but are driven by properties, e.g. the design of a paper machine, where the quality of the product cannot be described by composition alone. The quality of paper is described by the properties of the paper, such as opacity, reflectivity etc. Paper consists primarily of cellulose and hemicellulose, however the properties of the paper cannot be directly related to the composition of cellulose, i.e. even 100% pure cellulose may not satisfy the quality requirements as they are functions of fiber length, fibrocity etc. Recently, the concept of clustering has been introduced to enable the conserved tracking of surrogate properties (Shelley and El-Halwagi, 2000; Eden *et al.*, 2002b). Hence, the process design can be optimized based on integrating properties instead of chemical species. Property integration is defined as a functionality-based, holistic approach to the allocation and manipulation of streams and processing units based on tracking, adjustment, assignment, and matching of functionalities or properties throughout the process. As the property integration framework relies on techniques developed as part of the mass integration methodology (El-Halwagi, 1997), an updated terminology is given below. It is apparent that these terms can be used interchangeably when describing different problems as there is some inherent overlap between them.

- **Tracking** – Ability to identify and monitor how the properties of a stream change throughout the process.
- **Adjustment or interception** – Utilization of processing units to adjust the properties of a given stream to make it acceptable as feed to other units.
- **Assignment or allocation** – Routing of streams to appropriate process units in order to achieve certain process objectives.

- **Matching** – Ensuring that all constraints on process units or products are satisfied by the respective feed streams.

It should be emphasized that the property based process design methods presented in this work are NOT meant to replace conventional composition and component based methods. The property clustering techniques are intended to complement existing methods by allowing solution of problems, that conventional methods are not capable of handling. Furthermore, the emphasis is on development and application of visual techniques to obtain valuable insights that are traditionally hidden by not integrating the properties directly into the design procedure.

7.2 Problem Definition

Given a process with certain sources (streams) and sinks (units) along with their properties and constraints, it is desired to identify optimum strategies for allocation and interception that integrate the properties of sources, sinks, and interceptors so as to optimize a desirable process objective (e.g., minimum usage of fresh resources, maximum utilization of process resources, minimum cost of external streams) while satisfying the constraints on properties and flowrate for the sinks. The sources are divided into two classes, i.e. internal process resources, which are limited by their availability and external sources, which have to be utilized optimally to satisfy the process objectives.

7.3 Visualization and Solution Strategy

As stated in the problem definition, the starting point for the property based design methodology is the description of the system in terms of properties. Figure 7.1 illustrates how this property description may be obtained. Two approaches exist for obtaining the stream properties, i.e. the properties may be estimated from pure component property values and a composition based description of the streams or experimental data may be available for the desired properties. From either alternative, the stream properties are identified and can then subsequently be converted to cluster values and visualized according to the procedure given in table 5.1. It is important to point out that once the problem is described in the cluster domain, any design calculations can now be performed on a composition free basis. It should be emphasized, however, that once the design problem has been solved, the cluster based solution can only be converted back to the corresponding property domain. Hence, if the stream properties were given as experimental values, it is not possible to directly obtain the component compositions unless a suitable predictive model can be identified.

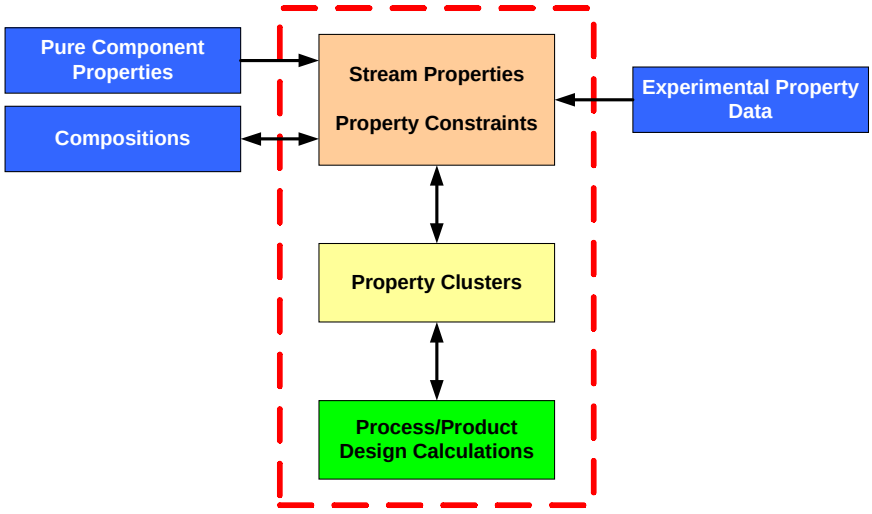

Figure 7.1: Conversion of properties to clusters

7.4 VOC Recovery from Metal Degreasing

To illustrate the application of constitutive or property based modeling, a case study of a metal degreasing facility is presented. The metal degreasing process presented in figure 7.2 uses a fresh organic solvent in the absorption column and another one in the degreaser (Shelley and El-Halwagi, 2000). Currently, the off-gas Volatile Organic Compounds (VOCs) evaporating from the degreasing process are simply flared, leading to economic loss and environmental pollution.

In this case study, the objective is to explore the possibility of condensing and reusing the off-gas VOCs, thus optimizing the usage of fresh solvents and simultaneously identify candidate solvents for both units. Three properties are examined to determine the suitability of a given organic process fluid for use in the absorber and/or degreaser:

- **Sulfur content** (S) - for corrosion considerations, expressed as weight percent.
- **Density** (ρ) - for hydrodynamic and pumping aspects.
- **Reid Vapor Pressure** (RVP) - for volatility, makeup and regeneration, defined as the vapor pressure at 100,° F, corresponding to 310.93K.

The solvents to be synthesized are pure component fluids, thus the sulfur content of these streams is zero. The constraints on the inlet conditions of the feed streams to the absorber and degreaser respectively are given in tables 7.1 and 7.2, while the property operator mixing rules are given in equations 7.1-7.3. In addition, experimental data are available for the degreaser off-gas condensate.

Samples of the off-gas were taken, and then condensed at various condensation temperatures ranging from 280K to 315K, providing measurements of the three properties as well as the flowrate of the condensate. The obtained data are presented in figures 7.3 and 7.4 (Shelley and El-Halwagi, 2000).

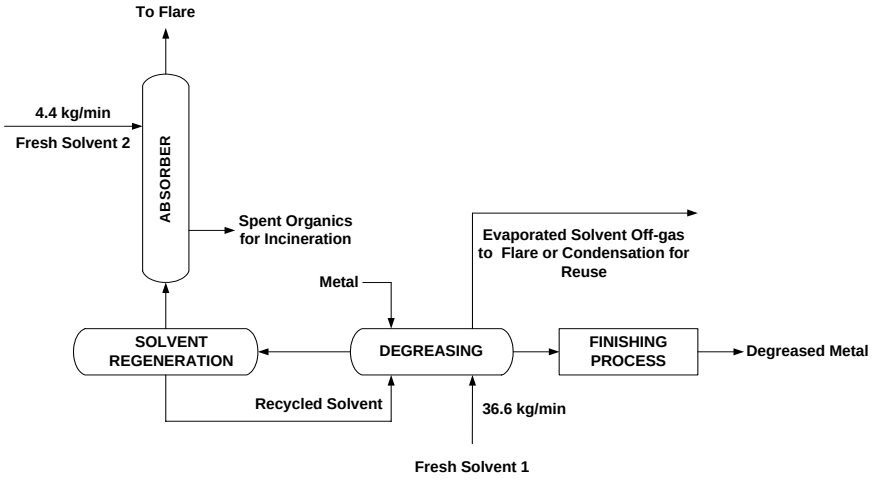

Figure 7.2: Schematic representation of original metal degreasing process

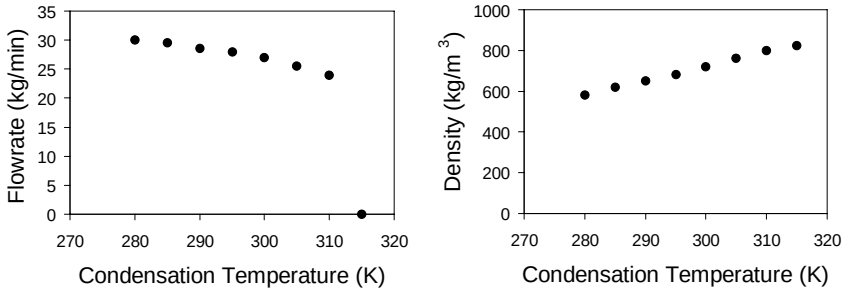

Figure 7.3: VOC condensation data for flowrate and density

$$S_M = \sum_{s=1}^{N_s} x_s \cdot S_s \qquad , S^{ref} = 0.5\text{wt.\%} \tag{7.1}$$

$$\frac{1}{\rho_M} = \sum_{s=1}^{N_s} x_s \cdot \frac{1}{\rho_s} \qquad , \rho^{ref} = 1000\text{kg/m}^3 \tag{7.2}$$

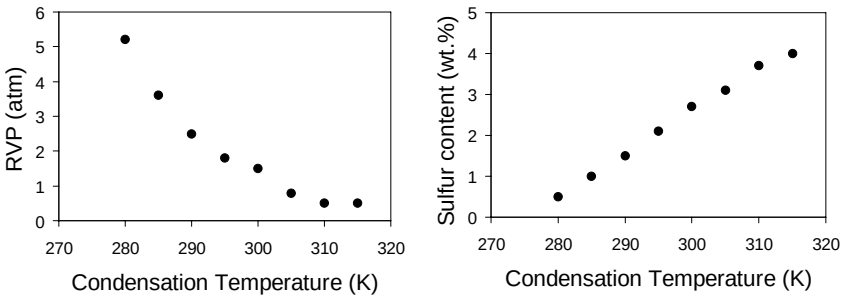

Figure 7.4: VOC condensation data for Reid vapor pressure and sulfur content

$$RVP_M^{1.44} = \sum_{s=1}^{N_s} x_s \cdot RVP_s^{1.44} \qquad , RVP^{ref} = 1.0\text{atm} \tag{7.3}$$

Property	Lower Bound	Upper Bound
S (wt. %)	0.00	0.10
ρ (kg/m^3)	530	610
RVP (atm)	1.5	2.5
Flowrate (kg/min)	4.4	6.2

Table 7.1: Feed constraints for absorber

Property	Lower Bound	Upper Bound
S (wt. %)	0.00	1.00
ρ (kg/m^3)	555	615
RVP (atm)	2.1	4.0
Flowrate (kg/min)	36.6	36.8

Table 7.2: Feed constraints for degreaser

The data in figures 7.3 and 7.4 correspond to the condensation route given in figure 7.5 and were converted to cluster values according to table 5.1. The unit feed constraints were converted to cluster values according to rule 5.6 yielding the two regions for the absorber and degreaser respectively. The cluster data was plotted and the feasible mixing paths identified. Since the fresh process fluids contain no sulfur, any feasible solution will be on the C_2-C_3 axis. Lever-arm analysis is employed to identify the minimum flow solutions.

At a condensation temperature of 280 K the condensate flowrate is 30.0 kg/min. Since the minimum flowrate requirement for the degreaser feed is 36.6 kg/min, the target for minimum fresh solvent usage is 6.6 kg/min assuming that a suitable solvent can be identified. Lever-arm analysis is employed to identify the minimum feasible flowrate, i.e. for a solvent that when mixed with the condensate actually satisfies the constraints for the degreaser. This analysis showed that of the feasible mixing paths the minimum feasible flowrate of the fresh solvent is 11.8 kg/min. In order to investigate whether a better solution exists, the same analysis was performed for a condensation temperature of 285 K. At this temperature the condensate flowrate is slightly reduced (29.5 kg/min), thus resulting in a target for minimum fresh solvent usage of 7.1 kg/min. When performing the lever-arm analysis to identify the minimum feasible solvent flowrate the result was also 7.1 kg/min, i.e. the target can be matched at this condensation temperature. It should be noted that using this approach the flowrate of the fresh material has been reduced by approximately 80%.

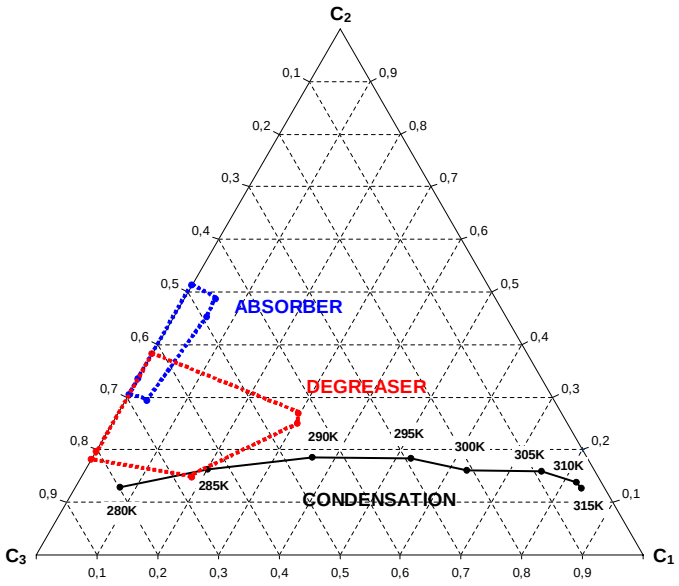

Figure 7.5: Ternary representation of metal degreasing problem

The analysis showed that the cluster solutions to the degreaser problem correspond to the degreaser points on the C_2-C_3 axis. Since all the condensate has been recycled to the degreaser, the solution for the absorber is a simple molecular design problem. Using the information obtained from the cluster diagram mapping analysis, a computer-based tool ProCAMD, which is a part of the ICAS software package (CAPEC, 2003), was invoked to synthesize candidate

process fluids. Not allowing phenols, amines, amides or compounds containing silicon, sulfur or halogens, due to safety and health considerations, reduced the search space. The CAMD algorithm (Harper, 2000; CAPEC, 2003) yielded a series of candidate solvents for each of the process units. Of the candidate compounds identified by the software, iso-Pentane was chosen for the absorber and n-Butane for the degreaser.

It is important to point out that the case study is solved in terms of properties only, i.e. no component information or compositions were needed to obtain the design targets. The reason for this is that experimental data was available for the properties of the individual streams. Therefore it is straightforward to convert the property values to cluster values using table 5.1. The design calculations follow the methodology outlined previously, furthermore in terms of reverse simulation, the conditions of operation (intensive variables) for the condenser, i.e. condensation temperature, is identified, instead of the unit operation. The objective of the case study was to investigate the possibilities of using condensation of the degreaser off-gas as a substitute solvent, thus the unit operation was fixed, however the operating conditions were not known.

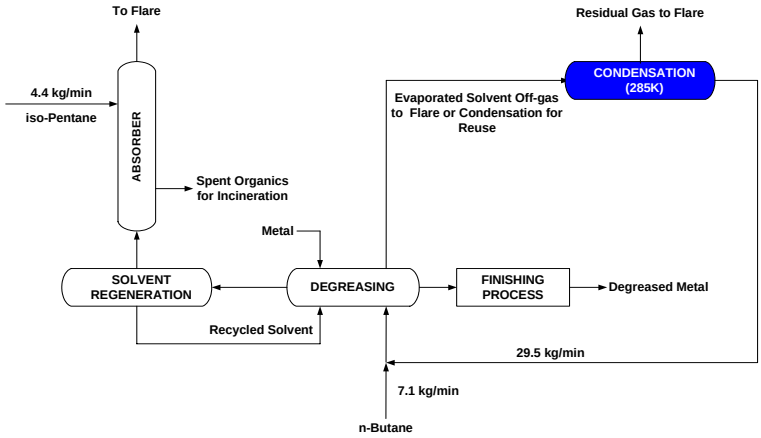

Figure 7.6: Schematic representation of metal degreasing process after property integration

7.5 Resource Conservation in Papermaking

To further illustrate the usefulness of constitutive or property based modeling, a case study of a papermaking facility is presented. A schematic of the process is given in figure 7.7. Wood chips are chemically cooked in a Kraft digester using white liquor (containing sodium hydroxide and sodium sulfide as main active ingredients). The spent solution (black liquor) is converted back to white liquor via a recovery cycle (evaporation, burning, and caustification). The digested

pulp is passed to a bleaching system to produce bleached pulp (fiber). The paper machine employs 100 ton/hr of the fibers. As a result of processing flaws and interruptions, a certain amount of partly and completely manufactured paper is rejected. These waste fibers are referred to as broke. The reject is passed through a hydropulper followed by a hydrosieve with the net result of producing an underflow, which is burned, and an overflow of broke, which goes to waste treatment. It is worth noting that the broke contains fibers that may be partially recycled for papermaking (Eden *et al.*, 2003a,b).

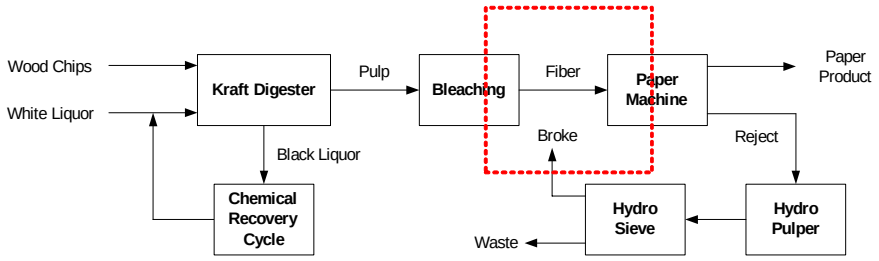

Figure 7.7: Schematic representation of pulp and paper process

The objective of this case study is to identify the potential for recycling the broke back to the paper machine, thus reducing the fresh fiber requirement and maximize the resource utilization. This gives rise to two solution strategies, each with different questions to be answered:

- **Direct recycle and reallocation** - What is the optimal allocation of the fiber sources (fresh fiber and broke) for a direct recycle/reuse situation, i.e. not allowing for new equipment?

- **Interception of broke** - To maximize the use of process resources and minimize wasteful discharge (broke), how should the properties of the broke be altered so as to achieve its maximum recycle?

Three primary properties determine the performance of the paper machine and thus consequently the quality of the produced paper (Willets, 1958; Brandon, 1981; Biermann, 1996):

- **Objectionable material (OM)** - refers to undesired species in the fibers, expressed as mass fraction.

- **Absorption coefficient (k)** - intensive property providing a measure of absorptivity of light into the fibers (black paper has a high value of k).

- **Reflectivity (R_∞)** - defined as the reflectance of an infinitely thick material compared to an absolute standard, which is Magnesium Oxide (MgO).

Hemicellulose and cellulose have very little absorption of light in the visible region, however, lignin has a high absorbance. Therefore, light absorbance is mostly attributed to lignin. The light absorption coefficient is a very useful property in determining the opacity of the fibers. Opacity ($C_{0.89}$) is defined as the ratio of reflectance of a single sheet, which is backed by a black body, compared to a sheet that is backed by a white body at 89% reflectance. Values and relationships of opacity, reflectivity, and the adsorption coefficient can be determined using the Kubelka-Munk theory (Biermann, 1996), which relates the basis weight of paper, opacity, and reflectivity of paper to one another.

In order to convert property values from raw property data to cluster values, property operator mixing rules are required. According to Brandon (1981), the mixing rules for objectionable material (OM) and absorption coefficient (k) are linear, while a nonlinear empirical mixing rule for reflectivity (R_∞) has been developed (Willets, 1958). The property operator mixing rules are presented in equations 7.4-7.6 along with the chosen reference values. The property data for the fresh fibers along with the broke is given in table 7.3, while the constraints of the paper machine are presented in table 7.4.

$$OM_M = \sum_{s=1}^{N_s} x_s \cdot OM_s \qquad , OM^{ref} = 0.01 \tag{7.4}$$

$$k_M = \sum_{s=1}^{N_s} x_s \cdot k_s \qquad , k^{ref} = 0.001 \tag{7.5}$$

$$R_{\infty M}^{5.92} = \sum_{s=1}^{N_s} x_s \cdot R_{\infty s}^{5.92} \qquad , R_\infty^{ref} = 0.01 \tag{7.6}$$

Property	Fibers	Broke
OM (mass fraction)	0.000	0.115
k (m^2/g)	0.0012	0.0013
R_∞ (fraction)	0.82	0.90
Flowrate (ton/hr)	100	30

Table 7.3: Properties of fiber sources

From table 7.4 it is apparent that the target for minimum resource consumption of fresh fibers is 70 ton/hr (100-30) assuming that all the broke can be recycled to the paper machine. The problem is visualized by converting the property values to cluster values using equations 5.1 and 5.4-5.6. The paper machine constraints are represented as a feasibility region, which is identified according to the procedure outlined in section 5.4 for the data given in table 7.4. The resulting ternary diagram is shown in figure 7.8, where the dotted line represents the feasibility region for the paper machine feed. The relationship between

Property	Lower Bound	Upper Bound
OM (mass fraction)	0.00	0.02
k (m^2/g)	0.00115	0.00125
R_∞ (fraction)	0.80	0.90
Flowrate (ton/hr)	100	105

Table 7.4: Feed constraints for paper machine

the cluster values and the corresponding AUP values ensures uniqueness when mapping the results back to the property domain.

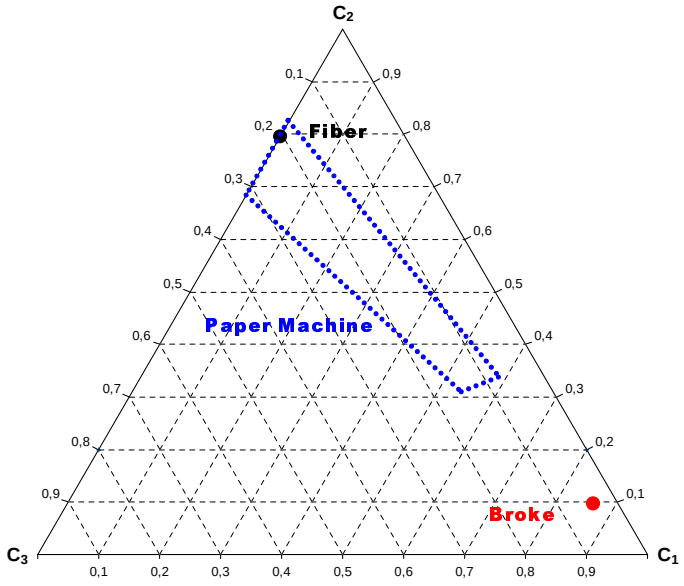

Figure 7.8: Ternary representation of pulp and paper problem

Since the optimal flowrates of the fibers and the broke are not known, a reverse problem is solved to identify the clustering target corresponding to maximum recycle. In order to minimize the use of fresh fiber, the relative cluster arm for the fiber has to minimized, i.e. the optimum feed mixture will be located on the boundary of the feasibility region for the paper machine. The cluster target values to be matched by mixing the fibers and broke are identified graphically and represented as the intersection of the mixing line and the feasibility region in figure 7.9. From the calculation of the feasibility region the cluster and AUP values for the mixing point are known. Using these results the stream fractions can be calculated from equation 5.17. The resulting mixture is calculated to consist of 83 ton/hr of fiber and 17 ton/hr of broke.

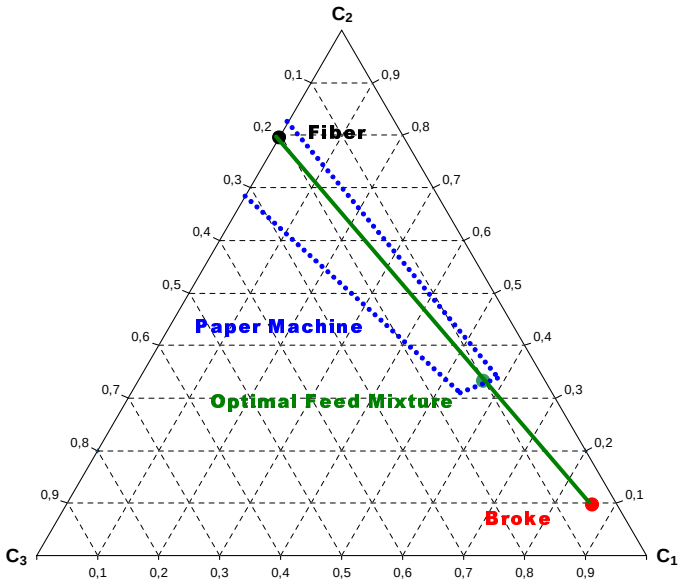

Figure 7.9: Identification of optimal feed to paper machine using direct recycle

Direct recycle can thus reduce the fiber usage from 100 ton/hr to 83 ton/hr, however it does NOT achieve the minimum fiber usage target of 70 ton/hr. Therefore the properties of the broke will have to be altered to match the maximum recycle target. Assuming that the feed mixture point is unchanged, and since the fractional contribution of the fibers and the intercepted broke are 70% and 30% respectively, the cluster "compositions" or cluster lever-arms (β_s) can be calculated from equation 5.14. Now the cluster values for the intercepted broke can be readily calculated from equation 5.13, and the resulting point is shown in figure 7.10. This reverse problem identifies the clustering target, which can then be converted to a set of property targets as given in table 7.5. Note that for each mixing point on the boundary of the feasibility region, a clustering target exists for the intercepted broke, so the reverse problem formulation technique is actually capable of identifying all the alternative product targets that will solve this particular problem.

Property	Original Broke	Intercepted Broke
OM (mass fraction)	0.115	0.067
k (m^2/g)	0.0013	0.0011
R_∞ (fraction)	0.90	0.879

Table 7.5: Properties of intercepted broke achieving maximum recycle

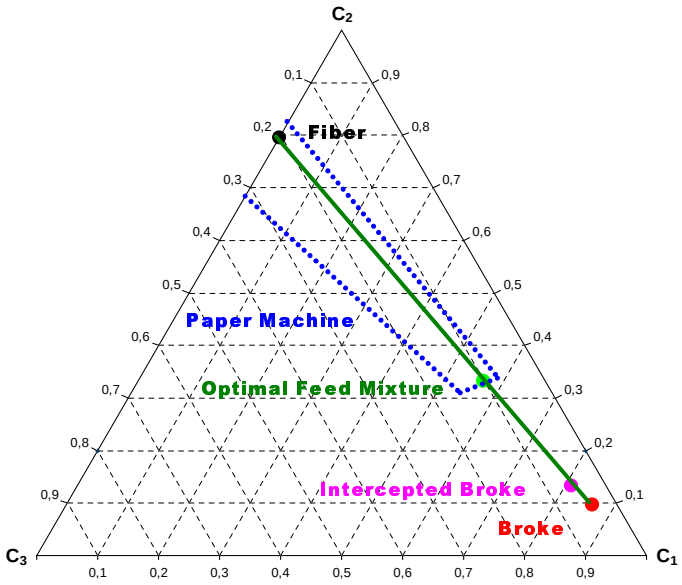

Figure 7.10: Identification of property interception targets for maximum recycle to paper machine

Solution of the second reverse problem, i.e. identification of the processing steps required for performing the property interception described by table 7.5, is not presented in this work. Most processes for altering or fine tuning paper properties are considered proprietary material, however the interception can be performed chemically and/or mechanically (Brandon, 1981; Biermann, 1996).

7.6 Water Conservation in Microelectronics Facility

To illustrate the application of property clustering techniques for pollution prevention purposes, a case study of a microelectronics manufacturing facility is presented. A schematic of the process is given in figure 7.11. A water treatment facility processes 2700 gallons per minute (gpm) of municipal water in several stages to produce a total of 2000 gpm of Ultra Pure Water (UPW). The purified water product from the treatment facility is distributed equally to two sections of the microelectronics manufacturing plant, i.e. the wafer fabrication section and the combined Chemical and Mechanical Processing section (CMP). The effluent from the wafer fabrication section can be treated by municipal waste water treatment plants, however the effluent from the CMP is too polluted for the municipal waste water treatment system and is sent to an industrial waste

water treatment facility for purification.

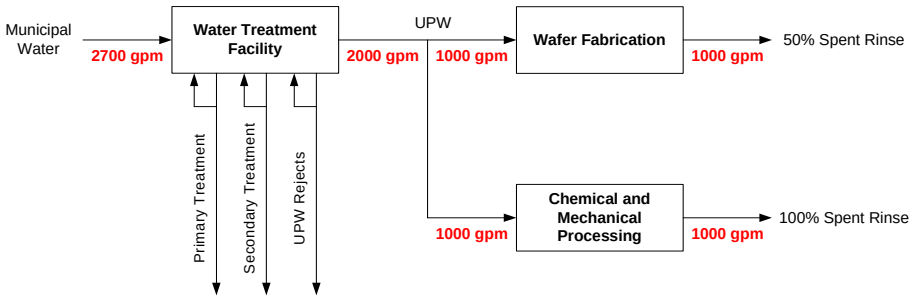

Figure 7.11: Schematic representation of microelectronics processing

The objective of this case study is to identify the potential for partly recycling the effluent from the wafer fabrication, thus reducing the UPW requirement. Three primary properties determine whether a water stream can be fed to the wafer fabrication section and the CMP section (Gabriel *et al.*, 2003a):

- **pH Factor** (pH) - for corrosion purposes.
- **Resistivity** (R) - measure of ionic content of the water $(k\Omega/cm)$.
- **Total Organic Content** (TOC) - constitute primary pollutants (ppm).

In order to convert property values from raw property data to cluster values, property operator mixing rules are required. According to Gabriel *et al.* (2003a), the property operator mixing rules can be described as presented in equations 7.7-7.9 along with the chosen reference values. The property data for UPW and the 50% Spent Rinse are given in table 7.6, while the constraints of the wafer fabrication and the CMP sections are presented in tables 7.7 and 7.8, respectively.

$$pH_M = \sum_{s=1}^{N_s} x_s \cdot 10^{-pH_s} \qquad , pH^{ref} = 7 \tag{7.7}$$

$$R_M = \sum_{s=1}^{N_s} x_s \cdot R_s^{-1.00786} \qquad , R^{ref} = 20000 \tag{7.8}$$

$$TOC_M = \sum_{s=1}^{N_s} x_s \cdot TOC_s \qquad , TOC^{ref} = 1 \tag{7.9}$$

The problem is visualized by converting the property values to cluster values using equations 5.1 and 5.4-5.6. The wafer fabrication and CMP section constraints are represented as feasibility regions, which are identified according to

Property	Ultra Pure Water (UPW)	50% Spent Rinse
pH	6.5	6.1
R ($k\Omega/cm$)	18000	5000
TOC (ppm)	2	20
Flowrate (gal/min)	2000	1000

Table 7.6: Properties of water sources

Property	Lower Bound	Upper Bound
pH	6.0	7.0
R ($k\Omega/cm$)	18000	20000
TOC (ppm)	1	2
Flowrate (gal/min)	1000	—

Table 7.7: Feed constraints for wafer fabrication section

Property	Lower Bound	Upper Bound
pH	5.5	6.5
R ($k\Omega/cm$)	12000	18000
TOC (ppm)	2	5
Flowrate (gal/min)	1000	—

Table 7.8: Feed constraints for CMP section

the procedure outlined in section 5.4 for the data given in tables 7.7 and 7.8. The resulting ternary diagram is shown in figure 7.8, where the dotted lines represent the feasibility regions for the water feed streams.

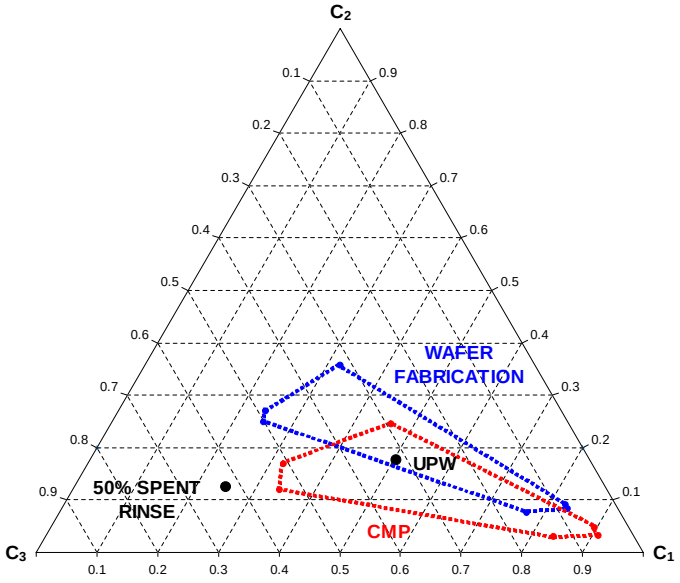

Figure 7.12: Ternary representation of microelectronics problem

From figure 7.8 it is readily seen that the UPW is very close to the boundary of the feasibility region of the wafer fabrication section, thus only a minimal benefit will be obtained by recycling the 50% Spent Rinse to this unit. The optimal mixing point would be right on the feasibility region boundary, and the relative arm for the UPW will be close to 1, thus the potential for recycle of the 50% Spent Rinse is negligible. However, there is a significant distance between the UPW point and the boundary of the feasibility region of the CMP section. Therefore it is worthwhile to investigate the potential for recycling the 50% Spent Rinse to the CMP section along with some UPW.

Lever-arm analysis dictates, that the optimal mixing point between UPW and the recycled 50% Spent Rinse is located on the boundary of the feasibility region for the CMP as indicated in figure 7.13. This mixture would consist of 71.5% UPW corresponding to 715 gpm, a reduction of 285 gpm from the base case design.

Unfortunately, as pointed out several times in this work, a match of the cluster values, is a necessary but NOT sufficient criteria for having matched the property constraints of a given sink. When evaluating the AUP values of the designed mixture and the feasibility region, the values do not match. The AUP value of the designed mixture is 13.60, while the AUP value for

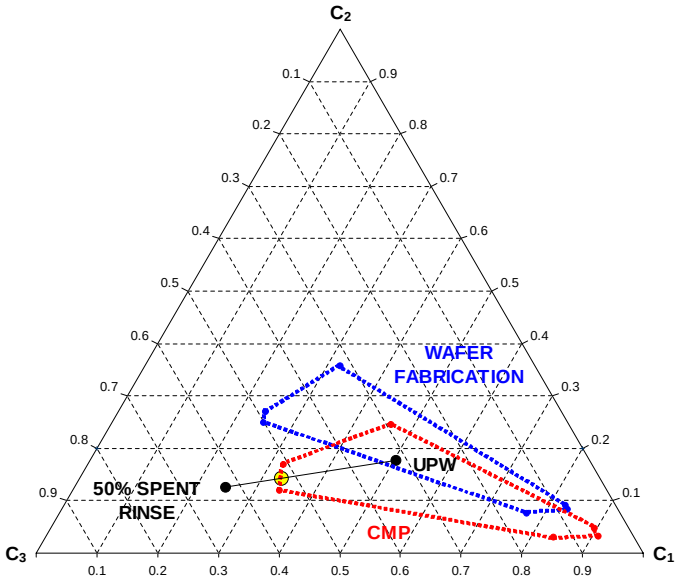

Figure 7.13: Optimal mixing point from visual analysis

the same cluster point on the feasibility region is 9.60, i.e. this mixing point is not feasible. However, since the optimal point has to be located on the straight line connecting the two streams and has to be contained within the feasibility region it is straightforward to move the mixing point closer towards the UPW point along the same line until the CMP constraints are matched. The resulting mixing point is depicted in figure 7.14, which corresponds to a CMP feed consisting of 83.3% UPW or 833 gpm, a reduction of 167 gpm from the base case design. The updated flowsheet along with the new flowrates is presented in figure 7.15. The lowered UPW requirements has lowered the municipal water requirement by 225 gpm, which on an annual basis, assuming 8000 hours of operation, corresponds to approximately 108 million gallons. Furthermore, due to the implementation of recycle, the annual load on the municipal waste water treatment system is reduced by approximately 80 million gallons. Finally, since the UPW producing facility has a capacity of 2700 gpm, the flexibility and reliability of this process is also increased, due to the reduced operating load.

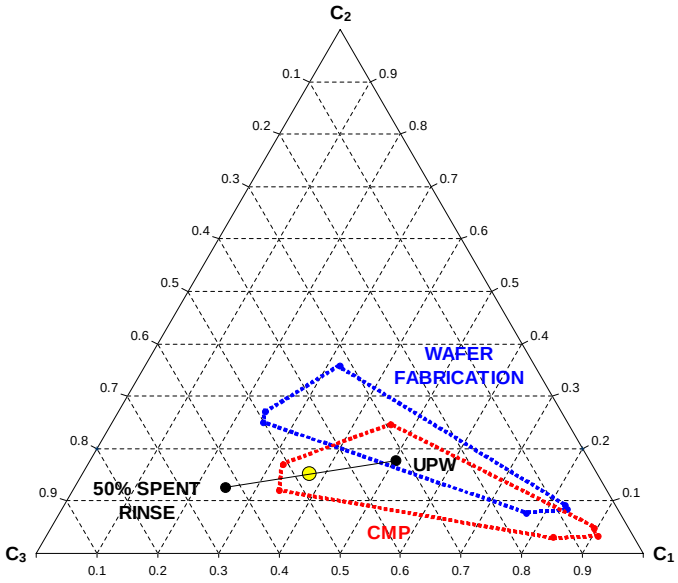

Figure 7.14: Optimal mixing point after AUP validation

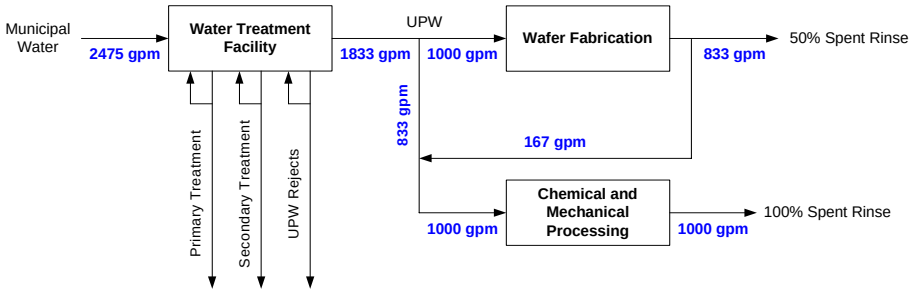

Figure 7.15: Schematic of microelectronics processing after property integration

7.7 Waste Minimization in Orange Juice Processing

In many product design problems, the quality constraints on the desired product are given not as absolute property values, but as a range of acceptable values. An example of such a problem is the production of orange juice, where a number of raw juice sources are mixed and diluted to produce several products. Figure 7.16 shows the schematic of a typical orange juice processing plant, where the fresh oranges, through several extraction and evaporation steps, are converted to three intermediate products or raw juices. These intermediate products designated 60 Brix juice, 65 Brix juice and pasteurized juice are then mixed and diluted to achieve the two products, i.e. unsweet pasteurized orange juice and frozen concentrated orange juice. The Brix value is named after Adolf F. Brix, who developed a method for measuring the sugar content of liquids in 1870. Measuring the Brix value of fruit juices provides a measure of the ripeness of the fruit, which is also used in quality assurance regulations, e.g. within the European Union.

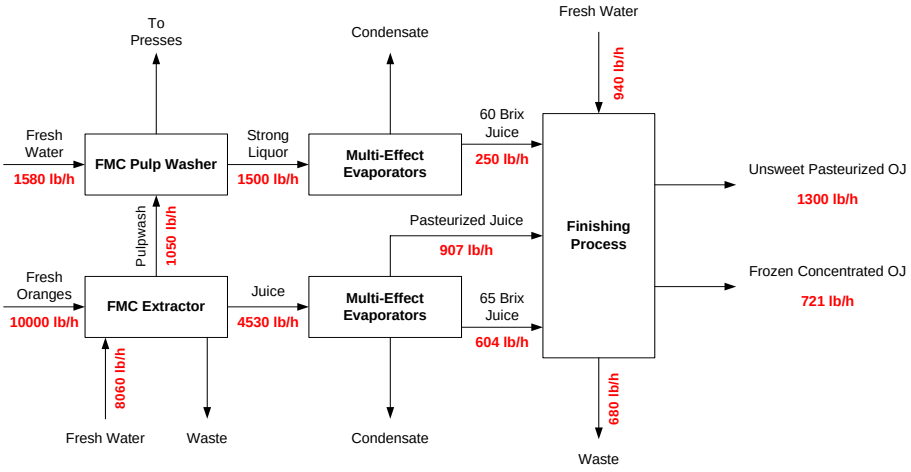

Figure 7.16: Schematic representation of orange juice manufacturing

The objective of this case study is to identify the optimum allocation of the process resources in order to achieve maximum utilization, i.e. minimum waste production. Three primary properties are used to characterize the raw orange juice sources (Gabriel *et al.*, 2003b):

- **Degrees Brix** (B) - total soluble solids determined by Brix hydrometer.
- **Brix to Acid Ratio** (BAR) - degrees Brix to anhydrous citric acid amount.

- **Pulp (P)** - amount of pulp in the juice (percent).

According to Gabriel *et al.* (2003b), the property operator mixing rules can be described as presented in equations 7.10-7.12 along with the chosen reference values. The property data for the juice sources are given in table 7.9, while the property constraints of the two products, i.e. unsweet pasteurized orange juice and frozen concentrated orange juice, are presented in tables 7.10 and 7.11, respectively.

$$B_M = \sum_{s=1}^{N_s} x_s \cdot B_s \qquad , B^{ref} = 5 \tag{7.10}$$

$$BAR_M = \sum_{s=1}^{N_s} x_s \cdot \frac{B_s}{BAR_s} \qquad , BAR^{ref} = 0.2 \tag{7.11}$$

$$P_M = \sum_{s=1}^{N_s} x_s \cdot P_s \qquad , P^{ref} = 1 \tag{7.12}$$

Property	Pasteurized Juice	60 Brix Juice	65 Brix Juice
B	11.95	60	65
BAR	14.18	20	10
P	11.9	4.4	7
Flowrate (lb/hr)	907	250	604

Table 7.9: Properties of juice sources

Property	Lower Bound	Upper Bound
B	11	16
BAR	11.5	18.0
P	5	10

Table 7.10: Product quality constraints for unsweet pasteurized orange juice

Property	Lower Bound	Upper Bound
B	41.8	46.8
BAR	11.5	49.5
P	8	12

Table 7.11: Product quality constraints for frozen concentrated orange juice

The process outlined in figure 7.16 does not give any information about how the intermediates are distributed to the products, but since the objective of this case study is to identify the optimum allocation of the process resources the problem can be represented as outlined in figure 7.17.

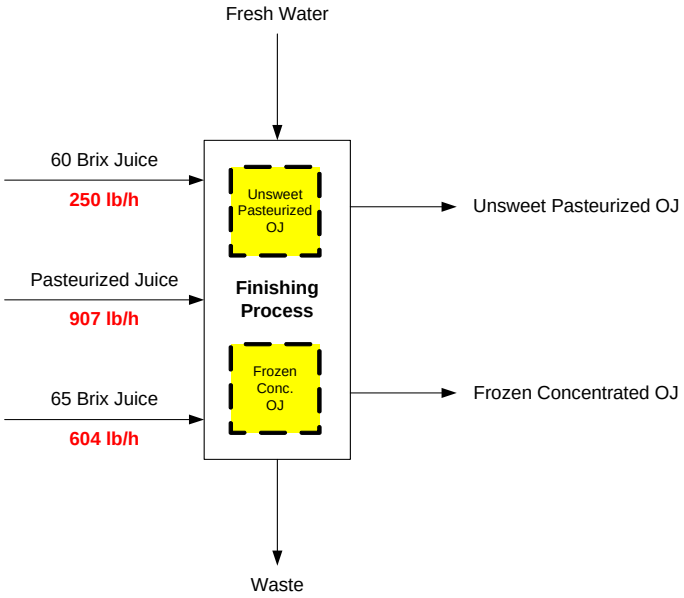

Figure 7.17: Overall juice blending problem in the finishing process

The problem is visualized by converting the property values to cluster values using equations 5.1 and 5.4-5.6. The product quality constraints are represented as feasibility regions, which are identified according to the procedure outlined in section 5.4 for the data given in tables 7.10 and 7.11. The resulting ternary diagram is shown in figure 7.18, where the dotted lines represent the feasibility regions for the two desired products.

It should be noted, that the mixing rule for the Brix to acid ratio is a function of two of the targeted properties as shown in equation 7.11. Therefore it is necessary to ensure that the calculation of the minimum and maximum values of the dimensionless operator values for use in the procedure in section 5.4 is correct, i.e. includes all combinations of the property constraint values, in order to obtain the true feasibility region.

According to Mims *et al.* (2000), the pasteurized orange juice intermediate has the lowest economic value, due to its lower concentration of juice. The reasoning behind this is that often the dilution of the concentrate is performed off-site, therefore lower concentration products will have a relatively higher transportation cost and thus also production cost even though the sales price of the diluted product is unchanged. Based on these considerations the optimum

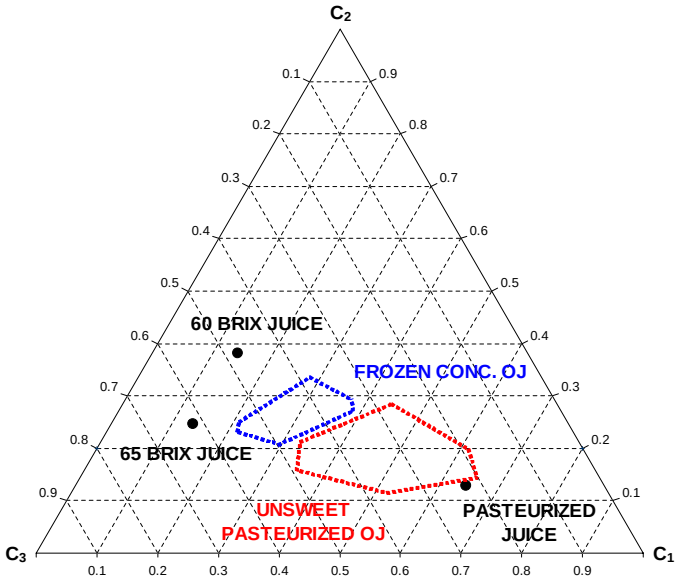

Figure 7.18: Ternary representation of orange juice blending problem

usage of the intermediate juice sources is to maximize the use of 60 Brix juice and 65 Brix juice. Applying lever-arm analysis, the resulting optimum mixing points, producing unsweet pasteurized orange juice and frozen concentrated orange juice, are presented in figures 7.19 and 7.20, respectively. Therefore the overall allocation problem to be solved can be represented as given in figure 7.21, where both optimum mixing solutions are included.

It is apparent from figures 7.19-7.21, that only two unique mixtures of 60 Brix juice and 65 Brix juice are viable candidates for mixing with the pasteurized orange juice in order to achieve the target. These points can be evaluated by comparing the slopes of the lines from the target to the pasteurized orange juice and from the mixture of 60 Brix juice and 65 Brix juice to the target. Using the Cartesian coordinates for each point the slopes are easily evaluated and it is possible to solve for the mixture ratios that result in matching slopes. It is also possible to solve for the mixture ratios graphically if the resolution of the plot is sufficiently high. By measuring or calculating the relative cluster arms, β_s, e.g. using equations 5.29 and 5.30, one can solve for the mixture ratios by rearranging equation 5.14 as shown in equation 7.13. Since the AUP value of the mixture is the same, regardless of which arm is evaluated, it is possible to calculate the mixture ratios as given by equation 7.14, which is obtained by equating the two expressions for AUP_M and rearranging to solve for x_s. The resulting mixtures of 60 Brix juice and 65 Brix juice for each of the products are given in table 7.12.

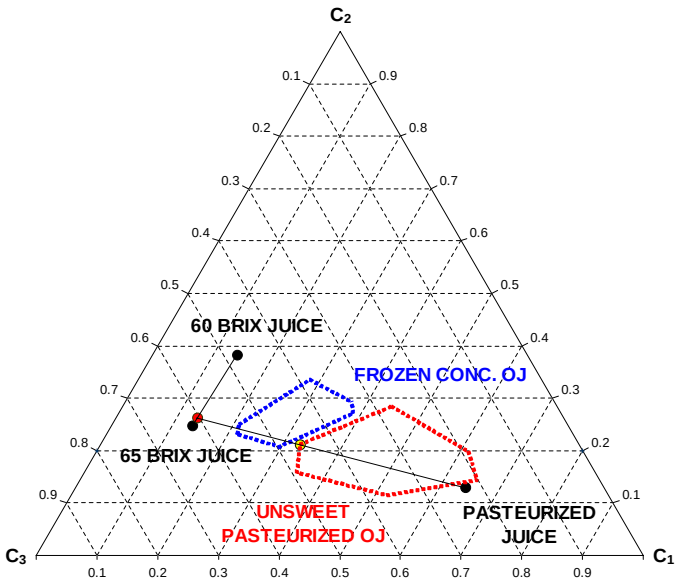

Figure 7.19: Optimal mixing point for unsweet pasteurized orange juice product

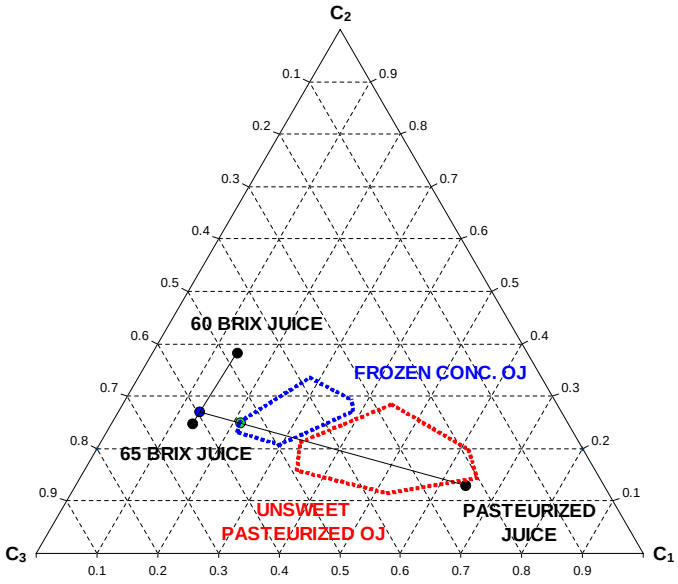

Figure 7.20: Optimal mixing point for frozen concentrated orange juice product

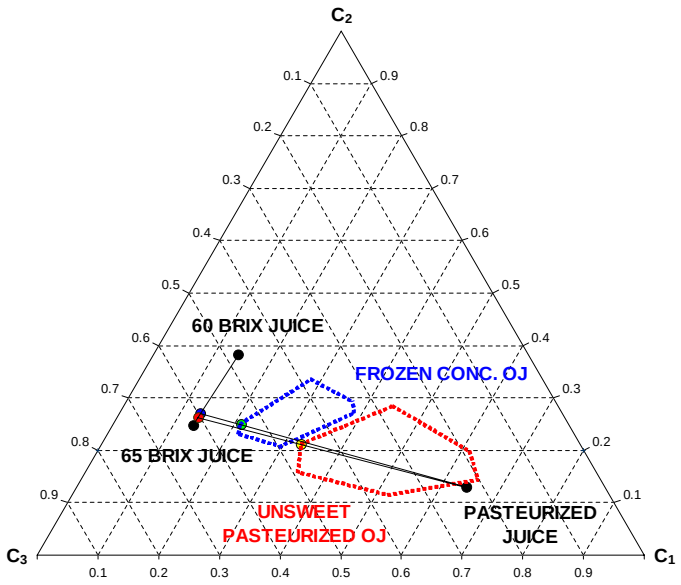

Figure 7.21: Optimal mixing points for both orange juice products

Analogous to the identification of the mixtures given in table 7.12, the relative contributions of these mixtures and the pasteurized orange juice to the mixture, which results in the optimal mixing points outlined in figures 7.19-7.21, can be identified. The mixtures of 60 Brix juice and 65 Brix juice are designated 60-65 Brix juice and the resulting mixtures are presented in table 7.13

$$AUP_M = \frac{x_s \cdot AUP_s}{\beta_s} \tag{7.13}$$

$$x_1 = \frac{AUP_2}{\frac{\beta_2}{\beta_1} \cdot AUP_1 + AUP_2} \tag{7.14}$$

Product	Fraction 60 Brix	Fraction 65 Brix
Unsweet Pasteurized OJ	0.1707	0.8293
Frozen Concentrated OJ	0.2470	0.7530

Table 7.12: Candidate mixtures of 60 Brix juice and 65 Brix juice

Once the optimal mixture has been obtained, fresh water is added in order to satisfy the property constraints. The dilution process does not change the location of the mixing point, i.e. the cluster values are unchanged, however the AUP value is altered to match the property constraints.

Product	Fraction 60-65 Brix	Fraction Pasteurized
Unsweet Pasteurized OJ	0.377	0.623
Frozen Concentrated OJ	0.685	0.315

Table 7.13: Candidate mixtures of 60-65 Brix juice and pasteurized juice

Based on the values given in tables 7.12 and 7.13 along with the data given in table 7.9 the AUP value of the mixture prior to dilution can be calculated using equation 5.17 as shown in equation 7.15. Once the AUP value of the mixture has been identified, the fractional contribution of the juice mixture to the diluted juice product can be calculated as presented in equation 7.16. This calculation utilizes the fact that water will have an AUP value of zero as it has zero values for all the targeted properties.

$$AUP_M^{Calc} = x_{60-65} \cdot AUP_{60-65} + x_{Past.OJ} \cdot AUP_{Past.OJ} \tag{7.15}$$

$$AUP_M^{Target} = x_{Mixture} \cdot AUP_M^{Calc} + x_{Water} \cdot AUP_{Water} \tag{7.16}$$

Using the data obtained from the visual analysis along with equations 7.15 and 7.16, a mathematical optimization problem can be formulated. It is important to emphasize that the optimization problem is formulated solely from initial data and values obtained from the visualization, i.e. the visualization provided the insights required to reduce the search space and therefore resulted in a well-defined optimization problem. Furthermore it should be noted that the equations and constraints used in the formulation are unchanged regardless of whether the objective function is defined as the minimization of waste or the maximization of profit. The formulated optimization problems are presented in Appendices A and C for the waste minimization and profit maximization cases, respectively. A commercially available solver package, LINGO 8.0 (LINDO, 2003), was used to solve the problems and the outputs from the solver are presented in Appendices A and C for the waste minimization and profit maximization cases, respectively. For the case of profit maximization, the sales price of the two juice products were specified as 0.60 $/lb and 0.33 $/lb for unsweet pasteurized orange juice and frozen concentrated orange juice, respectively (Mims *et al.*, 2000). The results obtained by solving the optimization problems are summarized in table 7.14. Based on the information reported by Mims *et al.* (2000), the cost of the unused intermediate juice sources are estimated as 0.08 $/lb of pasteurized orange juice, 0.15 $/lb of 60 Brix orange juice and 0.20 $/lb of 65 Brix orange juice. Assuming, that the feedstock of 10000 lb/hr of fresh oranges is unchanged, the profit is calculated as the difference between the value of the obtained products and the value of the waste streams. The waste minimization case results are illustrated in figures 7.22 and 7.23, while the profit maximization results are presented in 7.24 and 7.25.

A few key observations can be inferred from the optimization results. First of all, in the profit maximation case only the high value product, i.e unsweet pasteurized orange juice is produced, while in the waste minimization case frozen concentrated orange juice is produced as well. Furthermore, the profit maximation solution results in an added profit of 72 \$/hr compared to the minimum waste solution, which corresponds an increase of approximately 4.5%. This small increase, however is achieved at the expense of an increased fresh water consumption of almost 200 lb/hr along with an added waste generation of roughly 205 lb/hr.

The cost of the waste is calculated based on the value of the unused intermediates, however these estimates are actually too low. The reason is that if the unused intermediates are truly considered as waste, a disposal cost would have to be included and if the unused intermediate juices are to be stored for later use, an additional cost related to the transportation and storage should be added to the value. Similarly, large variations in the cost of the fresh water used in the dilution process, e.g. due to environmental regulations or changes in taxations, will also affect the design.

	Base Case	Min. Waste	Max. Profit
Water Usage (lb/hr)	940.0	1222.3	1422.2
Juice Production (lb/hr)	2021.0	2883.9	2878.0
Waste Generation (lb/hr)	680.0	99.4	305.1
Profit (\$/hr)	—	1601.6	1673.6

Table 7.14: Summary of optimization results for orange juice production

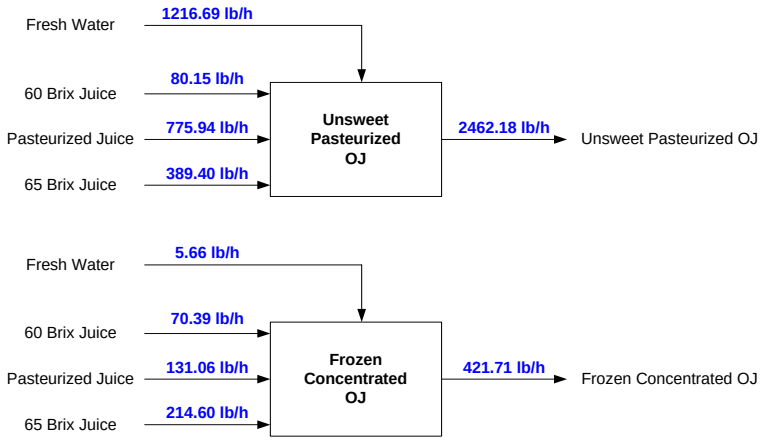

Figure 7.22: Optimal allocation of resources achieving minimum waste

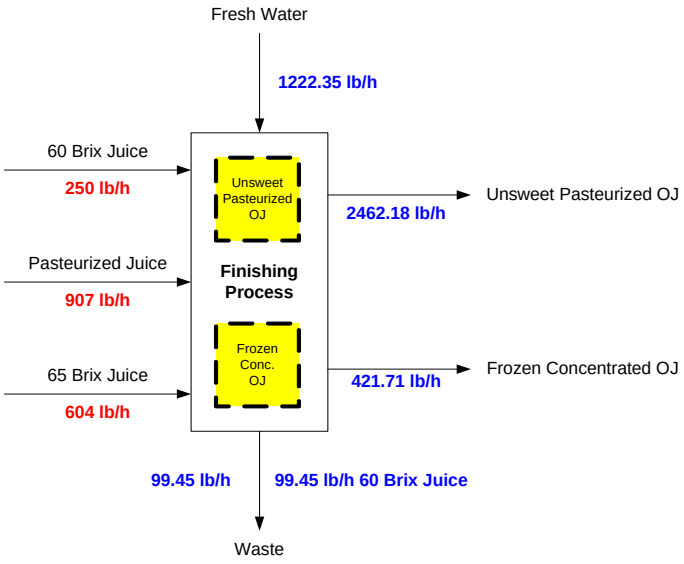

Figure 7.23: Overall minimum waste solution

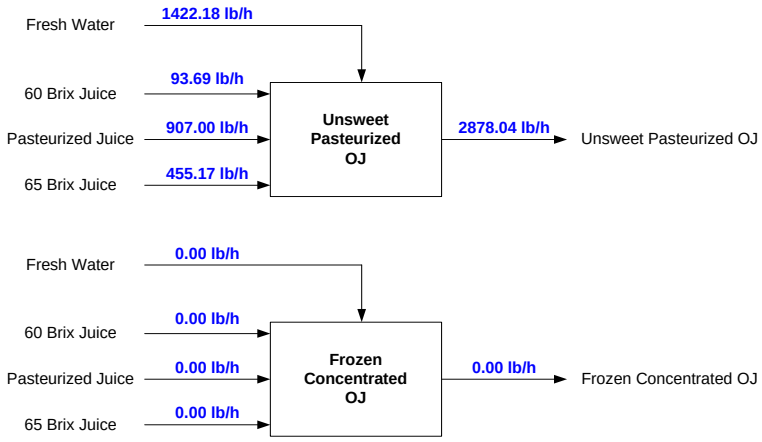

Figure 7.24: Optimal allocation of resources achieving maximum profit

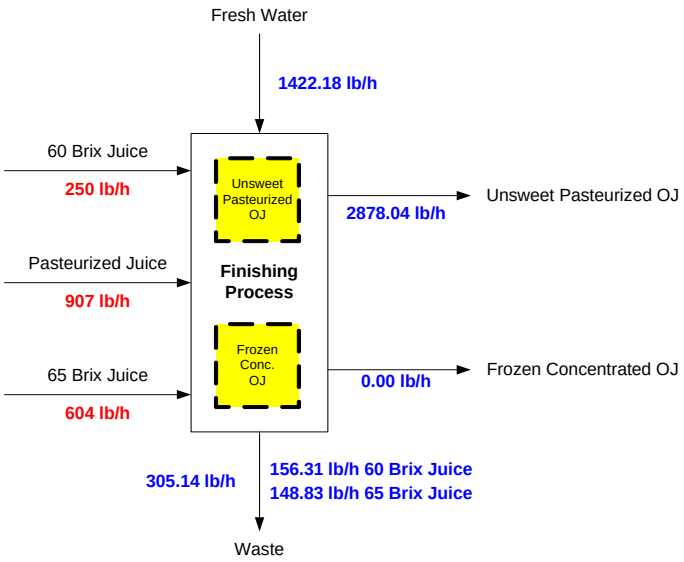

Figure 7.25: Overall maximum profit solution

7.8 Summary

In this chapter, the application of property based design techniques has been illustrated through a series of case studies. The examples highlight the use of the property clustering methodology in a reverse problem formulation environment and shows how the developed methods can be used for solving process design problems, where the design objectives are described by properties rather than chemical constituency. Visualization of the problems allows for easy identification of optimum recycle strategies using optimization rules based on lever-arm analysis.

In the metal degreasing example, the process design problem is solved to obtain the operating conditions of the condensation unit along with the properties of the required external solvent. The design does not commit to any components until the end, and thus the problem is solved by tracking the properties resulting in optimum performance. This approach exploits the benefits of the interface between process and product design problems by only targeting the properties required to achieve the desired performance. In the papermaking example the focus is on resource conservation, and the example highlights the use of reverse problem formulations for the identification of optimal direct recycling strategies and also for targeting the required interception that would enable maximum recycle. The water conservation problem highlights the feed feasibility conditions outlined in chapter 5, and shows how to proceed if the validation of the AUP values fails. Finally, the orange juice manufacturing example highlights how the product design, i.e. the formulation of the juice

products, can determine the process design by identifying the optimal allocation of the resources and also shows how visualization of the problem can assist in the formulation of well-defined optimization problems.

It should be emphasized, that conventional composition and component based methods for process design would not be feasible for handling the problems presented in this chapter. First of all, the process objectives and constraints were given in terms of physical properties, which conventional methods fail to adequately integrate in the design procedures. Furthermore, some of the properties, e.g. in the papermaking example, can not be described by composition alone. Property clustering provides the systematic framework required for handling such property driven problems. Using the developed stepwise procedures and visualization tools, the problems can be solved in a consistent manner.

8

PROPERTY BASED PRODUCT DESIGN

8.1 Introduction

Designing compounds and mixtures with specific physical and chemical properties are typical product design problems faced in chemical engineering. Since the properties of the compound or mixture determine whether or not the design is useful, the basis for solution approaches in this area should be the properties themselves. Computer Aided Molecular Design (CAMD) techniques have been successfully applied to design of compounds and mixtures with desired properties (Gani *et al.*, 1991; Odele and Macchietto, 1993; Marcoulaki and Kokossis, 1998; Harper and Gani, 2001). These techniques are based on identification of molecular building blocks resulting in chemical structures possessing the desired properties. In the following, a novel method for the synthesis of mixtures is presented. The methodology is meant as a supplement to the CAMD techniques, and focuses on the formulation of mixtures and blends from a known set of candidate constituents (Eden *et al.*, 2003c).

8.2 Synthesis and Design of Formulations

Formulation, the mixing of materials to achieve a new or improved product, is practised in many different industries, including paints and dyes, foods, personal care, detergents, plastics, and pharmaceuticals. Often the formulations are selected based on qualitative engineering knowledge and/or experience; however the effectiveness of such an approach is determined by the available data and absence of bias towards specific solutions. Anderson and Whitcomb (1999) have presented a methodology for formulation design using Design Of Experiments (DOE) software, which is based on rigorous optimization routines. When applied to the synthesis of a three component mixture problem, the data can be visualized using ternary diagram, where each of the vertices represents the pure components or mixtures to be used in the formulation. As functions of composition the property values are plotted as contour plots and for each desired property of the formulation a new ternary diagram must

be generated. Furthermore, this has to be repeated for all candidate ternary mixtures attainable from the pure component set. This approach will quickly suffer combinatorial problems if the number of candidate components/mixtures increases. Thus there is a need for fast, reliable and systematic screening methods capable of identifying candidate formulations without suffering from combinatorial explosion. A framework capable of handling multiple candidate components/mixtures is needed to synthesize promising formulations.

8.3 Visualization and Solution Strategy

Given a set of candidate compounds (constituents) NC, that are mutually totally miscible, determine a set of mixtures with 2 to NC compounds that, in addition to being totally miscible with a specified product, also satisfy a set of property targets given in equation 8.1. For visualization purposes, the number of properties (or property functions) NP is set to 3, however the method presented here is not limited to 3 properties, but for higher order systems mathematical optimization is required.

$$P_{jM}^{\min} \leq P_{jM} \leq P_{jM}^{\max} \qquad j = 1, 2, ..., NP \tag{8.1}$$

The problem is visualized by employing the procedure for conversion to property clusters outlined in chapter 5 and the strategy for identification of candidate formulations is presented in table 8.1. It should be emphasized, that the stepwise procedure outlined in table 8.1 is for the case, where the property targets to be matched are specific values. If the product constraints are given as intervals as in equation 8.1, then step 1 is replaced by the procedure for identification of the feasibility region boundary presented in section 5.4.

Step	Description	Equation
1	Calculate dimensionless target property values	5.4
2	Calculate target AUP index	5.5
3	Calculate ternary cluster values for the target	5.6
4	Plot the target point on the ternary cluster diagram	—
5	Calculate dimensionless property values of candidates	5.4
6	Calculate the AUP index for each candidate	5.5
7	Calculate ternary cluster values for each candidate	5.6
8	Plot the candidates on the ternary cluster diagram	—
9	Visually identify feasible mixtures (binary, ternary etc.)	—
10	Validate the candidate mixtures	—

Table 8.1: Stepwise procedure for identification of candidate formulations

It should be noted that in step 9, it is possible to visually identify even multicomponent mixtures, however for such mixtures a mathematical optimization

procedure is preferable. The solution of such an optimization problem is very simple as all mixing operations are described by linear models (Shelley and El-Halwagi, 2000; Eden *et al.*, 2002a). Furthermore, it should be emphasized that the validation step originates from rule 5.1 not including the flowrate condition (at least initially). The procedures for validating candidate binary and ternary mixture formulations are presented as parts of the case study below.

8.4 Formulation of Mixtures

To illustrate the use of the property cluster based formulation design method, a simple visual "mixture" design is presented. The objective of this investigation is to identify binary and ternary mixtures of pure components resulting in a target mixture with a solubility parameter of approximately 18.5 $MPa^{0.5}$ and properties defined by the values given in table 8.2.

Property	Operator	Reference Value	Target Value
Density (g/cm^3)	$\psi(\rho_s) = \rho_s^{-1}$	0.800	0.867
Boiling Point (K)	$\psi(T_{boil,s}) = T_{boil,s}$	300	383.75
Melting Point (K)	$\psi(T_{melt,s}) = T_{melt,s}$	150	178.25

Table 8.2: Property targets and characterization variables for mixture formulation

In order to identify candidate components that would be feasible constituents matching the target formulation, a database search was performed using the CAPEC database, which consists of experimental data for over 20,000 compounds (CAPEC, 2003). The search space was limited to compounds with solubility parameters between 18.0 and 19.0 $MPa^{0.5}$, thereby ensuring mutual miscibility of all the compounds found. The search yielded many compounds and the 25 compounds with solubility parameters closest to 18.5 $MPa^{0.5}$ have been considered. For illustrative purposes, however only the 6 components given in table 8.3 are highlighted in this study. The components were selected solely based on their relative placement on the ternary diagram to ensure a reasonable spread in order to illustrate the visual formulation synthesis methodology.

By employing the procedure presented in table 8.1, the property data given in tables 8.2 and 8.3 is converted to property cluster values for ternary representation. The resulting diagram is given in figure 8.1, where the individual components are denoted by shaded dots along with the component ID, while target is shown as a white dot. By drawing straight lines between the pure component points through the target cluster, the candidate formulations can be identified. The feasible binary and ternary candidate formulations are shown in figures 8.2 and 8.3. Again, it must be emphasized that matching the cluster target is a necessary but NOT sufficient criterion for matching the property

ID	Name	Density	Boiling Point	Melting Point
1	Ethyl-mercaptan	0.839 g/cm^3	308.25 K	125.35 K
2	Dimethyl-sulfide	0.848 g/cm^3	310.45 K	174.85 K
3	Benzene	0.879 g/cm^3	353.15 K	278.65 K
4	m-Xylene	0.868 g/cm^3	412.25 K	225.35 K
5	Diethyl-disulfide	0.990 g/cm^3	427.25 K	171.65 K
6	Cyclohexyl-benzene	0.950 g/cm^3	513.25 K	280.45 K

Table 8.3: Candidate formulation constituents and pure component property values

targets. The AUP values of the formulated mixture and the target must also match in order to match the property targets. In the following, the validation procedures for binary and ternary mixture formulations are presented. The procedures can be extended to quaternary and multi-component formulations using visual as well as mathematical optimization techniques. The first step is visual identification of the candidate constituents for e.g. a quaternary mixture, thus reducing the search space by screening out inherently infeasible combinations. Next the feasible mixture compositions are identified by evaluating the AUP values of the formulated mixtures and finally a performance criterion is added to identify the optimal mixture.

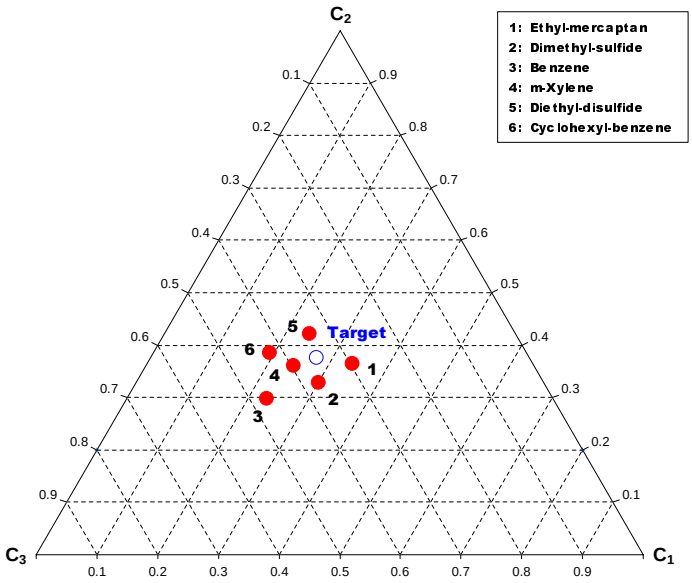

Figure 8.1: Ternary visualization of formulation synthesis problem

8.4.1 Validation Procedure for Binary Mixtures

From figure 8.2 it can be noted that only one binary mixture (components 1 and 6, Ethyl-mercaptan and Cyclohexyl-benxene respectively) is able to match the desired target formulation. The procedure for verifying that a binary formulation (A-B) satisfies the property targets as well as the cluster targets is given in table 8.4.

Step	Description	Equation
1	Calculate relative cluster arms β_s for A and B	5.31,5.32
2	Calculate fractional contributions x_s for A and B	5.14
3	Calculate the A-B mixture AUP value	5.17
4	Compare mixture AUP with target AUP	—

Table 8.4: Validation of candidate binary formulations

The final step in table 8.4 is the actual validation step. If the two AUP values are equal then the feasibility of the designed formulation is validated and the property targets have been matched. The AUP value for the formulated mixture is found to be 3.38, while the AUP value for the target is 3.39. This slight discrepancy is within the accuracy that can be expected by graphical lever-arm analysis. When the clustering solution is converted back to physical property values the estimated mixture properties are within $\pm 1.5\%$ of the target values. The binary solvent mixture consists of 65% Ethyl-mercaptan and 35% Cyclohexyl-benzene. Furthermore, the method also provides a quick screening of candidates, including e.g. implicitly infeasible binary pairs such as: (6-5), (6-4), (6-2), (6-3), (5-4), (5-3), (4-2), (4-3), (5-1), (4-1), (3-1), and (2-1).

8.4.2 Validation Procedure for Ternary Mixtures

In figure 8.3 all the candidate ternary mixtures have been identified. The mixture synthesized in figure 8.4 is denoted 1-5-3, i.e. a mixture of components 1 and 5 is mixed with component 3 to achieve the target. The procedure for verifying that a ternary formulation denoted A-B-C satisfies the property targets as well as the cluster targets is given in table 8.5. Note that in the following, the intermediate mixture between A and B is denoted AB. The final step in table 8.5 is the actual validation step. If the two AUP values are equal then the feasibility of the designed formulation is validated and the property targets have been matched.

For the ternary mixture 1-5-3 the AUP value calculated in step 3 (corresponding to the required mixing AUP for the AB mixture to achieve the formulation target) is 3.27, while the AUP value calculated in step 6 is 3.10. The difference is too large to be attributed to reading lever-arm values from the plot, i.e. this ternary mixture is NOT feasible for matching the property targets even though the cluster targets were matched. Once again, it should be emphasized that matching the cluster targets is a necessary but NOT sufficient criterion for

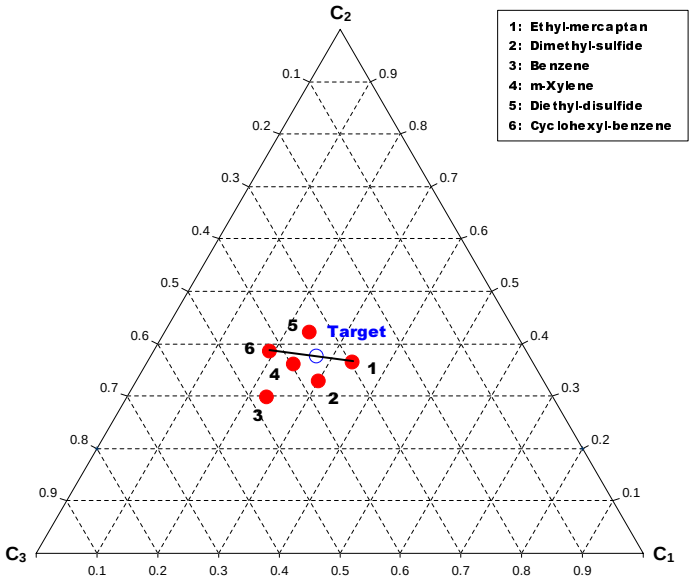

Figure 8.2: Candidate binary mixtures

Step	Description	Equation
1	Calculate relative cluster arms β_s for AB and C	5.31,5.32
2	Calculate fractional contributions x_s for AB and C	5.14
3	Calculate the AB mixture AUP value	5.17
4	Calculate relative cluster arms β_s for A and B	5.31,5.32
5	Calculate fractional contributions x_s for A and B	5.14
6	Calculate the AB mixture AUP value	5.17
4	Compare AUP from step 3 with AUP from step 6	—

Table 8.5: Validation of candidate ternary formulations

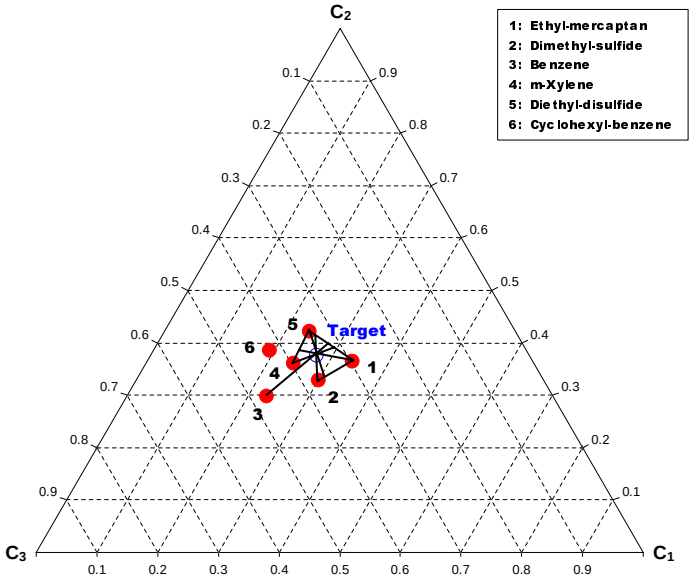

Figure 8.3: Candidate ternary mixtures

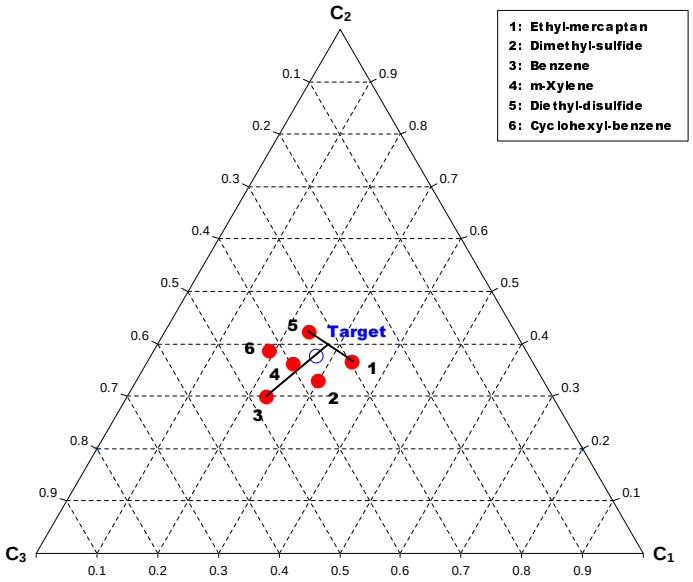

Figure 8.4: Naming convention for ternary mixtures

matching the property targets. This analysis is performed on all the identified ternary mixtures and the results are summarized in table 8.6.

Mixture (A-B-C)	Feasible	A	B	C
1-5-3	No	—	—	—
1-5-4	Yes	37%	20%	43%
1-5-2	No	—	—	—
1-2-5	No	—	—	—
4-5-1	No	—	—	—

Table 8.6: Candidate ternary mixtures and fractional contributions of the constituents to the final mixture

It should be noted that when the clustering solution for the ternary mixture of Ethyl-mercaptan, Diethyl-disulfide and m-Xylene (1-5-4) is converted back to physical property values, the estimated mixture properties are within ±2% of the target property values.

If the same investigation should have been performed using the design of experiments approach, then the number of contour plots to be generated can be calculated from equation 8.2:

$$N_{Plots} = NP \cdot \frac{NC!}{NP! \cdot (NC - NP)!} \tag{8.2}$$

For the simple problem investigated here with 3 targeted properties and 6 candidate constituents, the number of plots to be generated is 60, which then would have to be evaluated simultaneously to identify sweetspots within the search space. In table 8.7 the rapidly increasing number of plots to be generated is presented for systems with 3 targeted properties (for visualization purposes) and increasing number of candidate constituents.

Number of constituents	Number of contour plots
3	3
4	12
5	30
6	60
7	105
10	360
15	1365
20	3420

Table 8.7: Required number of contour plots in DOE methodology

It is apparent, that even for this very simple problem, the number of plots to be evaluated makes the DOE approach impractical for synthesis purposes.

It should be emphasized that design of experiments is a very effective tool to investigate the optimum blend of the selected constituents, however as a visual screening tool for identification of the constituents, the method is not practical, due to the number of plots that need to evaluated. Using property clustering techniques, the problem can be visualized in a single plot regardless of the number of constituents, thus providing a fast and easy screening tool for identification of promising candidates as well as inherently infeasible formulations. Once promising formulations have been identified, then DOE methods may be applied to utilize their advanced statistical and optimization methods to identify the optimum blending ratios. The initial screening performed using clustering techniques, reduces the number of candidate formulations, thereby reducing the overall computational intensiveness of solving the formulation problem.

8.5 Formulation of Polymer Blends

Selection of polymers for specific tasks is often performed by trial and error and commonly relies on prior engineering experience or on database searches. Employing systematic methods for the preliminary screening and selection of candidate polymers that possess the desired properties is therefore desirable. By combining molecular design techniques and mathematical programming a framework for designing polymers with specific properties has been developed (Vaidyanathan and El-Halwagi, 1994, 1996; Vaidyanathan *et al.*, 1998).

Mixing two or more polymers together to produce a blend or alloy is a well-established strategy for achieving a specified set of physical properties without having to resort to specialized polymer systems. From the design community the focus has been to develop empirical methodologies and tools that can reduce the number of laboratory trials by estimating the properties of the polymer blend. An overview of the different techniques used in producing polymer blends as well as rheological aspects of the blends is given by Folkes and Hope (1993).

When designing polymer blends a number of decisions have to made by the designer, most important are the selection of appropriate polymers to be blended and the optimum blending ratio. It is imperative that the polymers that constitute the blend are compatible, i.e. miscible. Unfortunately, most polymers are incompatible when mixed in their unmodified form, although some empirical data has been published describing feasible binary blends, e.g. Cowie (1991). The identification of possibly miscible blends is based primarily on specific intermolecular interactions, such as hydrogen bonds, dipole-dipole and ion-dipole interactions. It is assumed that certain groups when incorporated in polymer chains will exhibit these intermolecular interactions and thus enhance miscibility.

In this work, however it is assumed that polymers to be used in the formulation are miscible or can be made miscible by the addition of an external agent, e.g. a solvent. The physical properties of a blend are functions of the charac-

teristics of the constituent polymers, i.e. they are dependent on the properties of the polymers to be blended. Several predictive models for the estimation of blend properties have been reported, e.g. Paul and Newman (1978) and Folkes and Hope (1993).

For any microscopically homogeneous system of polymers, a simple linear mixing rule can used to estimate the blend properties. Unfortunately polymer systems are rarely homogeneous, therefore more advanced models for the prediciton of the mechanical properties of polymer blends have been developed. In equation 8.3, M_M represents the mechanical property of the blend, while C_1, M_1, C_2, M_2 represent the concentration in the blend and mechanical properties of polymers 1 and 2, respectively. This expression calculates the upper bound of the mechanical property of the blend, which is called the Voigt bound, while the lower bound, called the Reuss bound, can be found using equation 8.4 (Vaidyanathan and El-Halwagi, 1996).

$$M_M^{upper} = C_1 \cdot M_1 + C_2 \cdot M_2 \tag{8.3}$$

$$M_M^{lower} = \frac{M_1 \cdot M_2}{C_1 \cdot M_2 + C_2 \cdot M_1} \tag{8.4}$$

The properties that can be estimated using these prediction methods are the bulk modulus (K) and the shear modulus (G). From these values the Young's modulus (E) and the Poisson ratio (ν) can be calculated for isotropic materials using equations 8.5 and 8.6, respectively.

$$E = \frac{9KG}{3K + G} \tag{8.5}$$

$$\nu = \frac{3K - 2G}{2\,(3K + G)} \tag{8.6}$$

The framework for synthesizing polymer blends using property clustering techniques is analogous to the general strategy presented in section 8.3. Along with the mechanical properties, three physical properties are often used for the characterization of polymers:

- **Glass transition temperature (T_g)**
- **Density (ρ)**
- **Molecular weight (MW)**

The property operator mixing rules for the targeted properties are presented in equations 8.7-8.9. It should be noted that for certain polymers the glass

transition temperature, T_g can be calculated from a simple linear mixing rule, but since these polymers are relatively few, the generalized nonlinear form of the mixing rule is used (Cowie, 1991).

$$\frac{1}{T_{gM}} = \sum_{s=1}^{N_s} x_s \cdot \frac{1}{T_{gs}} \tag{8.7}$$

$$\frac{1}{\rho_M} = \sum_{s=1}^{N_s} x_s \cdot \frac{1}{\rho_s} \tag{8.8}$$

$$MW_M = \sum_{s=1}^{N_s} x_s \cdot MW_s \tag{8.9}$$

Once property data for the polymer constituents have been obtained and the internal miscibility verified, e.g. using Brandrup and Immergut (1989), then the cluster-based formulation technique provides a fast and easy screening method for the identification of candidate blends, that should be selected for further investigation. It should be emphasized that the formulation design can be performed visually analogous to the mixture formulation example presented in section 8.4, however it also requires the same validation steps as outlined in tables 8.4 and 8.5 for binary and ternary blends, respectively.

8.6 Summary

In this chapter, the use of property clustering techniques for product design, in particular the synthesis and design of formulations, has been presented. In most product design problems, the desired objectives are described by property constraints rather than compositions or components. Therefore the property clustering methodology is ideally suited for the solution of such problems. Compared to conventional methods such as design of experiments (DOE), the methodology presented in this thesis does not suffer combinatorial problems as the number of candidate constituents increases. Converting the property data of the constituents along with the target property constraints to clusters enables the visualization of the problem. Irrespective of how many candidate constituents are considered, the problem can be visualized in a single plot. Since all mixing operations are straight lines within the ternary cluster diagram, the method provides a quick screening tool for identification of inherently infeasible pairings as well as promising formulations that should be investigated further. Stepwise procedures for the identification and validation of binary and ternary formulations have been presented, and although it is possible to extend the procedures to include quarternary and multicomponent formulations, it is not practical to solve such problems visually. For multicomponent formulations it

is more efficient to use mathematical programming methods, however the optimization problems are easily formulated and solved as the mixing operations are all described by linear equations.

It should be noted, that further development and application of the property based product design techniques are contingent upon the availability and/or development of appropriate property operator descriptions. However the framework outlined in this chapter is capable of handling any type of formulation problem for which such operators exist. If the targeted product can be described by three properties then the formulation can be performed visually, and if more properties are required to adequately describe the design objectives, mathematical optimization methods can be invoked.

9

CONCLUSIONS

9.1 Achievements

The main achievement of this work is the development of the general reverse problem formulation framework, which facilitates more efficient solution of process and product synthesis and design problems. The developed framework differs significantly from conventional design methods as it is not iterative nor is it based solely on mathematical optimization techniques. Investigating the different roles, which process and property models play in design revealed that using the property model in a solve role in addition to the traditional service and advice roles would be more efficient. By reformulating a conventional forward problem into two reverse problems, the iterative nature of the traditional design methods is relieved. The reformulation strategy is based on decoupling the constitutive equations from the balance and constraint equations. The first reverse problem solves the balance and constraint equations in terms of the constitutive variables thus providing the design targets. The second reverse problem solves the constitutive equations to identify the unit operations, operating conditions and/or components that match the design targets. It is important to emphasize that if the design targets are matched, then it is not necessary to solve the balance and constraint equations again. Solving for the constitutive variables directly also reduces model complexity as the constitutive equations often include composition dependent terms, which have to be included in every iteration when solving the conventional forward problem. In the reverse problem formulation framework, the composition dependent terms are systematically removed from the equations by substituting for the constitutive variables. Furthermore, since most of the nonlinearity of a system is related to the constitutive equations, the two-stage solution methodology of the reverse problem formulation framework simplifies the problem significantly, i.e. regardless of whether the property model is nonlinear, the balance and constraint equations are simply solved for the property values. Subsequently, the property model is solved to match the obtained property design targets. The reverse problem formulation framework constitutes a paradigm shift as it challenges the notion of design being inherently iterative.

Another key achievement of this work is the development of a systematic framework for solution of process and product design problems, where the

objectives are driven by properties or functionalities rather than chemical constituency. Conventional methods are not capable of handling such problems, as the design objectives can not be described by composition alone. The introduction of property clustering enables the systematic tracking of properties throughout the process by using property operator functions, which describe the physical properties. The clustering technique is based on the capability to describe the properties by linear operator mixing rules, where the operators themselves may be linear or nonlinear. Hence the nonlinearity of the property models is embedded within the operator description which then enables all the design calculations to be performed on linear equations. The nonlinear expressions are only used to back calculate the physical property values after the design has been identified. For visualization purposes the number of targeted properties are limited to three, thus allowing for a one-to-one mapping between the property operators and the clusters. However the developed framework is capable of handling as many properties as necessary, however for more than three, mathematical programming techniques are called for. The solution of such optimization problems is relatively simple as they consist of linear equations for the cluster based design calculations. The clusters are tailored to have the attractive features of intra-stream and inter-stream conservation, thereby enabling the development of consistent additive rules. The emphasis in this thesis has been on developing visual design techniques and optimization rules based on lever-arm analysis. Systematic procedures for converting the original property data to cluster values for the streams and units and back again have been developed. The mapping from properties to clusters is unique, while uniqueness of the reverse mapping from clusters to properties is ensured through the AUgmented Property index. A feasible match within the cluster domain is a necessary but not sufficient criterion for having identified a match in the property domain. The required conditions for ensuring that the solution also satisfies the original property constraints have been presented and discussed.

Application of the developed techniques to process design problems has been demonstrated through a series of case studies. The different case studies show the usefullness of the framework for solving property driven problems, but the individual cases highlight different aspects and benefits of the methodology. Primarily the examples have focused on visualization of debottlenecking and resource conservation problems and employing the reverse problem formulation solution strategy for identification of the optimum process. For example in the metal degreasing problem, the objective was to identify the optimum operating conditions of the condenser in order to facilitate the maximum recycle of the recovered organics while at the same time identifying the property targets for an external solvent. In the papermaking example, it was desired to achieve minimum usage of fresh fibers. As direct recycle strategies targeting maximum recycle of fibers were incapable of achieving the desired design targets, a second reverse problem was solved to identify the interception required in order to match the target. The problem of ensuring that a solution

identified in the cluster domain also satisfies the original property constraints was highlighted in the microelectronics manufacturing example. Finally, the orange juice manufacturing case study shows how visualization and lever-arm analysis can assist in the formulation of a well-defined optimization problem that targets the optimum allocation of the process resources.

Chemical product design is an area ideally suited for the developed framework, as the design objectives are often described by the desired properties of the product. Computer Aided Molecular Design (CAMD) techniques have proven very useful for the identification of compounds that possess certain properties, however in many cases the product is to be found as a mixture of several compounds from a list of candidate constituents. The design of experiments (DOE) approach, which is based on statistical methods, has been advocated as an efficient way of synthesizing promising formulations. Unfortunately, this approach quickly suffers combinatorial problems as the number of candidate constituents increases. Thus there is a need for fast, reliable and systematic screening methods capable of identifying candidate formulations and thereby reduce the number of subsequent laboratory trials. In this work, a new simple, yet effective, systematic method to synthesize and design formulations has been developed. For any formulation design problem, the target and the raw materials are identified on the ternary property cluster diagram. Since formulation design involves only mixing operations, the optimal formulation is easily determined together with all possible solutions by determining the mixing operations that will match the target. It should be noted that the design of the formulation and the simulation of the mixing operation are performed simultaneously. Since all mixing operations are straight lines within the ternary diagram it is possible to visually identify the binary, ternary and multi-component mixtures, which are capable of matching the desired target properties. Lever-arm analysis can be employed to optimize the fractional contributions from each stream in terms of e.g. cost, environmental impact or any other performance indicator. Thus business decision making can be facilitated, as the methodology allows for easy screening of alternatives and identification of candidate formulations that should be selected for further, more rigorous investigation including laboratory trials. The significance of this method is that irrespective of how many components and or mixtures are handled, the problem is solved visually on a ternary diagram. Also, the same method (but with different cluster properties) is applicable to wide range of problems, such as, solvent mixtures, oil blends, coatings for paints, additives for drugs and many more.

9.2 Challenges and Future work

The development and application of the general reverse problem formulation framework and property clustering techniques represent considerable advances in the state of the art in process and product design. However, both methodologies have only recently been introduced, i.e. the methods are still in their

infancy, thus several areas exist, where further work is required. The work presented in this thesis provides the foundation as well as the general methods, and in the following sections, an overview of the challenges is presented along with possible means of addressing these challenges, that were not resolved in this thesis.

9.2.1 Reverse Problem Formulations

As stated earlier in this thesis, the concept of reverse problem formulations constitutes a paradigm shift in process design, as it challenges the way property models have been used in the past. In this work, the constitutive variables are represented using property clustering techniques, but this is only one way of representing these variables. For applications, where the process objectives are not property driven, it may be beneficial to use a different representation. Research efforts should be devoted to the identification of characteristic constitutive variables for different unit operations and analogously property models capable of handling these specific variables should be identified. This would lead to the creation of a library of corresponding variables and prediction methods for the individual unit operations. Such a knowledge base would be essential for fast solution of future reverse problem formulations.

The reverse problem formulation technique can also be used for model reduction purposes as the decoupling of the constitutive variables reduces the model complexity. The size of the process model is defined primarily by the number of stages (or discretized point locations), the number of phases and the number of components present in the process being modeled. Most model reduction techniques try to rearrange the model equations into a smaller set, where the number of equations in the reduced set is related to the number of dominant Eigen Values of the process model. Another form of reduction is obtained by reducing the number of discretized points or number of stages. Reverse problem formulations could lead to a new technique for model reduction that is based on rearranging a part of the model representing the constitutive equations. The rearrangement of the constitutive equations would lead to the definition of a new set of intensive variables, where the component compositions are replaced by reduction parameters in the process model. Since the number of components dominates the size of the traditional model equations, a significant reduction of the model size could be obtained through this new technique. Some interesting properties of this new technique is that the model reduction should not introduce any approximations to the model, it should not change the physical location of the process variables and it would provide a visualization of the process and operation that otherwise would not be possible. Gani and Pistikopoulos (2002) presented an example of such a model reduction, by solving a distillation design problem based on constitutive variables, i.e. the parameters in a cubic equation of state, thereby removing the composition dependent terms and reducing the overall problem size. Similar approaches should be investigated further, e.g. by determination of suitable

reduction parameters for different property models. The principal requirement of the reduction parameters is the ability to systematically remove the composition dependent terms. If successful, the model reduction methodology could also lead to new advances in the area of process control. In the example presented by Gani and Pistikopoulos (2002), the composition dependent terms reduce to only two parameters in the constitutive model, i.e. for such a system, proven two-dimensional control structuring methods could be employed to identify the optimum control structure, even if the original system is of much higher order.

9.2.2 Property Clustering

The property clustering techniques presented in this work are based on the use of property operator functions, which exhibit linear mixing rules even if the operators themselves are nonlinear. As long as the properties can be adequately described by such linear expressions, all the presented design rules are valid, therefore considerable efforts should be devoted to the identification of property operator mixing rules for a wide range of physical properties in order to establish a knowledge base for future use. The identification of property operator mixing rules can be achieved through analysis of existing property models or by evaluating experimental mixture data. There is a large number of regression and parameter estimation methods available for fitting data to specific model structures. One promising technique is basis expansion, which has been used successfully in different model identification studies, primarily in signal analysis (Giannakis and Tepedelenlioglu, 1998). The basis expansion method is similar in nature to the property operator mixing rules, as it attempts to fit nonlinear data to a weighted linear expression of linear or nonlinear terms. Basis expansion may provide a systematic method of identifying the required mixing rules as it can be applied not only to experimental data, but also as a means of generating linearized approximations to nonlinear property models.

The ability to track properties or functionalities throughout a process is a powerful way to analyze design problems, where the desired objectives are specified in terms of properties. If appropriate property operator mixing rules can be identified, the property clustering methodology can also be used for inclusion of parameters that traditionally have been evaluated in the post-design phase. An example of such parameters is the environmental impact, which can often be related to the physical properties of the streams and components in the system. Estimation of the environmental impact for a given system can be achieved by use of the WAste Reduction (WAR) algorithm (Cabezas *et al.*, 1999; Young and Cabezas, 1999; Young *et al.*, 2000). The algorithm consists of a series of additive rules describing the impact parameters based on contributions from the individual constituents in the system. The calculations are fairly straightforward, however obtaining the individual contributions may require extensive experimental work. The US-EPA has collected a large amount of environmental impact data, which has been implemented in a database for

use in environmental studies. An implementation of similar strategies in the clustering framework would allow targeted design of environmentally benign and sustainable processes and products. Analogous to the WAR algorithm Heikkilä (1999) has presented an index-based method for the evaluation of process safety. By evaluating the components present in the system as well as the operating conditions, a qualitative measure of the process safety is obtained through a set of additive rules. It would be a significant advance if such strategies could be incorporated in the clustering framework. However, it should be emphasized that it most likely will require development of robust optimization based strategies, if environmental impact and process safety aspects should be included in the design considerations, as the increased number of variables to track no longer allows for visual solution of the problem.

In this work the principles of constitutive modeling have been introduced, but considerable work needs to be done in this area to increase the applicability, especially with respect to handling dynamic systems such as batch operations. The potential benefits are very large as the problem complexity could be reduced significantly by only evaluating the property changes in the system. Since the properties often can be readily measured, such a representation could generate insights that otherwise would be hidden. The development of dynamic clustering models would also allow for control structure design and selection in a low-dimensional domain even if the original system is quite complex.

Finally, the methodologies for synthesis and design of formulations, which currently are limited to binary and ternary mixtures and blends, should be extended to multi-component formulations. As it is difficult, if not impossible, to visually identify candidate formulations with more than three constituents, there is a need to develop systematic mathematical optimization strategies for the identification and validation of such mixtures. The general framework for synthesizing promising formulations also utilizes property operator descriptions, so any increase the application range is contingent on the development of the appropriate operator mixing rules. However, the methodology can be applied to a wide range of formulation problems, where it is difficult to describe the design objectives adequately using composition based methods. An example could be the formulation of fragrances, where the design objectives are often quite subtle and hard to describe by composition alone, e.g. the degree of citrus scent of a given parfume. Using available knowledge and experimental data it may be possible to derive an empirical expression for the citrus scent as a function of the contributions from the individual constituents.

Appendices

A

LINGO Input for Orange Juice Waste Minimization

The input file used for solving the mathematical optimization problem formulated in section 7.7 with the objective of minimizing the total waste from orange; juice production. The optimization problem was solved using LINGO 8.0, LINDO Systems Inc, Chicago, IL, USA.

```
!Objective Function;

model:

min = WASTE;


!Definition of flows for unsweet pasteurized OJ product;

X_60_uns = F_60_uns / (F_60_uns + F_65_uns);
X_65_uns = F_65_uns / (F_60_uns + F_65_uns);
X_6065_uns = F_6065_uns / (F_6065_uns + F_Past_uns);
X_Past_uns = F_Past_uns / (F_6065_uns + F_Past_uns);

F_6065_uns = F_60_uns + F_65_uns;
F_mix_uns = F_6065_uns + F_Past_uns;


!Calculation of dilution for unsweet pasteurized OJ product;

AUP_mix_calc_uns = X_6065_uns * AUP_6065_uns
                   + X_Past_uns * AUP_Past;

AUP_mix_true_uns = (1-X_water_uns) * AUP_mix_calc_uns
                   + X_water_uns * AUP_water;
```

```
X_water_uns = F_water_uns / (F_water_uns + F_mix_uns);
F_juice_uns = F_water_uns + F_mix_uns;


!Definition of flows for frozen concentrated OJ product;

X_60_froz = F_60_froz / (F_60_froz + F_65_froz);
X_65_froz = F_65_froz / (F_60_froz + F_65_froz);
X_6065_froz = F_6065_froz / (F_6065_froz + F_Past_froz);
X_Past_froz = F_Past_froz / (F_6065_froz + F_Past_froz);

F_6065_froz = F_60_froz + F_65_froz;
F_mix_froz = F_6065_froz + F_Past_froz;


!Calculation of dilution for frozen concentrated OJ product;

AUP_mix_calc_froz = X_6065_froz * AUP_6065_froz
                  + X_Past_froz * AUP_Past;

AUP_mix_true_froz = (1-X_water_froz) * AUP_mix_calc_froz
                  + X_water_froz * AUP_water;

X_water_froz = F_water_froz / (F_water_froz + F_mix_froz);
F_juice_froz = F_water_froz + F_mix_froz;


!Overall mass balance;

BRIX_60_LEFT = 250 - (F_60_uns + F_60_froz);
BRIX_65_LEFT = 604 - (F_65_uns + F_65_froz);
PAST_OJ_LEFT = 907 - (F_Past_uns + F_Past_froz);

PROFIT = Cost_uns*F_juice_uns + Cost_froz*F_juice_froz
       - Cost_60*BRIX_60_LEFT - Cost_65*BRIX_65_LEFT
       - Cost_Past*PAST_OJ_LEFT;

WASTE = BRIX_60_LEFT + BRIX_65_LEFT + PAST_OJ_LEFT;
TOTAL_JUICE = F_juice_froz + F_juice_uns;
TOTAL_WATER = F_water_froz + F_water_uns;
```

```
!Known Values;

X_60_uns = 0.1707;
X_65_uns = 0.8293;
X_60_froz = 0.247;
X_65_froz = 0.753;
X_6065_uns = 0.377;
X_6065_froz = 0.685;
AUP_6065_uns = 48.89823;
AUP_6065_froz = 47.2883;
AUP_mix_true_uns = 15.15652174;
AUP_mix_true_froz = 37.70782609;
AUP_water = 0;
AUP_Past = 18.50368124;
Cost_uns = 0.6;
Cost_froz = 0.33;
Cost_60 = 0.15;
Cost_65 = 0.20;
Cost_Past = 0.08;


!Flowrate Constraints;

F_60_uns + F_60_froz <= 250;
F_60_uns + F_60_froz >= 0;
F_65_uns + F_65_froz <= 604;
F_65_uns + F_65_froz >= 0;
F_Past_uns + F_Past_froz <= 907;
F_Past_uns + F_Past_froz >= 1;
```

B

LINGO Output for Orange Juice Waste Minimization

The output file generated when solving the mathematical optimization problem formulated in section 7.7 with the objective of minimizing the total waste from orange juice production. The optimization problem was solved using LINGO 8.0, LINDO Systems Inc, Chicago, IL, USA.

```
Local optimal solution found at iteration:            61
Objective value:                                99.45397

                     Variable           Value        Reduced Cost
                        WASTE        99.45397            0.000000
                      X_60_UNS       0.1707000           0.000000
                      F_60_UNS        80.15257           0.000000
                      F_65_UNS        389.3997           0.000000
                      X_65_UNS       0.8293000           0.000000
                    X_6065_UNS       0.3770000           0.000000
                    F_6065_UNS        469.5523           0.000000
                    F_PAST_UNS        775.9445           0.000000
                    X_PAST_UNS       0.6230000           0.000000
                     F_MIX_UNS        1245.497           0.000000
                AUP_MIX_CALC_UNS      29.96243           0.000000
                 AUP_6065_UNS         48.89823           0.000000
                     AUP_PAST         18.50368           0.000000
                AUP_MIX_TRUE_UNS      15.15652           0.000000
                  X_WATER_UNS       0.4941490           0.000000
                    AUP_WATER        0.000000           0.000000
                  F_WATER_UNS        1216.685           0.000000
                  F_JUICE_UNS        2462.181           0.000000
                    X_60_FROZ       0.2470000           0.000000
                    F_60_FROZ        70.39346           0.000000
```

F_65_FROZ	214.6003	0.000000
X_65_FROZ	0.7530000	0.000000
X_6065_FROZ	0.6850000	0.000000
F_6065_FROZ	284.9938	0.000000
F_PAST_FROZ	131.0555	0.000000
X_PAST_FROZ	0.3150000	0.000000
F_MIX_FROZ	416.0493	0.000000
AUP_MIX_CALC_FROZ	38.22115	0.000000
AUP_6065_FROZ	47.28830	0.000000
AUP_MIX_TRUE_FROZ	37.70783	0.000000
X_WATER_FROZ	0.1343024E-01	0.000000
F_WATER_FROZ	5.663705	0.000000
F_JUICE_FROZ	421.7130	0.000000
BRIX_60_LEFT	99.45397	0.000000
BRIX_65_LEFT	0.000000	1.382015
PAST_OJ_LEFT	0.000000	0.000000
PROFIT	1601.556	0.000000
COST_UNS	0.6000000	0.000000
COST_FROZ	0.3300000	0.000000
COST_60	0.1500000	0.000000
COST_65	0.2000000	0.000000
COST_PAST	0.8000000E-01	0.000000
TOTAL_JUICE	2883.894	0.000000
TOTAL_WATER	1222.348	0.000000

Row	Slack or Surplus	Dual Price
1	99.45397	-1.000000
2	0.000000	-648.9283
3	0.000000	0.000000
4	0.000000	292.0921
5	0.000000	0.000000
6	0.000000	0.1461050
7	0.000000	0.000000
8	0.000000	0.000000
9	0.000000	0.000000
10	0.000000	0.000000
11	0.000000	0.000000
12	0.000000	-393.8657
13	0.000000	0.000000
14	0.000000	53.69979
15	0.000000	0.000000
16	0.000000	0.4065727E-01
17	0.000000	0.000000
18	0.000000	0.000000
19	0.000000	0.000000

20	0.000000	0.000000
21	0.000000	0.000000
22	0.000000	-1.000000
23	0.000000	0.3820149
24	0.000000	-1.000000
25	0.000000	0.000000
26	0.000000	-1.000000
27	0.000000	0.000000
28	0.000000	0.000000
29	0.000000	648.9283
30	0.000000	0.000000
31	0.000000	393.8657
32	0.000000	0.000000
33	0.000000	-292.0921
34	0.000000	-53.69979
35	0.000000	0.000000
36	0.000000	0.000000
37	0.000000	0.000000
38	0.000000	0.000000
39	0.000000	0.000000
40	0.000000	0.000000
41	0.000000	0.000000
42	0.000000	0.000000
43	0.000000	0.000000
44	0.000000	0.000000
45	0.000000	0.000000
46	99.45397	0.000000
47	150.5460	0.000000
48	0.000000	0.000000
49	604.0000	0.000000
50	0.000000	0.9115865
51	906.0000	0.000000

C

LINGO Input for Orange Juice Profit Maximization

The input file used for solving the mathematical optimization problem formulated in section 7.7 with the objective of maximizing the total profit from orange juice production. The optimization problem was solved using LINGO 8.0, LINDO Systems Inc, Chicago, IL, USA.

```
!Objective Function;

model:

max = PROFIT;


!Definition of flows for unsweet pasteurized OJ product;

X_60_uns = F_60_uns / (F_60_uns + F_65_uns);
X_65_uns = F_65_uns / (F_60_uns + F_65_uns);
X_6065_uns = F_6065_uns / (F_6065_uns + F_Past_uns);
X_Past_uns = F_Past_uns / (F_6065_uns + F_Past_uns);

F_6065_uns = F_60_uns + F_65_uns;
F_mix_uns = F_6065_uns + F_Past_uns;


!Calculation of dilution for unsweet pasteurized OJ product;

AUP_mix_calc_uns = X_6065_uns * AUP_6065_uns
                   + X_Past_uns * AUP_Past;

AUP_mix_true_uns = (1-X_water_uns) * AUP_mix_calc_uns
                   + X_water_uns * AUP_water;
```

```
X_water_uns = F_water_uns / (F_water_uns + F_mix_uns);
F_juice_uns = F_water_uns + F_mix_uns;


!Definition of flows for frozen concentrated OJ product;

X_60_froz = F_60_froz / (F_60_froz + F_65_froz);
X_65_froz = F_65_froz / (F_60_froz + F_65_froz);
X_6065_froz = F_6065_froz / (F_6065_froz + F_Past_froz);
X_Past_froz = F_Past_froz / (F_6065_froz + F_Past_froz);

F_6065_froz = F_60_froz + F_65_froz;
F_mix_froz = F_6065_froz + F_Past_froz;


!Calculation of dilution for frozen concentrated OJ product;

AUP_mix_calc_froz = X_6065_froz * AUP_6065_froz
                  + X_Past_froz * AUP_Past;

AUP_mix_true_froz = (1-X_water_froz) * AUP_mix_calc_froz
                  + X_water_froz * AUP_water;

X_water_froz = F_water_froz / (F_water_froz + F_mix_froz);
F_juice_froz = F_water_froz + F_mix_froz;


!Overall mass balance;

BRIX_60_LEFT = 250 - (F_60_uns + F_60_froz);
BRIX_65_LEFT = 604 - (F_65_uns + F_65_froz);
PAST_OJ_LEFT = 907 - (F_Past_uns + F_Past_froz);

PROFIT = Cost_uns*F_juice_uns + Cost_froz*F_juice_froz
       - Cost_60*BRIX_60_LEFT - Cost_65*BRIX_65_LEFT
       - Cost_Past*PAST_OJ_LEFT;

WASTE = BRIX_60_LEFT + BRIX_65_LEFT + PAST_OJ_LEFT;
TOTAL_JUICE = F_juice_froz + F_juice_uns;
TOTAL_WATER = F_water_froz + F_water_uns;
```

```
!Known Values;

X_60_uns = 0.1707;
X_65_uns = 0.8293;
X_60_froz = 0.247;
X_65_froz = 0.753;
X_6065_uns = 0.377;
X_6065_froz = 0.685;
AUP_6065_uns = 48.89823;
AUP_6065_froz = 47.2883;
AUP_mix_true_uns = 15.15652174;
AUP_mix_true_froz = 37.70782609;
AUP_water = 0;
AUP_Past = 18.50368124;
Cost_uns = 0.6;
Cost_froz = 0.33;
Cost_60 = 0.15;
Cost_65 = 0.20;
Cost_Past = 0.08;


!Flowrate Constraints;

F_60_uns + F_60_froz <= 250;
F_60_uns + F_60_froz >= 0;
F_65_uns + F_65_froz <= 604;
F_65_uns + F_65_froz >= 0;
F_Past_uns + F_Past_froz <= 907;
F_Past_uns + F_Past_froz >= 1;
```

D

LINGO Output for Orange Juice Profit Maximization

The output file generated when solving the mathematical optimization problem formulated in section 7.7 with the objective of maximizing the total profit from orange juice production. The optimization problem was solved using LINGO 8.0, LINDO Systems Inc, Chicago, IL, USA.

```
Local optimal solution found at iteration:            92
Objective value:                                1673.611

                     Variable           Value        Reduced Cost
                       PROFIT        1673.611            0.000000
                      X_60_UNS       0.1707000           0.000000
                      F_60_UNS        93.69019           0.000000
                      F_65_UNS        455.1686           0.000000
                      X_65_UNS       0.8293000           0.000000
                    X_6065_UNS       0.3770000           0.000000
                    F_6065_UNS        548.8587           0.000000
                    F_PAST_UNS        907.0000           0.000000
                    X_PAST_UNS       0.6230000           0.000000
                     F_MIX_UNS        1455.859           0.000000
                AUP_MIX_CALC_UNS        29.96243           0.000000
                AUP_6065_UNS        48.89823           0.000000
                     AUP_PAST        18.50368           0.000000
                AUP_MIX_TRUE_UNS        15.15652           0.000000
                  X_WATER_UNS       0.4941490           0.000000
                    AUP_WATER        0.000000           0.000000
                  F_WATER_UNS        1422.180           0.000000
                  F_JUICE_UNS        2878.039           0.000000
                    X_60_FROZ       0.2470000           0.000000
```

F_60_FROZ	0.000000	0.000000
F_65_FROZ	0.000000	0.000000
X_65_FROZ	0.7530000	0.000000
X_6065_FROZ	0.6850000	0.000000
F_6065_FROZ	0.000000	0.000000
F_PAST_FROZ	0.000000	0.000000
X_PAST_FROZ	0.3150000	0.000000
F_MIX_FROZ	0.000000	0.000000
AUP_MIX_CALC_FROZ	38.22115	0.000000
AUP_6065_FROZ	47.28830	0.000000
AUP_MIX_TRUE_FROZ	37.70783	0.000000
X_WATER_FROZ	0.1343024E-01	0.000000
F_WATER_FROZ	0.000000	0.000000
F_JUICE_FROZ	0.000000	0.1708618
BRIX_60_LEFT	156.3098	0.000000
BRIX_65_LEFT	148.8314	0.000000
PAST_OJ_LEFT	0.000000	2.099747
COST_UNS	0.6000000	0.000000
COST_FROZ	0.3300000	0.000000
COST_60	0.1500000	0.000000
COST_65	0.2000000	0.000000
COST_PAST	0.8000000E-01	0.000000
WASTE	305.1413	0.000000
TOTAL_JUICE	2878.039	0.000000
TOTAL_WATER	1422.180	0.000000

Row	Slack or Surplus	Dual Price
1	1673.611	1.000000
2	0.000000	27.44294
3	0.000000	0.000000
4	0.000000	-2152.791
5	0.000000	1066.422
6	0.000000	-0.1914650
7	0.000000	1.186120
8	0.000000	57.63298
9	0.000000	113.9327
10	0.000000	-3413.701
11	0.000000	0.6000000
12	-0.1208432E-07	0.000000
13	-0.1208432E-07	0.000000
14	0.000000	0.000000
15	0.000000	0.000000
16	0.000000	-0.1876500
17	0.000000	0.5076801
18	0.000000	0.000000

19	0.000000	0.000000
20	0.000000	0.000000
21	0.000000	0.5008618
22	0.000000	-0.1500000
23	0.000000	-0.2000000
24	0.000000	2.019747
25	0.000000	1.000000
26	0.000000	0.000000
27	0.000000	0.000000
28	0.000000	0.000000
29	0.000000	-27.44294
30	0.000000	0.000000
31	0.000000	0.000000
32	0.000000	0.000000
33	0.000000	4970.941
34	0.000000	0.000000
35	0.000000	21.72763
36	0.000000	0.000000
37	0.000000	-113.9327
38	0.000000	0.000000
39	0.000000	56.29975
40	0.000000	35.90534
41	0.000000	2878.039
42	0.000000	0.000000
43	0.000000	-156.3098
44	0.000000	-148.8314
45	0.000000	0.000000
46	156.3098	0.000000
47	93.69019	0.000000
48	148.8314	0.000000
49	455.1686	0.000000
50	0.000000	0.000000
51	906.0000	0.000000

List of definitions

Adjustment	The utilization of processing units to adjust the properties of a given stream to make it acceptable as feed to other units (also see interception)
Allocation	Routing of streams to appropriate process units in order to achieve certain process objectives (also see assignment)
Assignment	Routing of streams to appropriate process units in order to achieve certain process objectives (also see allocation)
AUP	AUgmented Property index – Sum of all dimensionless property operators for each stream
Balance equations	Conservational balances such as mass, energy and momentum
CAMD	Computer Aided Molecular Design – generation of compounds having specified properties from molecular fragments using a computerized technique
CAMS	Computer Aided Molecular Search – identification of compounds having specific properties by systematic database searching
CAPEC	Computer Aided Process Engineering Center – A research center at the Department of Chemical Engineering at the Technical University of Denmark. The work presented in this thesis was performed in CAPEC
Cartesian coordinates	Rectangular coordinate system with 2 axes
Cold stream	Process stream requiring heating
Constitutive equations	Property models relating constitutive variables and intensive variables
Constraint equations	Conditions of equilibrium, equipment constraints and other process constraints

Design of experiments	Identification of an optimal set of parameters through statistical analysis. The method systematically reduces the amount of possible combinations of parameters obtained from e.g. factorial design, thereby in the case of mixture design, reducing the number of laboratory trials required
Heat integration	Systematic methodology that provides a fundamental understanding of energy utilization within the process and employs this understanding in identifying energy targets and optimizing heat-recovery and energy-utility systems
HEN	Heat Exchange Network – A network of heat exchangers, where process streams are allowed to exchange heat
Hot stream	Process stream requiring cooling
ICAS	Integrated Computer Aided System – Software package developed by CAPEC
Inter-stream conservation	Fundamental conservation rule stating that when mixing two streams, the individual clusters of the mixture are conserved.
Interception	The utilization of processing units to adjust the properties of a given stream to make it acceptable as feed to other units (also see adjustment)
Intra-stream conservation	Fundamental conservation rule stating that for a given stream all individual clusters must sum to unity, thus for a ternary system, when two clusters are know the third one is implicitly given.
KKT	Karush-Kuhn-Tucker – Set of conditions that solutions to NLPs must satisfy
Lean stream	Process stream capable of accepting the targeted species
LP	Linear Programming – Definition of a linear optimization problem, where an optimal set of continuous variables, that satisfy a given objective, are determined

Mass integration	Systematic methodology that provides a fundamental understanding of the global flow of mass within the process and employs this understanding in identifying performance targets and optimizing the generation and routing of species throughout the process
Mass pinch analysis	Tool for determination of potential for internal mass exchange thereby reducing the need for external utilities. The methodology assists in identification of thermodynamic bottlenecks and systematically targets maximum resource utilization
Matching	Ensuring that all constraints on process units or products are satisfied by the respective feed streams
MEN	Mass Exchange Network – A network of mass exchangers, where process streams are allowed to exchange mass
MILP	Mixed Integer Linear Programming – Identical to a LP problem apart from the inclusion of discrete variables, i.e. they can only assume integer values
MINLP	Mixed Integer Non Linear Programming – Identical to a NLP problem apart from the inclusion of discrete variables, i.e. they can only assume integer values
MSA	Mass Separating Agent – Component or mixture added to a process stream to facilitate mass transfer, e.g. solvents, adsorbents, ion-exchange resins and stripping agents
NLP	Non Linear Programming – Definition of a nonlinear optimization problem, where an optimal set of continuous variables, that satisfy a given objective, are determined
OFAT	One Factor at A Time – Experimental strategy where only one factor is varied in each experiment
Pinch point	Thermodynamic bottleneck, that allows for decomposition of the resource, e.g. energy or mass,

	allocation problem into two sub-problems, i.e. above and below the pinch.
Process integration	Holistic approach to process design, retrofitting, and operation, which emphasizes the unity of the process by providing global insights
Product synthesis/design	The problem of identifying, generating, modifying or optimizing a chemical product
Property operator function	Functional description of physical property allowing for linear mixing rule with respect to relative contributions
Rich stream	Process stream requiring reduction in content of targeted species
Sink	Process unit capable of processing or accepting the sources
Source	Process stream carrying the targeted species. In pollution prevention studies sources are sometimes described as pollutant-laden streams
Source-sink mapping	Visualization tool allowing easy identification of direct recycle opportunities and extent of interception requirements
Ternary coordinates	Triangular coordinate system with 3 axes
Thermal pinch analysis	Tool for determination of potential for internal heat exchange thereby reducing the need for external utilities. The methodology assists in identification of thermodynamic bottlenecks and systematically targets maximum energy utilization
TLFD	Two-Level Factorial Design – Experimental strategy based on statistical methods where each factor is evaluated at two levels, i.e. its low and high values
Tracking	The ability to identify and monitor how the properties of a stream change throughout the process

Nomenclature

α_{Cost}	Relative cost of two external sources
$\bar{x}$	Vector of continuous variables
$\bar{y}$	Vector of integer variables
β_s	Relative cluster arm or cluster composition for stream s
ΔT_{min}	Minimum allowable temperature driving force
γ_{Flow}	Ratio of minimum feasible flowrates of two external sources
λ_j	Coefficients in general equation for describing the line eminating from the vertices in the ternary cluster diagram
ν	Poisson ratio
$\Omega^{\max}_{j,\text{sink}}$	Dimensionless upper bound on feed constraints of sink on property j
$\Omega^{\min}_{j,\text{sink}}$	Dimensionless lower bound on feed constraints of sink on property j
$\Omega_{jFSplit}$	Property split factor
Ω^{*}_{ji}	Dimensionless pure component property operator j for component i
Ω_{jM}	Dimensionless mixture property operator for property j
Ω_{jREAC}	Dimensionless property operator reaction term
Ω_{js}	Dimensionless property operator on the j'th property P_{js} of stream s
$\psi_j(P^{ref}_j)$	Reference property operator on the j'th property P_{js}
$\psi_j(P_{jM})$	Mixture property operator on the j'th property P_{jM}
$\psi_j(P_{js})$	Property operator on the j'th property P_{js} of stream s
θ_{I}	Stoichiometric coefficient for component I in general stoichiometric reaction
$\tilde{X}_v$	Cartesian X-coordinate of vertex point v
$\tilde{Y}_v$	Cartesian Y-coordinate of vertex point v

AUP_M	AUgmented Property index for mixture
AUP_s	AUgmented Property index of stream s
B_s	Degrees Brix of stream s
BAR_s	Brix to acid ratio of stream s
C_k	Concentration of polymer k in blend
$C_{j,\text{sink}}^{\max}$	Upper cluster bound on feed constraints of sink on property j
$C_{j,\text{sink}}^{\min}$	Upper cluster bound on feed constraints of sink on property j
C_{jM}	Mixture property cluster for property j
C_{js}	Property cluster for property j of stream s
$Cost_1$	Cost of stream S_1 per unit flow
$Cost_2$	Cost of stream S_2 per unit flow
$Cost_M$	Cost of mixture per unit flow
$Cost_{Total}$	Total cost for use of external source
E	Young's modulus
$f(\bar{x}, \bar{y})$	General objective function for MILPs and MINLPs
$f(\bar{x})$	General objective function for LPs and NLPs
$f_j(y_{is})$	Functional description of property j as a function of composition
F_s	Flowrate of stream s
F_s	Solvent flowrate
F_w	Wastewater flowrate
F_{obj}	Objective function to be minimized or maximized
G	Shear modulus
$g(\bar{x}, \bar{y})$	Vector of inequality constraints for MILPs and MINLPs
$g(\bar{x})$	Vector of inequality constraints for LPs and NLPs
$g_1(\mathbf{x})$	Process inequality constraints
$g_2(\mathbf{x}, \mathbf{y})$	Structural constraints

$h(\bar{x}, \bar{y})$	Vector of equality constraints for MILPs and MINLPs
$h(\bar{x})$	Vector of equality constraints for LPs and NLPs
$h_1\left(\frac{\partial \mathbf{x}}{\partial \mathbf{z}}, \mathbf{x}, \mathbf{y}\right)$	Process model including transport model
$h_2(\mathbf{x}, \mathbf{y})$	Process equality constraints
HC	Heat capacity flowrate
j	Property ID
K	Bulk modulus
k_s	Absorption coefficient of stream s
M	Mixture
M_k	Mechanical property of polymer k
MW	Molecular weight of polymer
N_C	Number of property clusters
N_s	Total number of streams
N_{Plots}	Number of contour plots to be constructed in DOE methodology
NC	Number of components or constituents
NP	Number of dimensionless property operators
NS	Number of possible separation sequences
NT	Number of potential separation techniques
OM_s	Amount of objectionable material in stream s, expressed as mass fraction
P_s	Amount of pulp in stream s
$P_{j,\text{sink}}^{\max}$	Upper bound on feed constraints of sink on property j
$P_{j,\text{sink}}^{\min}$	Lower bound on feed constraints of sink on property j
P_{ji}^{*}	Pure component property j for component i
pH_s	pH value of stream s
Q_i	Characterization point for identification of feasibility region boundary

Q_j	Characterization point for identification of feasibility region boundary
R_s	Resistivity of stream s
$R_{\infty s}$	Reflectivity of stream s
RVP_s	Reid Vapor Pressure of stream s
S	Solubility (solute mass per unit mass of solvent)
s	Stream ID
S_i	Split factor for component i
S_s	Sulfur content of stream s
T_g	Glass transition temperature
T_s	Supply temperature
T_t	Target temperature
$T_{ColdScale}$	Temperature scale for the process cold streams
$T_{HotScale}$	Temperature scale for the process hot streams
TOC_s	Total organic content of primary pollutants
X	Reaction conversion
X_1	Phenol supply concentration
X_2	Phenol target concentration
x_s	Fractional flowrate contribution of steam s
$X_{CC,mix}$	Cartesian X-coordinate of mixing point
$X_{CC,s}$	Cartesian X-coordinate for stream s
$Y_{CC,mix}$	Cartesian Y-coordinate of mixing point
$Y_{CC,s}$	Cartesian Y-coordinate for stream s
y_{is}	Composition of component i in stream s
$Y_{Scaling}$	Scaling factor for calculation of Cartesian Y-coordinate Y_{CC}
$\mathbf{A}$	Structural matrix
$\mathbf{x}$	Vector of optimization real variables
$\mathbf{y}$	Vector of optimization integer variables
ρ_s	Density of stream s

References

Aggarwal, A. and Floudas, C. (1990). Synthesis of General Distillation Sequences. *Computers and Chemical Engineering*, **14**, 631–653.

Anderson, M. J. and Whitcomb, P. J. (1996). Optimize Your Process-Optimization Efforts. *Chemical Engineering Progress*, **12**, 51–60.

Anderson, M. J. and Whitcomb, P. J. (1998). Find the Most Favorable Formulations. *Chemical Engineering Progress*, **4**, 63–67.

Anderson, M. J. and Whitcomb, P. J. (1999). Computer-Aided Tools for Optimal Mixture Design. *PCI Paint and Coatings*, **11**.

Andrecovich, M. J. and Westerberg, A. W. (1985). A MILP Formulation for Heat Integrated Distillation Sequence Synthesis. *AIChE Journal*, **31**, 1461.

Balakrishna, S. and Biegler, L. (1992). A Constructive Targeting Approach for the Synthesis of Isothermal Reactor Networks. *Industrial and Engineering Chemistry Research*, **31**, 300.

Barnicki, S. and Fair, J. (1990). Separation System Synthesis: A Knowledge-Based Approach - 1. Liquid Mixture Separations. *Industrial and Engineering Chemistry Research*, **29**, 421–432.

Barnicki, S. and Fair, J. (1992). Separation System Synthesis: A Knowledge-Based Approach - 2. Gas/Vapor Mixtures. *Industrial and Engineering Chemistry Research*, **31**, 1679–1694.

Bek-Pedersen, E. (2003). *Synthesis and Design of Distillation Based Separation Schemes*. Ph.D. thesis, CAPEC, Department of Chemical Engineering, Technical University of Denmark.

Biegler, L.; Grossmann, I. and Westerberg, A. (1997). *Systematic Methods of Chemical Process Design*. Prentice Hall PTR, One Lake Street, Upper Saddle River, New Jersey.

Biermann, C. J. (1996). *Handbook of Pulping and Papermaking*. Academic Press, San Diego, CA, USA.

Box, G. E. P.; Hunter, W. G. and Hunter, J. S. (1978). *Statistics for Experimenters*. Wiley, New York.

Brandon, C. E. (1981). Properties of Paper. In J. P. Casey, editor, *Pulp and Paper Chemistry and Chemical Technology*, volume III, 3rd Edition, pages 1739–1746, 1819–1886. John Wiley and Sons, New York, NY, USA.

Brandrup, J. and Immergut, E. H., editors (1989). *Polymer Handbook. Third Edition*, New York. John Wiley and Sons.

Briones, V. and Kokossis, A. (1999). Hypertargets: A Conceptual Programming Approach for the Optimization of Industrial Heat Exchanger Networks - I. Grassroot Design and Network Complexity. *Chemical Engineering Science*, **54**, 519–539.

Cabezas, H.; Bare, J. and Mallick, S. (1999). Pollution Prevention with Chemical Process Simulators: The Generalized Waste Reduction (WAR) Algorithm. *Computers and Chemical Engineering*, **23(4-5)**, 623–634.

CAPEC (2003). *ICAS User Manual*. Department of Chemical Engineering, Technical University of Denmark.

Chen, Y. and Fan, L. T. (1993). Synthesis of Complex Separation Schemes with Stream Splitting. *Chemical Engineering Science*, **48(7)**, 1251–1264.

Ciric, A. and Floudas, C. (1991). Heat Exchanger Network Synthesis Without Decomposition. *Computers and Chemical Engineering*, **15**, 385–396.

Cornell, J. (1990). *Experiments with Mixtures*. John Wiley and Sons, Inc, New York.

Cowie, J. M. G. (1991). *Polymers: Chemistry and Physics of Modern Materials*. Chapman and Hall, New York.

Douglas, J. (1985). A Hierarchical Decision Procedure for Process Synthesis. *AIChE Journal*, **31(3)**, 353–362.

Duvedi, A. P. and Achenie, L. E. K. (1996). Designing Environmentally Safe Refrigerants Using Mathematical Programming. *Chemical Engineering Science*, **51(15)**, 3727–3739.

Eden, M. R.; Koggersbøl, A.; Hallager, L. and Jørgensen, S. B. (2000). Dynamics and Control During Startup of Heat-Integrated Distillation Column. *Computers and Chemical Engineering*, **24**, 1091–1097.

Eden, M. R.; Jørgensen, S. B.; Gani, R. and El-Halwagi, M. M. (2002a). A Novel Framework for Simultaneous Separation Process and Product Design. In *Proceedings of Distillation and Absorption 2002*.

Eden, M. R.; Jørgensen, S. B.; Gani, R. and El-Halwagi, M. M. (2002b). Property Integration - A New Approach for Simultaneous Solution of Process and Molecular Design Problems. *Computer-Aided Chemical Engineering*, **10**, 79–84.

Eden, M. R.; Jørgensen, S. B. and Gani, R. (2003a). A New Modeling Approach for Future Challenges in Process and Product Design. *Computer Aided Chemical Engineering*, **14**, 101–106.

Eden, M. R.; Jørgensen, S. B.; Gani, R. and El-Halwagi, M. M. (2003b). Reverse Problem Formulation Based Techniques for Process and Product Design. *Computer Aided Chemical Engineering*, **15A**, 451–456.

Eden, M. R.; Jørgensen, S. B.; Gani, R. and El-Halwagi, M. M. (2003c). Property Cluster Based Visual Technique for Synthesis and Design of Formulations. *Computer Aided Chemical Engineering*, **15B**, 1175–1180.

Eden, M. R.; Jørgensen, S. B.; Gani, R. and El-Halwagi, M. M. (2004). A Novel Framework for Simultaneous Separation Process and Product Design. *Chemical Engineering and Processing*, **43**, 595–608.

El-Halwagi, M. M. (1997). *Pollution Prevention Through Process Integration: Systematic Design Tools*. Academic Press, San Diego, CA, USA.

El-Halwagi, M. M. and Maniousiouthakis, V. (1989). Synthesis of Mass Exchange Networks. *AIChE Journal*, **35**, 1233–1244.

El-Halwagi, M. M. and Spriggs, H. D. (1998). Solve Design Puzzles with Mass Integration. *Chemical Engineering Progress*, **94**, 25–44.

El-Halwagi, M. M.; Glasgow, I. M.; Eden, M. R. and Qin, X. (2004). Property Integration: Componentless Design Techniques and Visualization Tools. *AIChE Journal*. (Accepted).

Eppinger, S. D. (1991). Model-Based Approaches to Managing Concurrent Engineering. *Journal of Engineering Design*, **2(3)**, 283–290.

Floudas, C.; Ciric, A. and Grossmann, I. (1986). Automatic Synthesis of Optimum Heat Exchanger Network Configurations. *AIChE Journal*, **32**, 276.

Folkes, M. J. and Hope, P. S., editors (1993). *Polymer Blends and Alloys*, Glasgow, UK. Chapman and Hall.

Friedler, F.; Fan, L. T.; Kalotai, L. and Dallos, A. (1998). A Combinatorial Approach for Generating Candidate Molecules with Desired Properties Based on Group Contribution. *Computers and Chemical Engineering*, **22(6)**, 809–817.

Gabriel, F. B.; Harell, D. A. and El-Halwagi, M. M. (2003a). Pollution Targeting Via Functionality Tracking. *AIChE Spring Meeting New Orleans, LA*. Paper 106a.

Gabriel, F. B.; Harell, D. A.; Qin, X. and El-Halwagi, M. M. (2003b). Resource Conservation Targeting Techniques Using Componentless Design Approach. *AIChE Spring Meeting New Orleans, LA*. Paper 143c.

Gani, R. and Bek-Pedersen, E. (2000). Simple New Algorithm for Distillation Column Design. *AIChE Journal*, **46(6)**, 1271–1274.

Gani, R. and O'Connell, J. P. (2001). Properties and CAPE: From Present Uses to Future Challenges. *Computers and Chemical Engineering*, **25**, 3–14.

Gani, R. and Pistikopoulos, E. N. (2002). Property Modelling and Simulation for Product and Process Design. *Fluid Phase Equilibria*, **194-197**, 43–59.

Gani, R.; Nielsen, B. and Fredenslund, A. (1991). A Group Contribution Approach to Computer Aided Molecular Design. *AIChE Journal*, **37(9)**, 1318–1332.

Giannakis, G. B. and Tepedelenlioglu, C. (1998). Basis Expansion Models and Diversity Techniques for Blind Identification and Equalization of Time-Varying Channels. *Proceedings of Institute of Electrical and Electronics Engineers*, **86(10)**, 1969–1986.

Glasser, D.; Hildebrandt, D. and Crowe, C. (1987). A Geometric Approach to Steady Flow Reactors: The Attainable Region and Optimization in Concentration Space. *Industrial and Engineering Chemistry Research*, **26**, 1803–1810.

Glinos, K. N. and Malone, M. F. (1984). Minimum Reflux, Product Distribution and Lumping Rules for Multicomponent Distillation. *Ind. Eng. Chem. Process Des. Dev.*, **23**, 764–768.

Glinos, K. N. and Malone, M. F. (1985). Design of Side-Stream Distillation Columns. *Ind. Eng. Chem. Process Des. Dev.*, **24**, 822–828.

Glinos, K. N. and Malone, M. F. (1988). Optimality Regions for Complex Column Alternatives in Distillation Systems. *Chem. Eng. Res. Des.*, **66**, 229–240.

Greenbaum, J. and Kyng, M., editors (1991). *Design at Work*, Hillsdale, NJ. Lawrence Erlbaum Associates.

Grossmann, I. and Daichendt, M. (1996). New Trends in Optimization-Based Approaches to Process Synthesis. *Computers and Chemical Engineering*, **20**, 665–683.

Gundersen, T. and Naess, L. (1988). The Synthesis of Cost Optimal Heat Exchanger Networks: An Industrial Review of the State of the Art. *Computers and Chemical Engineering*, **12**, 503–530.

Harper, P. M. (2000). *A Multi-Phase, Multi-Level Framework for Computer Aided Molecular Design*. Ph.D. thesis, CAPEC, Department of Chemical Engineering, Technical University of Denmark.

Harper, P. M. and Gani, R. (2001). Computer Aided Tools for Design/Selection of Environmentally Friendly Substances. In M. M. El-Halwagi and S. Sikdar, editors, *Process Design Tools for the Environment*, Philadelphia, USA. Taylor and Francis.

Heikkilä, A.-M. (1999). *Inherent Safety in Process Plant Design - An Index-Based Approach*. Ph.D. thesis, Helsinki University of Technology.

Hermansen, R. D. (1994). Development of a Solvent Database Software Program. *Pollution Technology Review*, **212**, 159.

Hildebrandt, D. and Biegler, L. T. (1995). Synthesis of Reactor Networks. *AIChE Symposium Series*, **91**, 52–67.

Hohmann, E. (1971). *Optimum Networks for Heat Exchange*. Ph.D. thesis, University of Southern California, Los Angeles, USA.

Horn, F. (1964). Attainable Regions in Chemical Reaction Technique. In *Third European Symposium on Chemical Reaction Engineering*, London, UK. Pergamon.

Hostrup, M. (2002). *Integrated Approach to Computer Aided Process Synthesis*. Ph.D. thesis, CAPEC, Department of Chemical Engineering, Technical University of Denmark.

Jaksland, C. A. (1996). *Separation Process Design and Synthesis Based on Thermodynamic Insights*. Ph.D. thesis, CAPEC, Department of Chemical Engineering, Technical University of Denmark.

Jaksland, C. A.; Gani, R. and Lien, K. (1995). Separation Process Design and Synthesis Based on Thermodynamic Insights. *Chemical Engineering Science*, **50(3)**, 511–530.

Joback, K. G. (1994). Solvent Substitution for Pollution Prevention. *AIChE Symposium Series*, **40**, 98.

Karush, N. (1939). Master's thesis, University of Chicago.

King, C. J. (1980). *Separation Processes*. McGraw-Hill Inc., New York.

Koggersbøl, A. (1995). *Distillation Column Dynamics, Operability and Control*. Ph.D. thesis, CAPEC, Department of Chemical Engineering, Technical University of Denmark.

Kokossis, A. C. and Floudas, C. A. (1994). Optimization of Complex Reactor Networks - II. Nonisothermal Operation. *Chemical Engineering Science*, **49**, 1037.

Krishna, R. and Sie, S. (1994). Strategies for Multiphase Reactor Selection. *Chemical Engineering Science*, **49**, 4029–4065.

Kuhn, H. and Tucker, A. (1951). Nonlinear Programming. In J. Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, page 481, Berkeley, CA. University of California Press.

LINDO (2003). *LINGO 8.0 User Manual*. LINDO Systems Inc. USA.

Linnhoff, B. and Hindmarsh, E. (1983). The Pinch Design Method for Heat Exchanger Networks. *Chemical Engineering Science*, **38**, 745–763.

Linnhoff, B.; Townsend, D.; Boland, D.; Hewitt, G.; Thomas, B.; Guy, A. and Marsland, R. (1982). *A User Guide on Process Integration for the Efficient Use of Energy*. Institution of Chemical Engineers, Rugby, UK.

Malone, M. F. and Doherty, M. F. (1995). Separation System Synthesis for Nonideal Liquid Mixtures. *AIChE Symposium Series*, **91**, 9–18.

Marcoulaki, E. C. and Kokossis, A. C. (1998). Molecular Design Synthesis Using Stochastic Optimization as a Tool for Scoping and Screening. *Computers and Chemical Engineering*, **22**, S11–S18.

McCarthy, E.; Fraga, E. S. and Ponton, J. W. (1998). An Automated Procedure for Multicomponent Product Separation Synthesis. *Computers and Chemical Engineering*, **22**, 77–88.

Mims, W.; Wysocki, A. and Weldon, R. (2000). Understanding NFC and RECON Orange Juice Demand. *Florida Cooperative Extension - Electronic Data Information Source (EDIS)*, **FE 175**.

Modi, A.; Aumond, J. P.; Mavrovouniotis, M. and Stephanopoulos, G. (1996). Rapid Plant-Wide Screening of Solvents for Batch Processes. *Computers and Chemical Engineering*, **20**, S375–S380.

Montgomery, D. C. (1991). *Design and Analysis of Experiments*. Wiley, New York.

Nielsen, J.; Hansen, M. and Jørgensen, S. (1996). Heat Exchanger Network Modelling Framework for Optimal Design and Retrofitting. *Computers and Chemical Engineering*, **20**, S249–S254.

Nishida, N.; Stephanopolous, G. and Westerberg, A. W. (1981). A Review of Process Synthesis. *AIChE Journal*, **27(3)**, 321–348.

Odele, O. and Macchietto, S. (1993). Computer Aided Molecular Design: A Novel Method for Optimal Solvent Selection. *Fluid Phase Equilibria*, **82**, 47–54.

Papoulias, S. and Grossmann, I. (1983). A Structural Optimization Approach in Process Synthesis - II. Heat Recovery Networks. *Computers and Chemical Engineering*, **7**, 707–721.

Parthasarathy, G. and El-Halwagi, M. M. (2000). Optimum Mass Integration Strategies for Condensation and Allocation of Multi-Component Volatile Organic Compounds. *Chemical Engineering Science*, **55**, 881–895.

Paul, D. R. and Newman, S., editors (1978). *Polymer Blends*, New York. Academic Press.

Pistikopoulos, E. N. and Stefanis, S. (1998). Optimal Solvent Design for Environmental Impact Minimization. *Computers and Chemical Engineering*, **22(6)**, 717–733.

Pretel, E. J.; López, P. A.; Bottini, S. B. and Brignole, E. A. (1994). Computer-Aided Molecular Design of Solvents for Separation Processes. *AIChE Journal*, **40(8)**, 1349–1360.

Russel, B. M.; Henriksen, J. P.; Jørgensen, S. B. and Gani, R. (2002). Integration of Design and Control Through Model Analysis. *Computers and Chemical Engineering*, **26**, 213–225.

Schembecker, G.; Dröge, T.; Westhaus, U. and Simmrick, K. (1995). A Heuristic-Numeric Consulting System for the Choice of Chemical Reactors. *AIChE Symposium Series*, **91**, 336–339.

Schuler, D. and Namioka, A., editors (1993). *Participatory Design*, Hillsdale, NJ. Lawrence Erlbaum Associates.

Shelley, M. D. and El-Halwagi, M. M. (2000). Component-Less Design of Recovery and Allocation Systems: A Functionality-Based Clustering Approach. *Computers and Chemical Engineering*, **24**, 2081–2091.

Stat-Ease (1999). *Design-Expert 6.0 for Windows*. Stat-Ease Inc.

Suh, N. P. (1990). *Principles of Design*. Oxford University Press, New York.

Thompson, R. and King, C. (1972). Systematic Synthesis of Separation Systems. *AIChE Journal*, **18**, 941.

Vaidyanathan, R. and El-Halwagi, M. M. (1994). Computer-Aided Design of High Performance Polymers. *Journal of Elastomers and Plastics*, **26**, 277–292.

Vaidyanathan, R. and El-Halwagi, M. M. (1996). Computer-Aided Synthesis of Polymers and Blends with Target Properties. *Industrial and Engineering Chemistry Research*, **35**, 627–634.

Vaidyanathan, R.; Gowayed, Y. and El-Halwagi, M. M. (1998). Computer-Aided Design of Fiber Reinforced Polymer Composite Products. *Computers and Chemical Engineering*, **22(6)**, 801–808.

Venkatasubramanian, V.; Chan, K. and Caruthers, J. M. (1995). Genetic Algorithmic Approach for Computer-Aided Molecular Design. *ACS Symposium Series*, **589**, 396–414.

Willets, W. R. (1958). Titanium Pigments. In *Paper Loading Materials*, number 19 in TAPPI Monograph Series, pages 96–114, New York, NY, USA. Technical Association of the Pulp and Paper Industry.

Winograd, T. (1996). Introduction. In T. Winograd, editor, *Bringing Design to Software*, pages xiii–xxv. Addison-Wesley.

Yee, T. and Grossmann, I. (1990). Simultaneous Optimization Models for Heat Integration - II. Heat Exchanger Network Synthesis. *Computers and Chemical Engineering*, **14**, 1165–1171.

Young, D. M. and Cabezas, H. (1999). Designing Sustainable Processes with Simulation: The Waste Reduction (WAR) Algorithm. *Computers and Chemical Engineering*, **23**, 1477–1491.

Young, D. M.; Scharp, R. and Cabezas, H. (2000). The Waste Reduction (WAR) Algorithm: Environmental Impacts, Energy Consumption, and Engineering Economics. *Waste Management*, **20**, 605–615.

Index