DTU Library

Systematic Optimization-Based Integrated Chemical Product–Process Design Framework

Cignitti, Stefano; Mansouri, Seyed Soheil; Woodley, John M.; Abildskov, Jens

Systematic optimization-based integrated chemical product-process design framework

Stefano Cignitti, Seyed Soheil Mansouri, John M. Woodley, Jens Abildskov
Process and Systems Engineering Centre (PROSYS), *Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads*, *Building 229, DK-2800 Kgs. Lyngby, Denmark*

Abstract

An integrated optimization-based framework for product and process design is proposed. The framework uses a set of methods and tools to obtain the optimal product-process design solution given a set of economic and environmental sustainability targets. The methods and tools required are property prediction through group contributions, unless supported with a database, computer-aided molecular and mixture/blend design for generation of novel as well as existing products and mathematical programming for formulating and solving multi-scale integrated process-product design problems. The application of the framework is demonstrated through three case studies: (i) refrigeration cycle unit for R134a replacement, (ii) a mixed working fluid design problem for R134a replacement, and (iii) pure solvent design for water-acetic acid LLE extraction. Through the application of the framework it is demonstrated that all solutions satisfy product, process, economic and environmental targets simultaneously. The solution is obtained through a direct deterministic mathematical optimization strategy. The framework proposed in this work is generic and can be applied to a wide range of problems where an integrated solution to process-product design is beneficial.

Keywords: Integrated chemical product-process design, optimization, CAMD, molecular design, organic Rankine cycle, ORC

1   Introduction

Global economic growth has resulted in an ever-increasing need for new chemical products given the scarcity of resources [1], as well as effective processes to manufacture them. Moreover, sustainable and renewable raw material from which new products can made is also important. Each of these needs may be described in terms of an independent design problem. A generic product design methodology was proposed by Cussler and Wei [2] that consists of four steps: (i) needs, (ii) ideas, (iii) selection and (iv) manufacture; each one of these steps requires sub-methodologies to fulfill their desired goals. The needs for a chemical product may be centered around its properties, its use or its sustainability. Typical chemical products may need to fulfill one or more of such needs. For example, cooking oil primary need is centered on its properties, such as smoke point and toxicity; solvents require specific product properties, but must be also used in several processes where the process performance is a need; refrigerants requirements are present in their sustainability as well, such as ozone depletion potential. Therefore, the transition from ideas to a selection of product candidates is challenging. Often, numerous product alternatives (ideas) can be generated and not enough resources to screen and select amongst available alternatives. The use of computer-aided methods and tools facilitates generation and selection of product candidates. Furthermore, such methods can handle the complexity of integrating process and sustainability requirements in the chemical product design problem.

Several computer-aided methods exist for the design of novel molecules and mixtures. Computer-aided molecular design (CAMD) methods utilize group contribution (GC) description of molecules, that enable the property prediction based on structure alone, to reverse engineer novel molecules and mixtures. Most conventional methods can be categorized in the following:

1. Generate and test
2. Mathematical optimization
3. Continuous molecular targeting

Generate and test methods use CAMD to generate a number of molecule candidates (ideas) and test (selection) them by enumeration for structural consistency, structural design criteria, property targets, mixture targets and/or process criteria. Harper et al. [3] proposed a method for the design of new molecules based on structure and property needs. The method only reaches the testing of pure component and some mixture property targets, but no process design objectives. Satyanarayana et al. [4] added feature to handle polymer design problems. Despite this, generate and test methods fall back when large number of alternatives are encountered. Optimization is not utilized and so, objectives functions for process performance cannot be included. To overcome some of these problems, graphical methods were introduced. Solvason et al. [5] presents a systematic method for CAMD through graphical approach. The graphical method in CAMD generate and test enabled the design of mixtures as well. However, the search space can still only allow for a low number of alternatives and requires human interaction for the graphical design. For designing mixtures using non-graphical methods, Yunus et al. [6] presented a methodology for mixed products, including blends of binary and ternary mixtures. Such generic methods for mixtures design allow for the design of many types of mixtures. However, issues such as process application design and evaluation and use of objective function is still not overcome. For this, mathematical optimization is needed. Duvedi and Achenie (1996) proposed a mathematical program to design novel refrigerants using computer-aided molecular design (CAMD). The design problem only concerns the product properties and does not handle the process needs for the final application of the designed refrigerant. This simplification results in a less computationally expensive process design; however, product targets set by heuristics generate good process criteria. This leads to sub-optimal

solutions if common design variables exist for the process and the product, such as temperature. Hostrup et al. [8] proposed a method for integration of the product and the process design problem. However, heuristics are applied to simplify the mathematical problem with thermodynamic knowledge. The heuristics lead to sub-optimal decisions and this hybrid method does not yield deterministic solutions. Thereby, a structured method of solution, which employs thermodynamic insights to reduce the complexity and size of the mathematical problem by eliminating redundant alternatives, has been developed for the hybrid method. Therefore, the methods uses principles of enumeration rather than rigorous optimization and it does not provide a ranked sub-set of solutions based on the value of the objective function. Instead, it generates a sub-set of feasible solutions in which the optimal solution is not identified. This trend is seen throughout literature, that either products are designed based on property targets, or process targets are introduced through heuristics or simplifications. Papadopoulos et al. [9–11] proposed a systematic design approach for working fluid generation for organic Rankine cycles (ORC). Here, a weighted objective function based on scaled pure properties was proposed and utilized. Papadopoulos et al. [9], utilized the weighted objective function using all the key pure fluid properties. Whereas in a later work by them [10], the problem was reformulated into a multi-objective optimization problem considering only five of the pure fluid properties. Palma-Flores et al. [12] also proposed a method for choosing the optimal pure working fluid for an ORC using CAMD approach. The objective functions were based on temperature dependent pure fluid properties. The drawbacks of these methods are that the design problem is split into multiple stages: a product design phase and a process evaluation/optimization phase. If a given problem has too many product candidates, not all alternatives can be evaluated due to combinatorial explosion and resource demand. Furthermore, the product properties targets do not exactly translate into targeted process criteria. Working fluid design for thermodynamic cycles is a very sensitive and correlated product-process design problem, which today is a particularly important industrial topic [13–15]. The works presented aim at optimizing all the fluid properties, resulting in a difficult task, given also the correlation between properties that may make the result only locally optimal [16,17]. This is also true when considering detailed design of certain components in the ORC that cannot solely be based on primary fluid properties [18].

Continuous molecular targeting methods can be used as a backward approach to eliminate the issue of relating process criteria with product properties. Here, the process is optimized first with the product properties as variable. In the subsequent step, promising product candidates are generated based on the target properties. Such methods have been reported in several works and frequently such a solution approach has been found necessary due to the complexity of the thermodynamic model employed [19–21]. However, the second step of finding the promising chemical product is difficult when multiple target properties involved, thus, finding candidates that match the properties exactly is not possible and the final process criteria changes in an unfavorable direction. This leads to sub-optimal solution [22] and finding a global optimal solution requires screening of all alternatives, making this method an unfavorable alternative.

In cases where the sensitivity and correlation between the product design and process design is high, addressing the two design tasks separately may lead to sub-optimal solutions, the primary limitation of a sequential design approach. An alternative, to ensure a truly optimal product and process design solution, is to carry out such optimization simultaneously using a dual objective function. This simultaneous approach can pave the way to solve product-process design problems, such that a chemical product leading to an optimal process performance is identified in the early stages of process design. This is a need in several product-process design areas as described by Gani and Ng (2015). For example, the design of solvents is highly affected by the process it is used for, whether it be for liquid-liquid extraction or extractive distillation [8,24]. Similarly, working fluids used in

3

thermodynamics cycles are also a type of design problem where product properties cannot be directly related to the process criteria [22,25].

The focus of this paper is on the design of chemical products that include process application needs, which requires the integration of process design and optimization for the optimal chemical product idea selection. The chemical products focus is on molecules and mixtures. Therefore, this paper presents an optimization-based framework for integrated chemical product-process design. The framework consists of four hierarchical steps and uses a set of methods, tools and databases for property prediction, novel product design and mathematical programming to find the optimal product-process design solution. The application of the framework is highlighted through an integrated working fluid design and process design of a refrigeration cycle for a pure and mixed fluid presented in two separate cases as well as a solvent selection for LLE. The paper is structured such that first, the key concepts are presented to introduce the methods and tools incorporated into the computer-aided framework. Next, the framework is presented with the developed steps, algorithms and workflows. Finally, the application of the framework is demonstrated through three different case studies.

2   Systematic computer-aided framework for integrated chemical product-process design

Integrated chemical product-process design is a method for finding the optimal chemical product and optimal process given a set of target specifications for both. For the integration to happen, there must be common variables in the product and process model and/or in the objective function. Furthermore, a complete mathematical program should contain all constitutive and conditional constraints and the solution of this should consider all constraints simultaneously. The integration of the two design problems is explained conceptually through a set of model equations describing products and processes, their interaction and their common variables.

The present framework is restricted to deal with process design models consisting of following equations:

$$0 = f_1\left(\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}, \mathbf{Y}\right) \tag{1}$$

$$0 = g_1\left(\mathbf{x}, \mathbf{y}, \mathbf{u}\right) - \boldsymbol{\theta} \tag{2}$$

$$0 \geq g_2\left(\mathbf{x}, \mathbf{u}, \mathbf{y}\right) \tag{3}$$

$$\left\{\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}\right\} \in \mathbb{R}, \mathbf{Y} \in \{0,1\}$$

Where f_1 represents balance equations, such as energy, mass and momentum conservation, $\mathbf{x}$ is a vector of continuous variables, $\mathbf{y}$ is a vector of output variables, $\mathbf{Y}$ is a vector of binary decision variables, $\mathbf{u}$ is a vector of design variables and $\boldsymbol{\theta}$ is a vector of product properties. The constitutive equations, such as property prediction models, are represented through g_1, g_2 represents conditional equations that the process solution must satisfy, such as mole fraction definition. This generic definition represents process design problems [26]. It can be clearly seen that the product properties influence the design and decisions made at the design phase through the constitutive equations. However, a subset of these properties will be fixed from a process design point of view, if the molecule/mixture has been pre-selected. In most cases only temperature dependent properties will remain unknown and are described by the constitutive equations, for example, activation energy and interaction parameters are known, but reaction rate constant and activity coefficients remain defined by g_2. A generic product design problem will comprise the following model equations:

4

$$0 = f_2\left(\mathbf{x},\mathbf{y},\mathbf{u},\boldsymbol{\theta},\mathbf{Y},\mathbf{Q}\right) \tag{4}$$

$$0 = g_1\left(\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{Q}\right) - \boldsymbol{\theta}$$

$$0 \geq g_2\left(\mathbf{x},\mathbf{u},\mathbf{y}\right)$$

$$\{\mathbf{x},\mathbf{y},\mathbf{u},\boldsymbol{\theta}\} \in \mathbb{R}, \mathbf{Y} \in \{0,1\}, \mathbf{Q} \in \mathbb{Z}$$

Where f_2 describes the product selection. This can be selection from a database, novel molecular generation through CAMD (Eq. (53)-(61)) and mixture selection. Thus, the added vector of integer variables $\mathbf{Q}$ describes the molecular structure type and number of mixture components, while g_1 describes constitutive equations, such as GC models, and g_1 the conditional constraints, such as mixture stability criteria and structure stability.
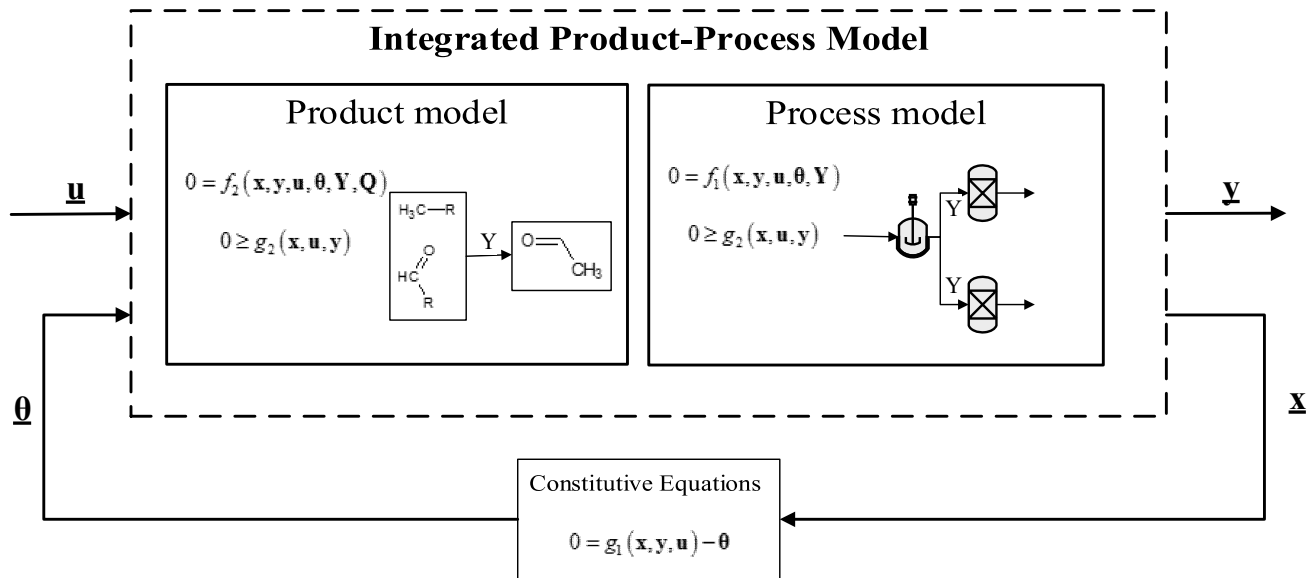

Figure 1: Integrated product-process model

The framework addresses product and process design needs simultaneously for optimal design through an optimization-based approach. The product design refers to the chemical design of pure, mixed or blended products, from now on simply referred to as products. Commonly, one is interested in optimizing a process variable, such as performance, subject to product and process specifications. A generic objective function is formulated as:

$$\max \sum_i w_i J_i\left(\mathbf{x},\boldsymbol{\theta}\right) \tag{5}$$

Where J_i is the objective function with weight w_i. With a single objective function, w_i is unity with a single element. With multi-objective optimization, the weights can be adjusted from 0 to 1 based on heuristic measures. In a generic form, the constraints for product and process design are given by:

$$0 = f_1\left(\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}, \mathbf{Y}\right)$$

$$0 = f_2\left(\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}, \mathbf{Y}, \mathbf{Q}\right)$$

$$0 = g_1\left(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{Q}\right) - \boldsymbol{\theta}$$

$$k_1^L \le k_1\left(\mathbf{Q}\right) \le k_1^U \tag{6}$$

$$k_2^L \le k_2\left(\mathbf{Q}, \mathbf{x}\right) \le k_2^U \tag{7}$$

$$k_3^L \le k_3\left(\mathbf{x}, \mathbf{Y}\right) \le k_3^U \tag{8}$$

$$k_4^L \le k_4\left(\mathbf{x}, \boldsymbol{\theta}\right) \le k_4^U \tag{9}$$

$$\boldsymbol{\theta}^L \le \boldsymbol{\theta} \le \boldsymbol{\theta}^U \tag{10}$$

$$\mathbf{Q} \in \mathbb{Z}; \{\mathbf{x}, \boldsymbol{\theta}\} \in \mathbb{R}; \mathbf{Y} \in \{0,1\}$$

Where k_1 constrains the specifications for molecular design in terms of group combinations, for example for number of functional groups, k_2 constrains mixture structural design, e.g. for number of components and composition, k_3 constrains the process variables to satisfy operational conditions, such as accounting for isothermal processes, k_4 constrains relations between continuous process variables and product properties and constraints temperature dependent or mixture properties. In an isolated view of a decomposed solution, the structural constraints are an MIP, the pure property constraints an LP (as integer solutions are provided from previous program), the mixture properties an NLP and the process models an NLP. An overview of the integrated product-process design framework is presented in Figure 2. As an example, a decomposed problem can be illustrated as follows. In program A, all pure product candidates are designed that will be used in a process that is designed and optimized in program B. Such an example is given in Table 1.

Table 1: Decomposed program example

Program A: Pure product design	Program B: Process design and optimization
$\max f_{obj} = 0$	$\max f_{obj} = J_1$
$f_2 : \sum_i^{N_i} Q_i v_i = 2$	$f_1 : \sum_j^{N_j} y_j - \sum_j^{N_j} u_j = 0$
$g_1 : \sum_i^{N_i} Q_i a_i - \theta_i = 0$	$k_3 : k_3^L \le \sum_j^{N_j} x_j \le k_3^U$
$k_1 : k_1^L \le \sum_i^{N_i} Q_i \le k_1^U$	$k_3 : k_3^L \le x_1 - \theta_1 \le k_3^U$

Starting with program A, all pure fluid candidates are desired. This means a *dummy* objective function is used, where the results is always constant. The, a structural equation is used where the integer variables are summed with the product of valency v_i. The constitutive equations is a predictive calculation of property θ_i, calculated from Q_i with arbitrary constant a_i. The structure is constrained with k_1. In program B, a balance equation is used (e.g. conservation of design variable u_j and y_j. The process variables are constrained in k_3. One θ_i element is then

constrained with respect to x_1 in k_3. A simultaneous program would contain a combination of program A and B without the dummy objective function. Although the programs here are separated in decomposed matter, this simple example feature common variable θ_1. In an LP, direct relation between θ_1 and J_1 could be made. With this, a proper k_4 constraint could be added in program A to narrow the pure fluid candidates. However, in very nonlinear NLP, or in NLP where heuristic or analytical knowledge is missing, such a constraint cannot be made. A simultaneous design path would be preferred in that case.
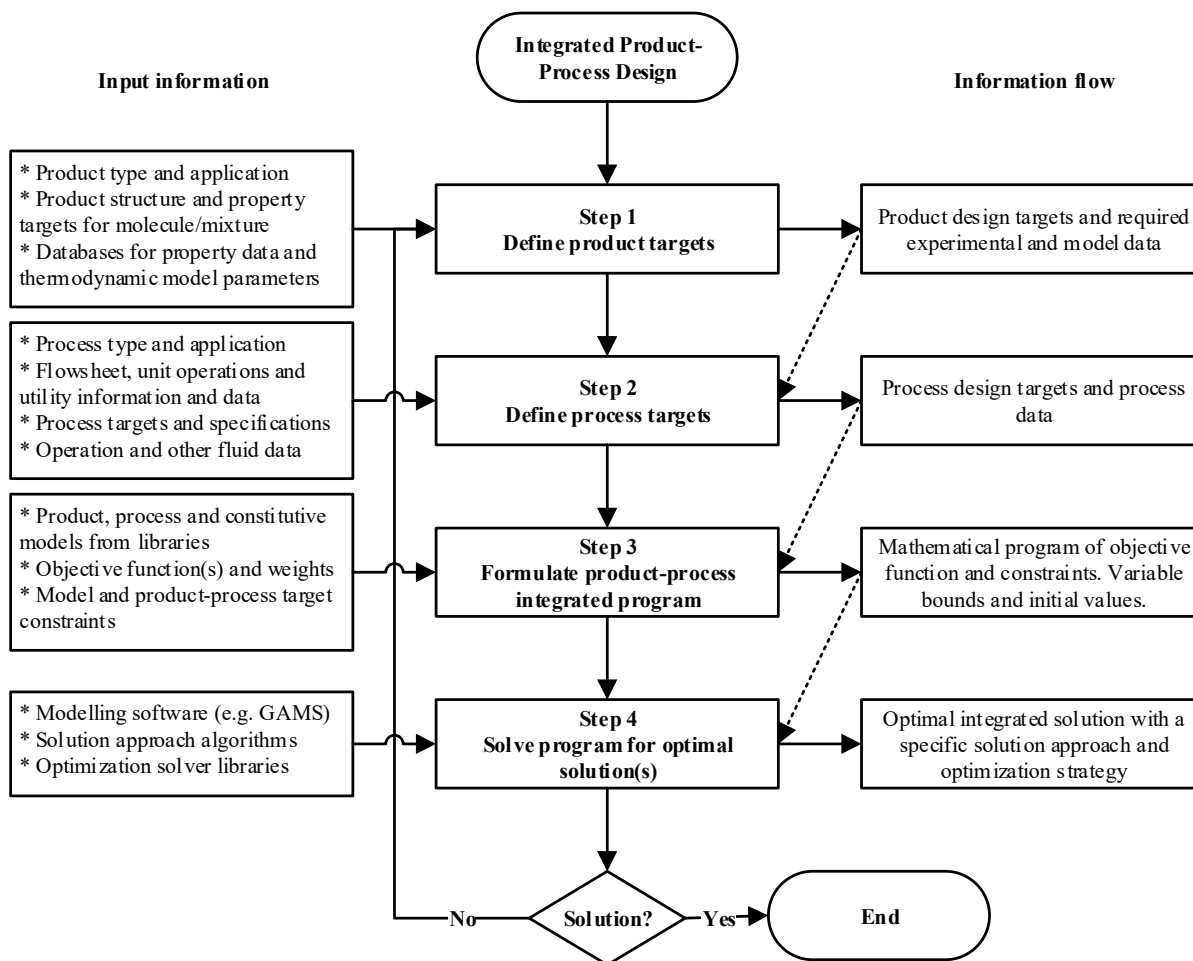

Figure 2: Integrated product-process design and optimization framework

2.1 *Step 1: Define product targets*

The first step in the framework is to define the product targets. It is assumed at this stage that the problem already concerns the integrated design of product and process through mathematical optimization. This step is completed by applying Algorithm 1.

7

Algorithm 1: Define product targets

Objective: To define the product targets for the integrated product-process design problem.

Step 1.1: Define product type, i.e. if the desired product is a molecule, mixture or a blend.

If (product type = molecule)

    Specify the molecular target specifications in terms of structure, functional groups, pure properties and sustainability metrics.

Elseif (product type = mixture/blend)

    Specify mixture/blend targets by defining number of components, mixture properties and mixture stability and phase equilibria conditions. Specify any fixed components and target specifications for the individual components (molecules).

end

Step 1.2: Search space is specified by selecting the option of using a compound database or a CAMD-based method for selection/generation of product.

If (search space = database)

    If (product type = molecule)

        Extract molecules from database based on targets.

    Elseif (product type = mixture/blend)

        Extract mixture components from database based on targets.

    end

    Extract data and specify which properties are missing and require prediction.

Elseif (search space = CAMD-based methods)

    Confine search space based on molecular/mixture component targets.

end

Data: The relevant data for the selected product type can be obtained from various chemical databases [27]*.*

2.2 *Step 2: Define process targets*

Given the product design targets, the process targets are defined in this step. The process is the system where the target product is applied and it is assumed that the separate design decisions are sensitive to each other, that is, one design decision affects the other design decision. The step is performed through an algorithm of sub-steps.

Algorithm 2: Define process targets

Objective: To define process targets for the integrated product-process design problem.

Step 2.1: Define the type of process needed and what the objective of the process is.

Step 2.2: Give the pre-defined design decisions made, i.e. fixed flowsheet information on unit operations, utility flows, auxiliary streams and fluids present and the connection between these.

Step 2.3: Define process design specifications, i.e. target process and/or sustainability metric and operating conditions.

Data: The data is collected from real operational data or from knowledge-based insights.

2.3   *Step 3: Formulate product-process integrated program*

To formulate the integrated mathematical program, Algorithm 3 must be used. To setup this algorithm, the product, process and constitutive models are first formulated. The product models should contain equations for molecular, mixture and/or blend design, depending on targets. Process models should contain balance equations i.e. mass, energy and momentum balances. Constitutional models are the property prediction methods for pure properties (e.g. GC methods), mixture/blend properties (e.g. mixing rules), phase equilibria and mixture stability models (e.g. g^E model and modified Raoults law). The choice of models requires to specify assumptions made on mixture type (e.g. ideal mixture) and on the operation (e.g. isentropic process). The output of the algorithm is the mathematical design program, which can be of several types, such as Linear Programming (LP), Mixed Complementarity Problem (MCP), Mixed Integer Nonlinear Programming (MINLP), Mixed Integer Linear Programming (MIP), Nonlinear Programming (NLP), Quadratically Constrained Program (QCP), Mixed Integer Quadratically Constrained Program (MIQCP) and Mathematical Program with Equilibrium Constraints (MPEC).

Algorithm 3: Formulate product-process integrated program

Objective: To formulate an integrated product-process mathematical program.

Step 3.1: Retrieve from model libraries or formulate product, process and relevant constitutive models. Define the objective function that is one or a set of product and/or process functions. Define constraints by the models from Step (i) and from a list of conditional equality and inequality constraints from defined product and process targets. Define variable types, upper and lower bounds, and initial values. In addition, decision variables are defined, i.e. variable selection of certain processing paths through either process options or super-structures.

Step 3.2: Retrieve the data for the model from Step 1 and 2 and for the constitutive models (e.g. thermodynamic parameters and group contributions).

Step 3.3: Analyze model to ensure at least there is one degree of freedom and identify the common variables between the two designs. Furthermore, identify the type of program.

Data: In addition to data from step I and II, data for constitutional models are retrieved from the sources of the models selected. The data and compounds specific correlations can be retrieved from several databases or handbooks [27–29].

2.4 *Step 4: Solve program for optimal solution(s)*

The integrated program is solved in this step. This requires first selection of appropriate solution approach, optimization strategy and solver. This is covered in the following algorithm.

Algorithm 4: Solve program for optimal solution(s)

Objective: To select appropriate solution approach, optimization strategy and solve the product-process program

Step 4.1: Implement integrated program of objective function(s) and constraints into a modelling based software for optimization.
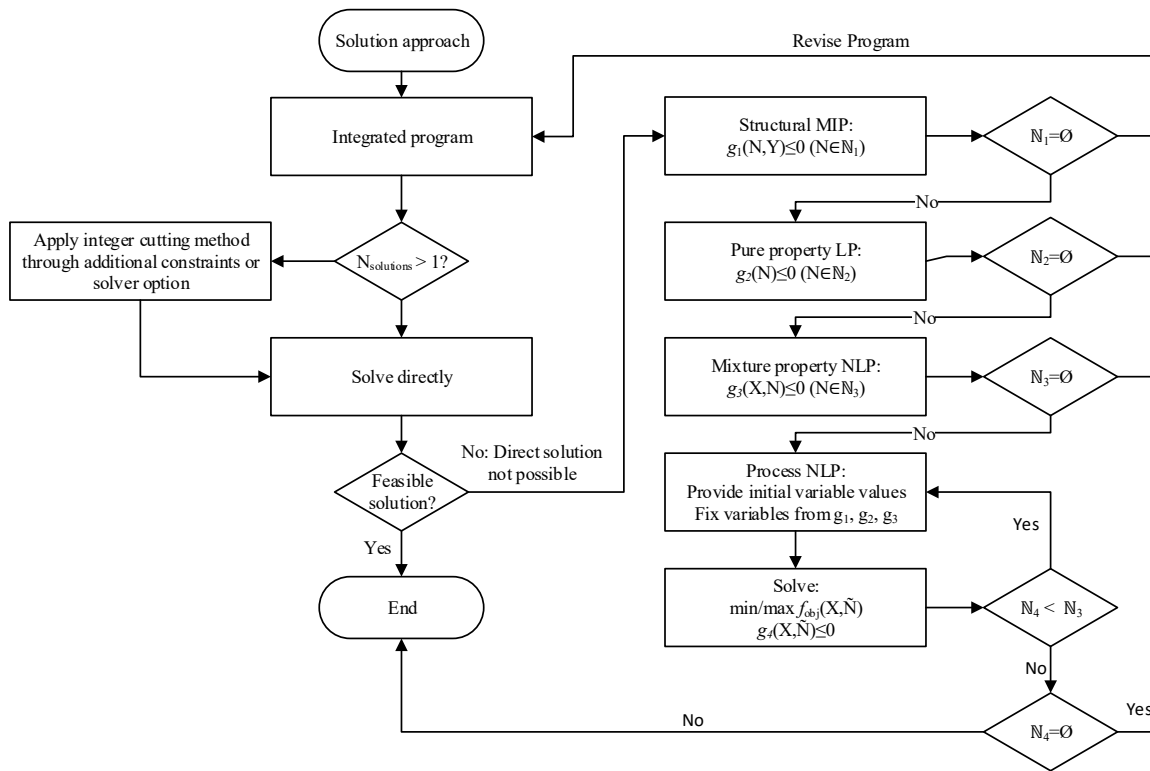

Figure 3: Solution approach selection workflow. A decomposed example is shown where each subprogram is kept separated and where it is necessary to redefine the initial variable values for the process NLP problem. This is because the process NLP problem will be a subset of N problems corresponding to N integer solution from the previous program. Each initial guess must be redefined as the value for a previous sub NLP may lead to infeasible starting values for current NLP. These initial values can be functions of fixed variables that were resulting from previous programs or simply the variable output result from previous NLP. This decomposition strategy has been applied in several works [24,30,31]

Step 4.2: Select solution approach. Figure 3 provides a workflow for selecting the solution approach. If the model can be solved directly and only the global optimal solution is desired, use direct optimization. If multiple solutions are required, one can implement integer cutting methods or using a solver that allows such option (e.g. CPLEX or BARON [32,33]). Decomposed strategies are needed when the problem cannot be solved directly. Here, the constraints and objective function are separated into smaller sets of mathematical programs: Structural constraints as MILP, pure component constraints as LP, mixture property constraints as LP or NLP, and objective function and process design as a sub-set of smaller NLP problems where a NLP problem is solved for each candidate chemical product. The five programs can be kept as five individual programs or some subprograms can be merged. It is best to group subprograms together that have common variables. More advanced decomposition methods can be selected that are case specific.

Step 4.3: Select optimization solver. This choice is based on the type of integrated program formulated in Step III and on the solution approach selected.

Step 4.4: Solve the integrated program and obtain the optimal solution(s) according to workflow in Figure 3.

Tools: An equation oriented algebraic software – for example GAMS [34]*.*

The framework is considered satisfactorily followed when desired solution(s) have been obtained. If no solution is found, return to step 1 for revising product targets, step 2 for process targets, step 3 for re-formulation of constraints, bounds or initial values, or return to step 4 for changing solver or solution approach.

3   Applications examples

In this section, application of the framework is highlighted through three case studies. Each case study is unique in terms of application area. This is to demonstrate the applicability and versatility of the proposed framework to a wide range of problems. The first example is a refrigeration process where the overall performance of the process is subject to the performance of the working fluid in power generation. The second case study features design of mixed working fluids for a heat pump process. Heat pumps are used to transfer energy from a low-temperature source to a high-temperature sink. The last case study is a classic example of integrated product-process application design problem, which is the solvent design for a liquid-liquid extraction process. However, the solution strategy is unique such that it is solved using the proposed framework in this work.

3.1   *Example I: Design of pure working fluid for a refrigeration process to replace R-134a*

The goal here is to design a refrigerant for a refrigeration (vapour-compression cycle) process. This is a replacement problem for a vapour-compression cycle using R-134a refrigerant. Due to recent EU regulations (No. 517/2014), R-134a and other commonly used refrigerants are going to be phased out [35]. A study suggest R-1234yf to be a promising replacement for R-134a [36] as well as other hydrofluoroolefins (HFOs) [37]. A simple schematic of the refrigeration process together with key operating parameters is depicted in Figure 4.
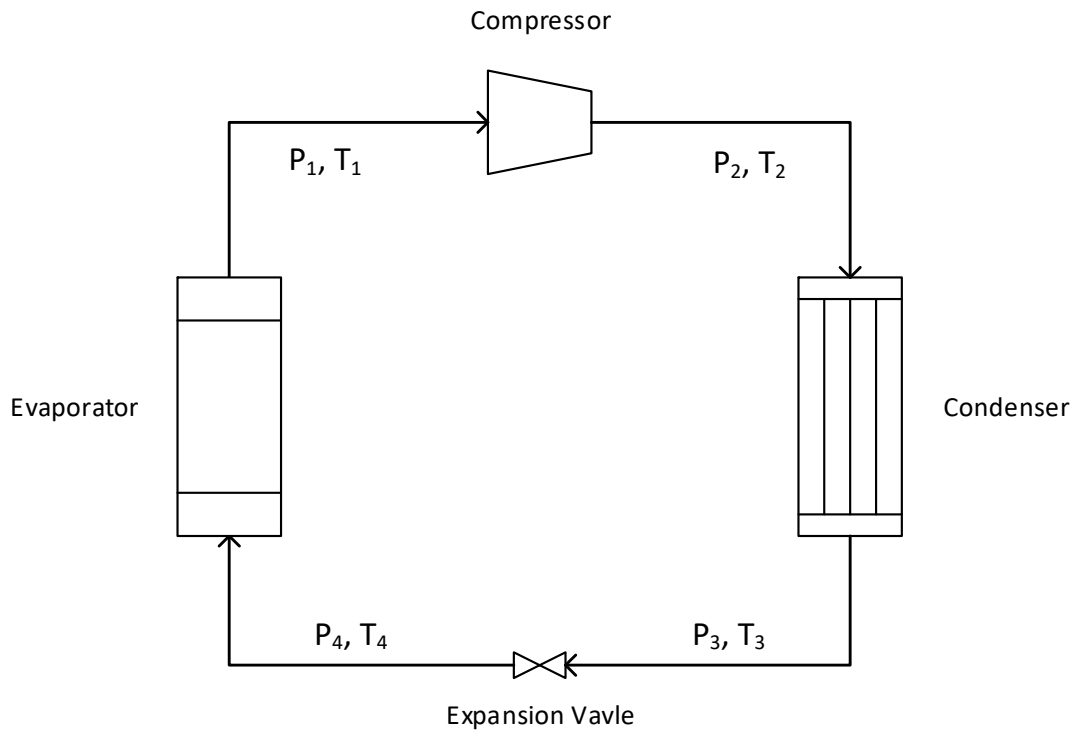
Figure 4: Vapor-compression cycle

Recent studies confirm that the design of novel refrigerants is a challenge. The future needs to find fluids that respect regulations and perform sufficiently well to replace existing refrigeration's that will be out-phased. However, given the high sensitivity between the fluid properties and the cycle process operating conditions and performance, this is a challenging tasks with existing heuristic approaches [38]. Recent literature point at the need for use of CAMD techniques for finding novel refrigerants [25].

Step i: Define product targets

The tasks in this example is to replace R-134a, a common refrigerant used for mobile and industrial refrigeration systems. The R-134a has an exceedingly high GWP that will make it an inappropriate fluid to use according to the F-gas regulation by 2020 [39]. Thus, this is a retrofitting problem where it is desired that the fluid utilizes existing equipment using R-134a. Therefore, Algorithm 1is applied to search for a pure fluid that has an acyclic structure. Given the low number of alternatives existing in literature, the desired search space includes all possible fluids within bounds set targets. The CAMD technique is therefore employed. The target properties are then selected based on targeting similarities with R-134a [40], as given in table 1.

Table 2: Product targets

Target product property	Lower bound	Upper bound
Molecular mass		110 g/mol
Normal boiling point*		250 K
Critical temperature*	350 K	400 K
Critical pressure*	30 bar	50 bar
Enthalpy of Vaporization at Tb*	200 kJ/kg	
Thermal conductivity (liquid, Tb)*	0.08 W/m.K	
ODP*		0
GWP*		1400
Atmospheric Lifetime **		14 years
Number of groups	1	10
Number of functional groups	0	2

* Properties calculated with Marrero and Gani (2001) and [42] method. ** Property must be looked up in literature.

3.1.1 Step ii: Define process targets

In this step, Algorithm 2 is applied. Ranges for existing common process operating conditions are selected for the process targets [36]. The values are translated into utility streams used for the condensation and evaporation of the designed fluid and they are listed in Table 2.

Table 3: Process utility temperatures

Process specifications	Inlet	Outlet
Condensation temperature	313.15 K	333.15 K
Evaporation temperature	280.15 K	265.65 K

3.1.2 Step iii: Formulate product-process integrated program

In this step Algorithm 3 is applied by which the product and process mathematical program is formulated. The objective function is the coefficient of performance (COP), which is the main process performance metric for this kind of processes:

$$\max Obj = \frac{H_1 - H_3}{H_2 - H_1} \tag{11}$$

Where H is the enthalpy of stream j. The product model is firstly described, which starts with the equations for molecular connectivity equations, as described in the key concepts [43]:

$$\sum_i n_i \left(2 - v_i \right) = 2 \tag{12}$$

13

$$1 \leq \sum_i n_i \leq 10 \tag{13}$$

$$0 \leq \sum_i n_i - n_i\left(CH_3\right) - n_i\left(CH_2\right) - n_i\left(CH\right) - n_i\left(C\right) \leq 5 \tag{14}$$

Where n is the number of i groups. The needed product property prediction equations are described by the following equations [42]:

$$T_c = T_{c0} \log\left(\sum_i n_i T_{c_i}\right) \tag{15}$$

$$P_c = \left(\frac{1}{\sum_i n_i P_{ci} + P_{c02}}\right)^2 + P_{c01} \tag{16}$$

$$\omega = \omega_A \left(\log \sum_i n_i \omega_i + \omega_C\right)^{\omega_B} \tag{17}$$

$$k = \sum_i n_i k_i \tag{18}$$

$$GWP = \log\left(\sum_i n_i GWP_i\right) \tag{19}$$

$$ODP = \log\left(\sum_i n_i ODP_i\right) \tag{20}$$

The model for $C_{p,ig}$ is given by the following equation [44]:

$$C_{p,ig} = \left(\sum_{i_1} n_{i_1} A_{i_1} - A_0\right) + \left(\sum_{i_1} n_{i_1} B_{i_1} + B_0\right) T_{st} + \left(\sum_{i_1} n_{i_1} C_{i_1} - C_0\right) T_{st}^3 + \left(\sum_{i_1} n_{i_1} D_{i_1} - D_0\right) T_{st}^3 \tag{21}$$

Where T_c is the critical temperature; P_c is the critical pressure; ω is the acentric factor; k is the thermal conductivity; GWP is the global warming potential; ODP is the ozone depletion potential; $C_{p,ig}$ is the ideal heat capacity. The

properties with subscript i refer to the group contribution of that property. Properties with subscript of number of letter are constants. For the relation between pressure P, temperature T and compressibility factor Z, for the equilibrium values and the calculation of enthalpy, entropy s and fugacity φ, the SRK equation of state is used [45]:

$$f(Z) = 0 = Z_{st,ph}^3 - Z_{st,ph}^2 + Z_{st,ph}\left(\Psi R \frac{T_c^2 P_{st}}{T_{st}^2 P_c}\alpha - \Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}} - \left(\Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}}\right)^2\right) - \Omega\Psi R \frac{T_c^3 P_{st}^2}{T_{st}^3 P_c^2}\alpha \tag{22}$$

$$\alpha = \left(1 + \kappa\left(1 - \sqrt{T_{st}/T_c}\right)\right)^2 \tag{23}$$

$$\kappa = 0.37464 + 1.54226\omega - 0.26992\omega^2 \tag{24}$$

$$h_{st,ph} - h_{ref} = \int C_{p,ig}\,dT_{st} + RT_{st}\left(Z_{st,ph} - 1\right) - \frac{\Psi R}{\Omega}\left(\sqrt{\alpha T_{st} T_c}\,\kappa + \alpha T_c\right)\ln\left(\frac{Z_{st,ph} + \Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}}}{Z_{st,ph}}\right) \tag{25}$$

$$s_{st,ph} - s_{ref} = \int \frac{C_{p,ig}}{T_{st}}\,dT_{st} - R\left(\ln\left(\frac{P_{st}}{P_0}\right) + \ln\left(Z_{st,ph} - \Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}}\right) - \frac{\Psi}{\Omega}\frac{\sqrt{\alpha T_c}\,\kappa}{\sqrt{T_{st}}}\right)\ln\left(\frac{Z_{st,ph} + \Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}}}{Z_{st,ph}}\right) \tag{26}$$

$$\varphi_{st,ph} = \exp\left(Z_{st,ph} - 1 - \ln\left(Z_{st,ph} - \Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}}\right) - \frac{\Psi R}{\Omega}\frac{T_c}{T_{st,ph}}\alpha \ln\left(\frac{Z_{st,ph} + \Omega \frac{T_c}{P_c}\frac{P_{st}}{T_{st}}}{Z_{st,ph}}\right)\right) \tag{27}$$

For the equation-based selection of the correct liquid and vapor phase compressibility factor roots, the following equations are employed [46]:

$$f(Z_{V1}) = 0$$
$$f'(Z_{V1}) \geq 0$$
$$f''(Z_{V1}) \geq 0$$
$$f(Z_{V2}) = 0$$
$$f'(Z_{V2}) \geq 0 \tag{28}$$
$$f''(Z_{V2}) \geq 0$$
$$f(Z_{L3}) = 0$$
$$f'(Z_{L3}) \geq 0$$
$$f''(Z_{L3}) \leq 0$$

The process model is now defined. A model for the compressor efficiency is defined as follows:

$$\eta = \frac{H_2 - H_1}{H_{2s} - H_1} \tag{29}$$

The isofugacity and isentropic constraints are defined as follows:

$$s_2 - s_1 = 0 \tag{30}$$

$$\varphi_{1,V} = \varphi_{1,L} \tag{31}$$

$$\varphi_{3,V} = \varphi_{3,L} \tag{32}$$

$$\varphi_{4,V} = \varphi_{4,L} \tag{33}$$

Isobaric and isothermal constraints:

$$P_2 = P_3, P_4 = P_1, T_4 = T_1 \tag{34}$$

Isenthalpic constraint over the expansion valve:

$$H_3 - H_4 = 0 \tag{35}$$

Energy balances over condenser and evaporator:

$$F_g C_{p_g} \left(T_{1g} - T_{4g} \right) = F_r \left(h_{6,2} - h_{3,1} \right) \tag{36}$$

16

$$F_g C_{p_g} \left(T_{1g} - T_{3g} \right) = F_r \left(h_{6,2} - h_{4,1} \right) \tag{37}$$

Pinch point constraints over heat exchangers:

$$\Delta T_{boiler} \leq T_{3g} - T_4 \tag{38}$$

$$\Delta T_{condenser} \leq T_1 - T_{1w} \tag{39}$$

3.1.3 *Step iv: Solve program for optimal solution(s)*

Several options arise from the possible solution strategies in Figure 4. Since a process model and objective function is used, the mathematical programming approach is used to solve the problem with the BARON optimizer [32] through direct optimization. The solutions are given in Table 4.

Table 4: Molecular design results

Property	R134a	R152a	3,3,3-trifluoropropyne	3,3-difluorobutene	2,2-difluorobutane
MW	102 g/mol	66 g/mol	94 g/mol	92 g/mol	94.1 g/mol
T_b	236 K	230 K	242 K	246 K	252 K
T_C	374 K	366 K	369 K	392 K	389 K
P_C	37.32 bar	46.75 bar	38.13 bar	39.27 bar	37.8 bar
Ω	0.33	0.27	0.28	0.14	0.18
K	0.082 W/m.k	0.097 W/m.k	0.086 W/m.K	0.093 W/m.K	0.092 W/m.K
Δh_v at T_b	26523 J/mol	28562 J/mol	18635 J/mol	21312 J/mol	26991 J/mol
ODP	0	0	0	0	0
GWP	1370 (ASHRAE, 2005)	133 (ASHRAE, 2005)	134.6	2.4	0.9
ATMOSPHERIC LIFETIME	13.4 years (ASHRAE, 2005)	1.5 years (ASHRAE, 2005)	-	-	-
COP	10.1	12.8	9.5	11	11.1

As listed in Table 3, the first solution is R134a. The second solution is a known chemical also called 1,1-Difluoroethane or R-152a. The third and fourth solutions are triple bond and double bond containing refrigerants, respectively. However, they do not perform as well with the COP objective function and self-polymerization may

become an issue. The fifth solution is 2,2-diFluorobutane, which has not been reported as a refrigerant. R-152a has very comparable properties to R134a, but performs better in terms of environmental properties. All three refrigerants are now used for the NLP process model to find optimum operating conditions. Two NLP problems, one for each candidate, are solved to establish the best refrigerant. The KNITRO optimizer in GAMS [47] is used. KNITRO utilizes an interior-point optimization method. It can be seen from the results that R152a obtains a higher COP value than R134a in addition to better environmental properties.

3.2 *Example II: Mixed fluid design for refrigeration cycle process*

Example I is now extended for the design of mixed working fluids. The design problem concerns the replacement of R-134a refrigerant. In this paper, the case study was extended to include an expanded search space, inclusion of utility streams and practical design criteria. Furthermore, both binary mixtures and molecules are investigated. The problem was solved through a simultaneous design strategy using the framework proposed here.
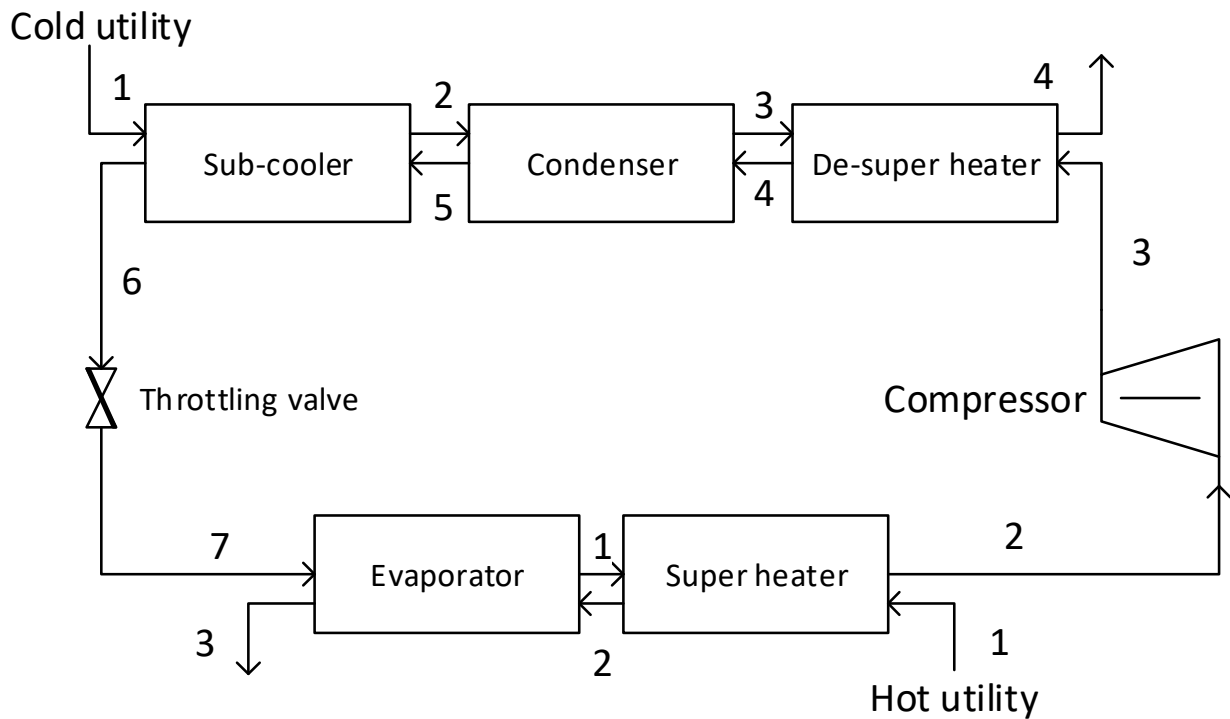

Figure 5: Overview of refrigeration cycle

3.2.1 *Step i: Define product targets*

The design criteria are the same as the one provided for example I (see Table 2). However, Algorithm 1 is followed differently with the selection of a mixture design target rather than a molecular design target.

18

3.2.2   *Step ii: Define process targets*

An overview of the cycle is provided in Figure . A sub-cooler, de-super heater and super heater are added in the cycle for practical reasons. It must be ensured that the temperature of fluids entering and exiting from the compressor exceed the saturation temperatures at the specified pressures. The sub-cooling is a necessary option to have as the pinch point match between cold utility and refrigerant may not lie at the saturation point. A minimum super-heating of 5K and minimum de-super heating of 5K was specified. The outlet of the evaporator and inlet and outlet of the condenser are constrained as saturated state points. The evaporation pressure is lower than the condensation pressure. Hot utility temperature was bounded between 265.65 K and 280.15 K, cold utility temperature has a lower bound of 313.15K and an upper bound was not defined. Minimum approach temperature was set to 10K. These numbers are based on common R-134a refrigeration applications [36]. Minimum temperature and pressure in the cycle are set to 200K and 1 bar, respectively. Extended bounds are allowed to avoid exclusion of optimal solutions, while the simultaneous solution approach helps us avoiding combinatorial explosions.

3.2.3   *Step iii: Formulate product-process integrated program*

In this step, the mathematical program is formulated. The GC method by Hukkerikar et al. [42] was utilized to predict T_c, P_c and ω (acentric factor). The GC method by Joback and Reid [44] was used to predict the ideal gas heat capacity. The SRK equation of state by Soave [45] was used to predict the PVT behavior and properties, particularly enthalpy (H), entropy (S) and fugacity (φ). The objective function is the coefficient of performance (COP):

$$\max COP = Q_{evaporator+superheater} / W_{compressor} \qquad (40)$$

As an equation-based method is needed for the selection of the vapour and liquid compressibility root of the equation of state, the method by Kamath et al. [48] was applied. Remaining equations are the same as example I. van der Waals mixing rules are used to adjust the equation of state (EoS) for mixtures. Much advanced models would be required to model the advantages of using a mixture, such as temperature glide during phase change. Nonetheless, the target remains the optimization of COP, therefore, a simultaneous solution strategy would imply the redundancy in modelling temperature glides.

3.2.4   *Step iv: Solve program for optimal solution(s)*

The program was implemented into GAMS where the simultaneous solution strategy was used. The solver utilized was SBB [49]. For results report generation, the gnuplotxyz and gdx2xls options are used. Table 5 summarizes the product results while Table 6 summarizes the cycle (process) results. Restrictions to pure fluids identified R-161 as a solution. Its results are also well given in Table 6. As it can be seen, the mixture offers significant improvement. A Q-T diagram and a log(P)-H diagram are given in Figure 6 for the mixture and pure component result. It can be seen that the pinches are satisfied and a temperature glide is prominent. The GWP of the mixture is significantly lower than R-134a (1430 kg CO_2 eq) [50].

Table 5: Binary mixture design product results (* predicted properties [42], ** Experimental value 14 kg CO_2 eq [50])
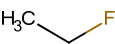
Components	Fluoroethane (R-161)	1,1-difluoroprop-1-ene (novel)
x	0.54	0.46
T_c*	360.84 K	437.09 K
P_c*	52.3 bar	43.29 bar
ω*	0.21	0.23
GWP100*	3.49 kg CO_2 eq **	0 kg CO_2 eq
ODP*	0 eqR11	0 eqR11

Table 6: Binary mixture cycle results

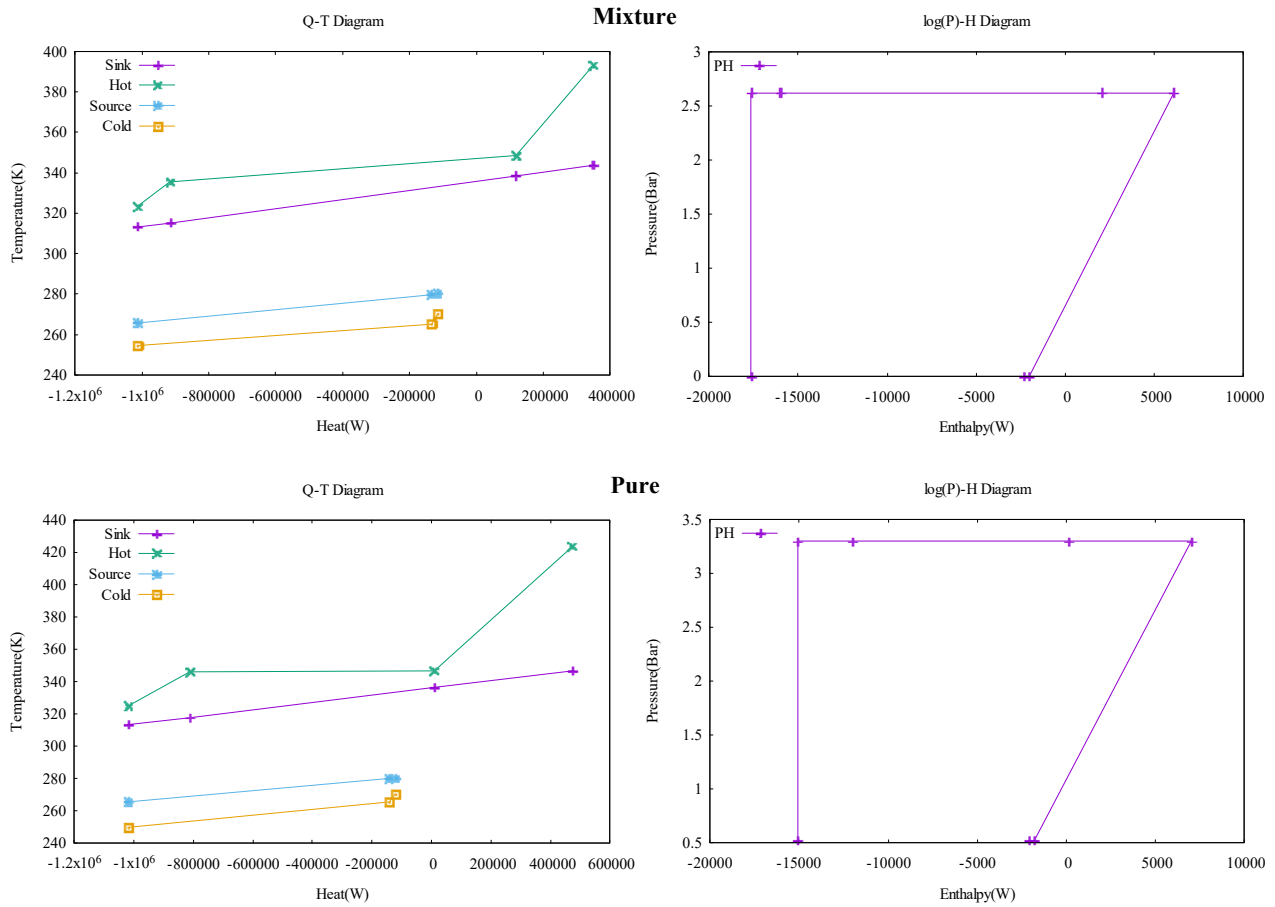|	COP	F [mol/s]	P_2 [bar]	P_3 [bar]	T_7 [K]	T_3 [K]
Mixture	1.93	57.29	1 bar	13.73	254.73	393.06
Pure	1.50	67.67	1.67	27.06	249.76	423.93

Figure 6: Q-T and log(P)-H diagram for the binary mixture and pure solution.

The COP of 1.93 is comparable to systems with similar operating conditions using R-134a [51]. A relaxed solution to the mixture design problems shows a COP value of 1.97. Thus, the discrete group variables are made continuous to obtain a theoretical best solution. As the COP value is close to the one obtained in Table 2, it indicates that a near optimal solution is obtained. However, global optimality can only be guaranteed with a design using a sequential approach and observe if the optimal solution of all feasible working pairs matches the one found in the simultaneous solution. This is subject to an upcoming work by the authors of this manuscript.

3.3   *Example III: Pure solvent for liquid-liquid extraction process using a one stage process representation*

In this case study, a solvent is designed for an extraction process. Given a mixture of water (H_2O) and acetic acid (CH_3COOH), find a suitable solvent that can extract acetic acid from water. The simple block flowsheet of the extraction process is given in Figure 7. The feed is the acetic acid/water binary mixture, the extract is the acetic acid rich phase, whereas the raffinate (R) is the water rich phase.
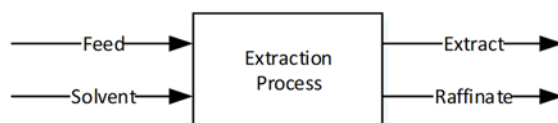
Figure 7: Extraction Process

3.3.1 Step i: Define product targets

It is desired to obtain a solvent that splits acetic acid from water. Therefore, an efficient solvent for removal of acetic acid from water is needed. As the extraction process is operated at 298K, it is needed that the solvent is in a liquid state at this temperature. The boiling point of acetic acid is 391K. In many cases, the solvent is separated from the solute in the extract phase in a solvent recovery system. Thus, the volatility between solute and solvent must be high.

As the needs have been defined, a CAMD problem is formulated. Here, the needs will be formulated through lower and upper bounds of the needed properties. It is specified here that, for the sake of this case study, only acyclic solvents containing C, H and/or O atoms are chosen. Thus, the group set u only contains acyclic groups (see Eq. 8) containing C, H and/or atoms ($u \in$ {CH$_3$, CH$_2$, CH, C, OH, CH$_3$OH, CHO, CH$_3$COO, CH$_2$COO, HCOO, CH$_3$CO, CH$_2$O, HCO, COOH, COO}). The size of the generated solvents will also be limited. The needs with their respective lower and upper bounds are given in Table 7. N_G is the total number of groups u, N_F is the total number of functional groups. The need regarding efficient solvent is here expressed as distribution coefficient m, solvent loss SL, solvent power SP, and selectivity β.

Table 7: Lower and upper bounds for specified needs

Need	Lower	Upper	Need	Lower	Upper
N_G	1	6	m	0.49	-
N_F	0	2	SL	-	0.0038
T_m	-	270K	SP	0.778	-
T_b	430K	-	β	11	-

Step ii: Define process targets

The combined flow of feed and solvent (F) is 2338.2 mol/h. Mole fractions x_i of acetic acid (A), water (B), and solvent (S) are 0.0077, 0.2951 and 0.6972, respectively. The extraction process is operated at 298 K. Every step of the framework from Figure 1 will be walked through in the following.

Step iii: Formulate product-process integrated program

As the objective function, the molar flow of water in the raffinate and of acetic acid in the extract should be maximized (Eq. 6). The process models will be molar balances for the extraction process and liquid-liquid phase equilibrium (Eq. 16-18).

$$f_{obj} = \max\left(x_A E + x_B R\right) \tag{41}$$

As in example I, the same equations are used for the molecular connectivity of groups. However, following pure property model is used for melting point:

$$102.425 \times \log \sum_i N_i T_{mi} \leq 270 \tag{42}$$

And boiling point:

$$430 \leq 204.359 \times \log \sum_i N_i T_{bi} \tag{43}$$

UNIFAC is utilized for the calculation of the activity coefficients and the liquid-liquid equilibria conditions [52]. The model for infinite dilution properties of the solvent, which can be evaluated without the process conditions, are given:

$$0.49 \leq \gamma_{A,B}^{\infty} MW_B / \left(\gamma_{A,S}^{\infty} MW_S \right) \tag{44}$$

$$1 / \gamma_{S,B}^{\infty} \leq 0.00381 \tag{45}$$

$$0.778 \leq 1 / \gamma_{A,S}^{\infty} \tag{46}$$

$$11 \leq \gamma_{B,S}^{\infty} / \gamma_{A,S}^{\infty} \tag{47}$$

The process equations for the extractive single stage are given:

$$\gamma_i^E x_i^E - \gamma_i^R x_i^R = 0, i \in \{A, B, S\} \tag{48}$$

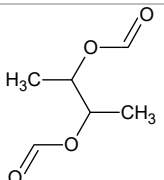$$x_i^F F - x_i^E E - x_i^R R = 0, i \in \{A, B, S\} \tag{49}$$

$$x_A^p + x_B^p + x_S^p = 1, p \in \{F, E, R\} \tag{50}$$

Step iv: Solve program for optimal solution(s)

The MINLP cannot be solved directly due to its size. Thus, the decomposed solution strategy is used (step 4.1 to step 4.4). In step 4.1 constraint 7 to 9 are solved as MIP, which yields 514 solutions. The solutions are screened in step 4.2 through constraint 10 to 11, which yields 196 solutions. These are further screened in step 4.3 in constraints 12 to 15, which yields 19 solutions. These are then solved in a set of 19 NLP with objective function from Eq. 6 with constraints 16 to 18. This yields the optimum solvent, given in Table 8. As it can be seen, a diester (butane-2,3-diyl diformate, groups: 2 CH3, 2 CH, 2 HCOO), was generated that satisfied the constraints and obtained a maximized objective function value. The recovery is 98.3%.

Table 8: Optimum designed solvent and key properties

T_m [K]	T_b [K]	m	SL	β	Possible Structure
236.75	455.22	0.7678	0.0021	14.7702	
SP	x_A^E	x_B^E	x_S^E	x_A^R	
2.1290	0.00859	0.20005	0.79136	0.00112	H_3C … CH_3 (O, O, O)
x_B^R	x_S^R	E [mol/h]	R [mol/h]	f_{obj} [mol/h]	
0.99700	0.00188	2059	278.86	295.72	

4 Discussion

The framework application solution the in the three examples have shown that viable solution can be found. In this discussion, the solution quality and the resources required are compared against conventional solution and solutions from conventional methods.

Comparison with other conventional methods

Heuristics

Example I and II

A heuristic approach for the design of refrigeration systems involving pure and mixed refrigerants is a conventional method of design. Navarro-Esbrí et al. [36] has performed experimental testing of R1234yf as a potential replacement for R134a. Although the fluid will comply with current and future regulations, it is reported it has a 19% lower COP than R134a. Our designed fluids show that we obtain same or higher COP with improved environmental properties of the fluid. However, the comparison is difficult between the example results in this paper and the results of the cycle process performance from other works. This is primarily due to differences in process conditions, product targets and the overall scope of the problem.

Example III

A heuristic approach for the acetic acid and water case study would require a sequence of decision making process for the search space confinement, selection of a handful of potential fluid candidates followed by model-based and experiment-based testing. This has several disadvantages. For one, it would require the search space confinement only using two chemical properties (melting and boiling point). However, as given in the case results, this would require the further evaluation of 512 fluids. This imposes a heavy burden on resources required for the experimental testing. Model-based testing through a process simulator can be feasible approach in this case, although very tedious. A final point, the final solution can only be as good as the knowledge-based capabilities and will not contain the design of new molecules. Classical solvent selection requires experienced people within a process studies group. This experience can be hard to maintain; however, PSE also requires experienced personnel. Many process studies groups are being cut these decades.

The framework designed and optimized a potential fluid (see Table 2). The separation problem of acetic acid and water is widely studied in literature. In [53] methyl ethyl ketone was designed as solvent for the same problem.

Sequential (decomposed approach)

Example I and II

As it was presented, working fluids are difficult fluids to design new alternatives for due to the high sensitivity between the product and process design variables. It is therefore the case that sequential, as with heuristic methods, obtain sub-optimal design solutions. Works that employ a sequential approach to replacement problems for R134a are found in literature. Seader et al. [54] performed an exergy analysis for R600a as a replacement for R134a and found a COP of 1.82. This value matches very well with our results and their existing system obtained a 54.95% improved compared to when used with R134a. However, as explained with detail in the work, the process required optimization and modification of the compressor.

Example III

Conventionally, the acetic-acid and water problem would have been solved using sequential approach (or decomposed approach). Where the integrated model would be solved in sequential steps involving the gradual confinement of the search space as follows: selection based on molecular groups, selection based on the pure compound properties and final selection based on the mixture properties. Such methods rarely involved process models or objective function if not only for verification purposes afterwards. Such conventional methods include the ProCAMD computer-aided tool [55]. Karunanithi etl a. (2005) designed 2-hexanone as a potential solvent for this problem using the sequential approach, however, it does not achieve same recovery.

SWOT analysis

The presented framework capabilities and weaker sides are summarized in a SWOT table.

Strengths	Weaknesses
• A holistic design approach can be achieved accounting for the design and optimization of product and process application. • The approach is particularly strong and outperforms other methods when there is a high design sensitivity between the product and process side. • The framework can accommodate different types of processes, products and search spaces.	• The method requires models for the product or a product database and a process model. • The problems need to be well bounded and a feasible initial point is needed. • The method is prone to uncertainties, as with any prediction method. • Large process models may need to be simplified down to smaller representative processes
Opportunities	**Threats**
• New products can be found and further verified that otherwise would not be considered. • The product and process synergies can be utilized to find a global or close to global optimal solution.	• The decomposed method can be preferred when all feasible fluid designed solutions are desired, e.g. for sensitivity studies. • Use of more complex and high fidelity thermodynamic models or process models

• More sustainable fluids can be designed that comply with international regulations.	can make the simultaneous solution of the MINLP infeasible.

5 Conclusions

A systematic, multi-level and multi-scale framework has been developed for the holistic design and optimization of process application systems. The framework can integrate and solve simultaneously product and process design models. The proposed framework is resilient and robust as it can be applied to design different types of problems at different levels of complexity and across scales. Furthermore, problems with high sensitivity between the product and process application design decision variables are overcome using this framework. For the refrigeration design problem, pure and mixed refrigerants were designed that outperforms the replacement fluid R134a both on regulatory level, but also on a performance level. The designed fluids are comparable with other proposed fluids in literature that were based on heuristic or conventional modelling approaches. A liquid-liquid extraction of acetic acid and water mixture was solved by the design and optimization of a pure solvent and single stage extraction process. The designed pure solvent outperforms on recovery with other solvents in literature reported for this problem. Overall, the framework and cases have shown a capable approach for designing new fluids and processes through optimization that can offer substantial demands in resources required for otherwise solving such difficult problems.

Acknowledgements

References

[1] S.S. Mansouri,, I.A. Udugama,, S. Cignitti,, A. Mitic,, X. Flores-Alsina,, K. V. Gernaey, Current Opinion in Chemical Engineering 18 (2017) 1–9. 10.1016/j.coche.2017.06.002.

[2] E.L. Cussler,, J. Wei, AIChE Journal 49(5) (2003) 1072–5. 10.1002/aic.690490502.

[3] P.M. Harper,, R. Gani,, P. Kolar,, T. Ishikawa, Fluid Phase Equilibria 158–160 (1999) 337–47. 10.1016/S0378-3812(99)00089-8.

[4] K.C. Satyanarayana,, J. Abildskov,, R. Gani, Computers and Chemical Engineering 33 (2009) 1004–13. 10.1016/j.compchemeng.2008.09.021.

[5] C.C. Solvason,, N.G. Chemmangattuvalappil,, M.R. Eden, Computers and Chemical Engineering 33 (2009) 977–91. 10.1016/j.compchemeng.2008.11.001.

[6] N.A. Yunus,, K. V. Gernaey,, J.M. Woodley,, R. Gani, Computers and Chemical Engineering 66 (2014) 201–13. 10.1016/j.compchemeng.2013.12.011.

[7] A.P. Duvedi,, L.E.K. Achenie, Chemical Engineering Science 51(15) (1996) 3727–39. 10.1016/0009-2509(96)00224-2.

[8] M. Hostrup,, P.M. Harper,, R. Gani, Computers and Chemical Engineering 23 (1999) 1395–414. 10.1016/S0098-1354(99)00300-2.

[9] A.I. Papadopoulos,, M. Stijepovic,, P. Linke,, P. Seferlis,, S. Voutetakis, Chemical Engineering Transactions 21 (2010) 61–6. 10.3303/CET10210011.

[10] A.I. Papadopoulos,, M. Stijepovic,, P. Linke, Applied Thermal Engineering 30(6–7) (2010) 760–9. 10.1016/j.applthermaleng.2009.12.006.

[11] A.I. Papadopoulos,, M. Stijepovic,, P. Linke,, P. Seferlis,, S. Voutetakis, Computer Aided Chemical Engineering 30(June) (2012) 66–70. 10.1016/B978-0-444-59519-5.50014-9.

[12] O. Palma-Flores,, A. Flores-Tlacuahuac,, G. Canseco-Melchor, Computers and Chemical Engineering 72 (2015) 334–49. 10.1016/j.compchemeng.2014.04.009.

[13] S. Cignitti,, J. Frutiger,, B. Zühlsdorf,, F. Bühler,, J.G. Andreasen,, F. Müller,, F. Haglind,, B. Elmegaard,, J. Abildskov,, G. Sin,, J. Woodley, Dansk Kemi 97(10) (2016) 10–2.

[14] J. Frutiger,, J. Abildskov,, G. Sin, Computer Aided Chemical Engineering 38 (2016) 283–8. 10.1016/B978-0-444-63428-3.50052-7.

[15] J.G. Andreasen,, U. Larsen,, T. Knudsen,, L. Pierobon,, F. Haglind, Energy 73 (2014) 204–13. 10.1016/j.energy.2014.06.012.

[16]   H. Chen,, D.Y. Goswami,, E.K. Stefanakos, Renewable and Sustainable Energy Reviews 14(9) (2010) 3059–67. 10.1016/j.rser.2010.07.006.

[17]   R. Prapainop,, K.O. Suen, International Journal of Engineering Research and Applications 2(4) (2012) 486–93.

[18]   J. Bao,, L. Zhao, Renewable and Sustainable Energy Reviews 24 (2013) 325–42. 10.1016/j.rser.2013.03.040.

[19]   M. Lampe,, M. Stavrou,, H. Bücker, Industrial and Engineering Chemistry Research 53 (2014) 8821−8830.

[20]   M. Lampe,, M. Stavrou,, J. Schilling,, E. Sauer,, J. Gross,, A. Bardow, Computers and Chemical Engineering 81 (2015) 278–87. 10.1016/j.compchemeng.2015.04.008.

[21]   D. Roskosch,, B. Atakan, Energy 81 (2015) 202–12. 10.1016/j.energy.2014.12.025.

[22]   S. Cignitti,, J.G. Andreasen,, F. Haglind,, J.M. Woodley,, J. Abildskov, Applied Energy 203 (2017) 442–53. 10.1016/j.apenergy.2017.06.031.

[23]   R. Gani,, K.M. Ng, Computers and Chemical Engineering 81 (2015) 70–9. 10.1016/j.compchemeng.2015.04.013.

[24]   S. Cignitti,, L. Zhang,, R. Gani, Computer Aided Chemical Engineering 37 (2015) 2093–8. 10.1016/B978-0-444-63576-1.50043-1.

[25]   J. Andreasen,, A. Meroni,, F. Haglind, Energies 10(4) (2017) 547. 10.3390/en10040547.

[26]   S.S. Mansouri,, J.K. Huusom,, R. Gani,, M. Sales-Cruz, AIChE Journal 62(9) (2016) 3137–54. 10.1002/aic.15322.

[27]   S. by Aic. Design Institute for Physical Properties, DIPPR Project 801 - Full Version, 2005.

[28]   J.P. O'Connell,, J.M. Prausnitz,, B.E. Poling, The properties of gases and liquids, McGraw-Hill Education; 5 edition (November 27, 2000), 2000.

[29]   D.W. Green,, R.H. Perry, Perry's Chemical Engineers' Handbook, 8th ed., McGraw-Hill, New York, 2008.

[30]   S. Kalakul,, S. Cignitti,, L. Zhang,, R. Gani, Computer Aided Chemical Engineering 38 (2016) 2343–8. 10.1016/B978-0-444-63428-3.50395-7.

[31]   L. Zhang,, S. Cignitti,, R. Gani, Computers & Chemical Engineering 78 (2015) 79–84. 10.1016/j.compchemeng.2015.04.022.

[32]   N. V. Sahinidis, Journal of Global Optimization 8(2) (1996) 201–5. 10.1007/BF00138693.

[33]   IBM Corp.,, IBM, International Business Machines Corporation 12(1) (2009) 481.

[34]   M. Bussieck,, A. Meeraus, Applied Optimization 88 (2004) 137–58. 10.1007/978-1-4613-0215-5_8.

[35]   A. Mota-Babiloni,, J. Navarro-Esbrí,, Á. Barragán-Cervera,, F. Molés,, B. Peris, International Journal of Refrigeration 52 (2015) 21–31. 10.1016/j.ijrefrig.2014.12.021.

[36]   J. Navarro-Esbrí,, J.M. Mendoza-Miranda,, a. Mota-Babiloni,, a. Barragán-Cervera,, J.M. Belman-Flores, International Journal of Refrigeration 36(3) (2013) 870–80. 10.1016/j.ijrefrig.2012.12.014.

[37]   J.S. Brown, ASHRAE Journal 51(8) (2009) 22–9.

[38]   J. Frutiger,, J. Andreasen,, W. Liu,, H. Spliethoff,, F. Haglind,, J. Abildskov,, G. Sin, Energy 109 (2016) 987–97. 10.1016/j.energy.2016.05.010.

[39]   M. Schulz,, D. Kourkoulas, Official Journal of the European Union (517) (2014) L150/195-230.

[40]   ASHRAE, ASHRAE Handbook: Fundamentals, American Society of Heating, Refrigerating and Air-conditioning Engineers, Inc., Atlanta, USA, 2009.

[41]   J. Marrero,, R. Gani, Fluid Phase Equilibria 183–184 (2001) 183–208. 10.1016/S0378-3812(01)00431-9.

[42]   A.S. Hukkerikar,, B. Sarup,, A. Ten Kate,, J. Abildskov,, G. Sin,, R. Gani, Fluid Phase Equilibria 321 (2012) 25–43. 10.1016/j.fluid.2012.02.010.

[43]   O. Odele,, S. Macchietto, Fluid Phase Equilibria 82 (1993) 47–54.

[44]   K.G. Joback,, R.C. Reid, Chemical Engineering Communications 57(1–6) (1987) 233–43. 10.1080/00986448708960487.

[45]   G. Soave, Chemical Engineering Science 27(6) (1972) 1197–203. 10.1016/0009-2509(72)80096-4.

[46]   R.S. Kamath,, L.T. Biegler,, I.E. Grossmann, Computers and Chemical Engineering 34(12) (2010) 2085–96. 10.1016/j.compchemeng.2010.07.028.

[47]   R.H. Byrd,, J. Nocedal,, R.A. Waltz, Energy 83 (2006) 1–25. 10.1007/0-387-30065-1_4.

[48]   R.S. Kamath,, L.T. Biegler,, I.E. Grossmann, Computers and Chemical Engineering 34(12) (2010) 2085–96. 10.1016/j.compchemeng.2010.07.028.

[49]   M.R. Bussieck,, A. Drud, GAMS Development Corp (2001).

[50]   M. Goedkoop,, R. Heijungs,, M. Huijbregts,, A. De Schryver,, J. Struijs,, R. Van Zelm, ReCiPe 2008, 2013.

[51]   S. Karagoz,, M. Yilmaz,, O. Comakli,, O. Ozyurt, Energy Conversion and Management 45(2) (2004) 181–96. 10.1016/S0196-8904(03)00144-4.

[52] A. Fredenslund,, R.L.R. Jones,, J.M.J. Prausnitz, AIChE Journal 21(6) (1975) 1086–99. 10.1002/aic.690210607.

[53] J.D. Seader,, E.J. Henley,, D.K. Roper, Separation process principles, Wiley, New York, 1998.

[54] M.M. Joybari,, M.S. Hatamipour,, A. Rahimi,, F.G. Modarres, International Journal of Refrigeration 36(4) (2013) 1233–42. 10.1016/j.ijrefrig.2013.02.012.

[55] P.M. Harper,, R. Gani, Computers and Chemical Engineering 24 (2000) 677–83. 10.1016/S0098-1354(00)00410-5.

[56] A.T. Karunanithi,, L.E.K. Achenie,, R. Gani, Industrial and Engineering Chemistry Research 44(13) (2005) 4785–97. 10.1021/ie049328h.

[57] S. Cignitti,, J.M. Woodley,, J. Abildskov, Computer Aided Chemical Engineering (2017) (In Press).

[58] R. Gani,, L. Zhang,, S. Kalakul,, S. Cignitti, Computer Aided Chemical Engineering, Vol. 39, Elsevier, 2017, pp. 153–96.

[59] A.S. Hukkerikar, Development of pure component property models for chemical product-process design and analysis. Techinical University of Denmark, 2013.

[60] J.W. Kang,, J. Abildskov,, R. Gani,, J. Cobas, Industrial & Engineering Chemistry Research 41(13) (2002) 3260–73. 10.1021/ie010861w.

[61] R. Gani,, P.P.M. Harper,, M. Hostrup, Industrial & Engineering Chemistry Research 44(44) (2005) 7262–9. 10.1021/ie0501881.

[62] S. Jonuzaj,, C.S. Adjiman, Chemical Engineering Science 159 (2017) 106–30. 10.1016/j.ces.2016.08.008.

Appendix: Key concepts

This section presents some of the key concepts that are used throughout the integrated chemical product-process design method, which is implemented through a computer-aided optimization framework. The concepts are used in parts of the algorithms and in product-process model constraints to generate novel products and to predict properties.

5.1 *Group-contribution (GC) methods*

GC methods are property prediction methods that only require the molecular structure for property prediction. The group contribution methods are generally formulated as follows [41]:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \tag{51}$$

Where X is a property of the function $f(X)$. N and M are the number of first and second order groups, respectively; while C and D are the contributions to the first and second order groups, respectively. The first order approximation utilizes representation of simple molecular groups and functional groups, such as CH_3 and OH. The second order approximations use the combination of first order groups to increase the model accuracy by considering adjacent first order groups, such as $(CH_3)_2CH$ group. Similarly, O and E refer to the third order approximation by its number of groups and its contribution, respectively. Second order groups provide an improved description of polyfunctional components and can differentiate between some isomers. Third order groups allows prediction of complex heterocyclic and larger polyfunctional acyclic components [59]. Finally, w and z are binary variables that determine if a certain order of approximation is to be included.

There are also GC methods for the prediction of mixture phase equilibria for mixtures and blends, such as UNIFAC or KT-UNIFAC [52,60]. These excess Gibbs free energy models provide activity coefficients needed for phase equilibria constraints and mixing rules for property prediction. In addition, they provide prediction for interaction parameters needed also in other thermodynamic models, such as equations of state (EoS). GC methods are also used for temperature dependent properties, such as ideal gas heat capacity proposed by Joback and Reid (1987), where Eq. (51) is modified to include temperature dependent terms.

There are certain limitations to GC methods. The methods require description of molecular groups. If a molecular group description is missing for a part of a molecule, the prediction is not possible. For obtaining the contribution of this new groups, it is required to find experimental data and re-do the regression. To overcome this, Gani et al. (2005) combined atom connectivity index (CI) method with GC to create missing groups and predict their contributions. This new type of method is called Group Contribution[+] Method (GC[+]). The CI method utilizes the following equation:

$$f(X) = \sum_i a_i A_i + b(^v\chi^0) + 2c(^v\chi^1) + d \tag{52}$$

Where a_i is the contribution of atom i and A_i is the number of occurrences of group i in the molecule. The parameter $^v\chi^0$ is the zeroth-order atom valence connectivity index, $^v\chi^1$ is the first-order bond connectivity index.

The symbols b and c represent adjustable parameters and d is a universal parameter. Hukkerikar et al (2012) extended the GC method of Marrero and Gani (2001) to be used as a GC$^+$ method. From this, environmental properties have also been included. These environmental properties are used for sustainability specifications.

5.2   *Computer-aided molecular design (CAMD)*

CAMD methods employ a reverse engineering approach to the GC methods; given a set of target properties find the novel molecules that satisfy these targets. CAMD methods use a set of constraints that describe chemically feasible and stable molecules in combination with GC property prediction functions. This results in MILP problems that often are solved through optimization or generate and test methods [23]. The octet rule provides a simple relation for the structural feasibility of a collection of groups [43] by classifying different structural groups based on their valencies (number of free attachments).

$$\sum_{i_2 \in G_1} \left(2 - v_{i_1}\right) n_{i_1} = 2q \tag{53}$$

$$\sum_{i_1 \neq i_2; i_1, i_2 \in G_1} n_{i_2} \geq n_{i_1}\left(v_{i_1} - 2\right) + 2 \quad \forall i_1, i_2 \in G_1 \tag{54}$$

Eq. (53) is the octet rule and eq. (54) ensures single bonds between groups. The parameter $n_i^{(1)}$ is the number of first-order groups G_1 of set i in the target molecule, v_i is the valency of group i and q is assigned the value of 1, 0 or -1 for acyclic, monocyclic or bicyclic groups, respectively. Bicyclic groups with or without branches require additional constraints. Additional constraints may be placed on the number ($n_i^{(1)}$) of groups to keep it within lower and upper bounds, n_i^L and n_i^U, respectively.

$$n_{i_1}^L \leq n_{i_1} \leq n_{i_1}^U \tag{55}$$

Another constraint may be imposed on the total number of groups making up a molecule.

$$n^L \leq \sum_{i_1 \in G_1} n_{i_1} \leq n^U \tag{56}$$

Eq. (53)-(56) are valid for structural combination of first order molecular groups. The drawback is that higher order groups cannot be incorporated and do not add the additional accuracy of prediction and structural information. Higher order groups are conventionally added in post-design phase through a combinatorial problem where all different combinations are tested [55]. Recently, Zhang et al. (2015) proposed an equation-based optimization approach for including higher order groups in CAMD problem inherently in the optimization problem. Eqs. (53)-(56) can be only used to generate acyclic or unbranched cyclic and bicyclic molecules. For branched cyclic and bicyclic molecules additional constraints are required [22]

$$n^L \leq \sum_{i_1 \in G_{Alicyclic}} n_i \leq \left(n^U - \left(n^U - 2\right)q\right) \tag{57}$$

$$-2q \le n_{CH_{(cyc)}} \tag{58}$$

Similarly, equations were defined for aromatic monocyclic and bicyclic molecules:

$$\sum_{i_1 \in G_{Aromatic}} n_i = 6 - 4q \tag{59}$$

$$-2q \le n_{aC} \tag{60}$$

An additional equation was defined for the number of branched chains on the cyclic molecule:

$$\sum_{i_1 \in G_{Acyclic}} n_{i_1} \le n_{branch}^U \left(\sum_{i_1 \in \{G_{Alicyclic}, G_{Aromatic} | v_{i_1} \ge 3\}} n_{i_1} \left(v_{i_1} - 2 \right) + 2q \right) \tag{61}$$

CAMD has been applied and extended in the area of molecular design for mixtures (Cignitti et al., 2015; Karunanithi et al., 2005). In addition, in the area of mixture-blend design (CAM^{b}D) where mixing rules and mixture stability criteria are used as foundation for generating novel mixtures from a set of molecules and target needs [5,6,62].