A Superstructure-Based Framework for Bio-Separation Network Synthesis

Wenzhao Wu, Kirti Yenkie, Christos T. Maravelias*

Dept. of Chemical and Biological Engineering and DOE Great Lakes Bioenergy
Research Center, University of Wisconsin-Madison, Madison, WI 53706

ABSTRACT

Modern biotechnologies enable the production of chemicals using engineered microorganisms. However, the cost of downstream recovery and purification steps is high, which means that the feasibility of bio-based chemicals production depends heavily on the synthesis of cost-effective separation networks. To this end, we develop a superstructure-based framework for bio-separation network synthesis. Based on general separation principles and insights obtained from industrial processes for specific products, we first identify four separation stages: cell treatment, product phase isolation, concentration and purification, and refinement. For each stage, we systematically implement a set of connectivity rules to develop stage-superstructures, all of which are then integrated to generate a general superstructure that accounts for all types of chemicals that can be produced using microorganisms. We further develop a superstructure reduction method to solve specific instances, based on product attributes, technology availability, case-specific considerations, and final product stream specifications. A general optimization model, including short-cut models for all technologies, is formulated. The proposed framework enables preliminary synthesis and analysis of bio-separation networks, and thus estimation of separation costs.

Keywords

Mixed integer nonlinear programming; process optimization; global optimization; renewable chemicals.

1. Introduction

The last decade has seen significant progress in metabolic engineering and synthetic biology (Tee, et al., 2014; Angermayr, et al., 2015; Knoop & Steuer, 2015; Julleson, et al., 2015; Liao, et al., 2016). These advances enable the use of engineered microorganisms such as *E.coli*, yeast and algae for the production of chemicals that are currently derived mainly from fossil fuel feedstocks (Gavrilescu & Chisti, 2005; Wilson & Roberts, 2014; Clark, et al., 2014; Bornscheuer & Nielsen, 2015; Zhang, et al., 2016). Advantages of bio-processes include mild production conditions and good selectivity toward

a specific product (Kiss, et al., 2015). However, the effluent of the bioreactor is typically dilute (containing less than 20 wt% product), and thus the downstream separation tends to be expensive, typically accounting for 60–80 % of the total production cost (Kiss, et al., 2015; Brandt & Schembecker, 2016). Thus, the synthesis of an effective downstream process is one of the major challenges (Nishida, et al., 1981; Siirola, 1996; Barnicki & Siirola, 2004; Westerberg, 2004; Li & Kraslawski, 2004; Noble & Agrawal, 2005; Grossmann & Guillen-Gosalbez, 2010; Kravanja, 2010; Cremaschi, 2015).

The synthesis of bio-separation processes is challenging for the following reasons: (1) multiple technologies are usually available for a given separation task, and thus a large number of alternative process configurations exists; (2) many bio-based chemicals require processing under mild conditions, and thus certain separation technologies (e.g., distillation) are sometimes excluded; and (3) the product physical properties and the bioreactor effluent composition are not uniform across chemicals, but rather specific. Methods used for process synthesis generally include enumeration of alternatives, evolutionary modification, and superstructure optimization (Biegler, et al., 1997). In enumeration of alternatives, alternative designs are generated and evaluated, which is only feasible when the number of alternatives is relatively small. In evolutionary modification, designers make changes to known flowsheets for similar processes to meet new objectives and constraints (King, et al., 1972; Douglas, 1985; Douglas, 1988; Smith & Linnhoff, 1988). Superstructure optimization is a model-based approach that compares alternative processes simultaneously (Grossmann & Daichendt, 1996; Biegler, et al., 1997; Barnicki & Siirola, 2004; Liu, et al., 2010; Cremaschi, 2015). A superstructure incorporates a large number of alternative units and relevant interconnections. It has been proposed for the synthesis of separation networks (Novak, et al., 1993; Eden, et al., 2004; Marquardt, et al., 2008; Dowling & Biegler, 2015), and the development of bio-refineries and bio-processes (Kokossis, et al., 2010; Kokossis & Yang, 2010; Schaber, et al., 2011; Rizwan, et al., 2013; Yuan, et al., 2013; Kelloway & Daoutidis, 2013; Kim, et al., 2013; Mascia, et al., 2013; Gonzalez-Delgado, et al., 2015; Kong, et al., 2016). However, these studies were mostly performed for specific products, on a case-by-case basis. The work presented herein focuses on the development of a general framework for bio-separation network synthesis.

In terms of superstructure representation, which is the first step in addressing the process synthesis problem, there are multiple approaches: State Task Network (STN) (Kondili, et al., 1993); State Equipment Network (SEN) (Smith, 1996; Yeomans & Grossmann, 1999); P-Graph (Friedler, et al., 1992; Friedler, et al., 1993); R-Graph (Farkas, et al., 2005); and the recently proposed Unit-Port-Conditioning Stream (UPCS) approach (Wu, et al., 2016).

Based on the selected representation approach, a superstructure is then generated, often by combining specific processes, or subsystem superstructures (e.g. reaction networks, separation

networks, heat recovery networks, etc.) each pre-designed via system-specific methodologies (Stephanopoulos & Westerberg, 1975; Kocis & Grossmann, 1989; Yee & Grossmann, 1990; Kravanja & Grossmann, 1990; Kravanja & Grossmann, 1993; Lakshmanan & Biegler, 1996; Schweiger & Floudas, 1999). There are also general rule-based systematic approaches that aim to generate simple superstructures containing all relevant structural alternatives (Friedler, et al., 1992; Friedler, et al., 1993; Tula, et al., 2015; Wu, et al., 2016).

Based on this superstructure, an optimization model is then formulated. In general, this model includes sets of equations describing units, their interconnections, equations for thermodynamic property calculations, etc. The model is often formulated as a mixed integer non-linear programming (MINLP) problem. The modeling of units can be performed in various ways, generally using simplifications, such as shortcut models (e.g. Fenske-Underwood equations for distillation columns) (Vausa & Marquardt, 2000; Caballero & Grossmann, 2001; Bao, et al., 2011) and surrogate models (Caballero & Grossmann, 2008; Caballero & Grossmann, 2008; Henao & Maravelias, 2011; Zhang & Sahinidis, 2012; Boukouvala & Ierapetritou, 2013; Agarwal & Biegler, 2013; Eason & Cremaschi, 2014; Cozad, et al., 2014; Biegler, et al., 2014; Yu, et al., 2015; Rogers & Ierapetritou, 2015). Also, in pharmaceutical product and process development, ontologies have been used for efficient model development and management (Venkatasubramanian, et al., 2006; Suresh, et al., 2010; Hailemariam & Venkatasubramanian, 2010). The solution of the optimization model identifies the best process along with the optimal operating conditions for all units. For a review of MINLP and global optimization solution algorithms, the reader is pointed to past works (Arora, et al., 1995; Tawarmalani & Sahinidis, 2004; Floudas & Chrysanthos, 2009; Burer & Letchford, 2012; Horst & Pardalos, 2013; Nallasivam, et al., 2013; Trespalacios & Grossmann, 2014; Boukouvala, et al., 2016).

In this work, we develop a framework for bio-separation network synthesis, which allows the generation of the optimal separation processes for liquid and solid chemicals that can be produced using microorganisms. The framework provides guidance on the preliminary synthesis of separation networks, thereby aiding the analysis of bio-based chemical production technologies.

The remaining of the paper is structured as follows. In the next section, we present our superstructure representation and generation approaches. In Section 3, we provide a general overview of our framework. In Section 4, we discuss the generation of stage-superstructures. In Section 5, we present the stage-superstructures. In Section 6, we discuss the general superstructure. In Section 7, we present methods to generate a reduced superstructure for specific instances. In Section 8, we discuss our modeling approach, and in Section 9, we present three case studies.

2. Background

In this work, we adopt the superstructure representation and generation methods developed by Wu et al. (Wu, et al., 2016), which are briefly summarized in Figure 1. In terms of representation (“UPCS”), three basic elements are adopted (Figure 1A): units (u), ports (p), and conditioning streams (s). A unit is indexed by a unit type (ut) and a unit number (un), e.g. “Dst,1” refers to the unit of unit type “Dst (distillation)” and unit number “1”. We also refer to unit type as “technology” hereafter. In addition to general units, source units and sink units are included as auxiliary elements whose functions are to provide raw materials and collect the final product stream and wastes. Every unit type has a predefined set of inlet and outlet ports, e.g. membrane units have one inlet port- for the feed stream, and two outlet ports- one for the permeate stream and the other for the concentrate stream. The inlet ports are generally treated as mixers, and the outlet ports are treated as splitters. Thus, a port is indexed by a unit, a port type (pt : in/out) and a port number (pn). Conditioning streams (referred to as “streams” hereafter) act as connections between outlet and inlet ports, while also performing temperature and pressure conditioning tasks. A stream is indexed by an inlet and an outlet port. Since the focus of this work is on preliminary synthesis and analysis of separation networks, we adopt simple models. Specifically, we neglect the conditioning function of streams and regard them as simple connections instead. Calculation of the stream conditioning duties is incorporated into the unit models, with simplifying assumptions. For example, the reboiler heating duty of a distillation unit is approximated as the latent heat of vapor in the column plus the energy required to heat up the feed stream from the feed temperature (usually 25 °C) to its boiling point (see Supplementary Material for example). Since each connection between ports represents a stream, we use the terms “connection” and “stream” interchangeably hereafter.

Figure 1. Superstructure representation and generation methods developed by Wu et al. (A) Representation; (B) generation. “Src”=source; “Snk”=sink. In (B), P_{in}/P_{out} denotes inlet/outlet port; $p' \in P_p$ indicates that port p' is connected to port p .

In terms of generation (see Figure 1B), four connectivity rules based on “minimal” and “feasible” component sets (components that are “required” and “allowed”, respectively, to be present in the ports during the normal operation of their units) are implemented for the generation of simple superstructures (Wu, et al., 2016).

3. Framework overview

Chemical products can be divided into three classes: commodity chemicals, fine chemicals, and specialty chemicals (Pollak, 2011; Smith, 2015). Commodity and fine chemicals are pure chemical substances with specified purity grades. Specialty chemicals are formulations of chemicals containing one or more pure chemical substances as active ingredients. Our framework does not incorporate the blending of chemicals for the formulation of specialty chemicals containing more than one pure ingredients. Instead, we assume a single chemical substance as being the *product*. Moreover, some technologies are usually suitable only for fine and specialty chemicals in industrial production (e.g. chromatography). Hence, as a distinction, we refer to fine and specialty chemicals with single ingredients as “SPC”, and commodity chemicals as “CMD”.

We consider a general effluent (i.e. the “initial product stream”) from an upstream bio-conversion unit, which typically contains cells, the product (if extracellular), large amount of water, and other impurities. It must be concentrated and purified to a specific industrial grade before entering the market (thus becoming the “final product stream”). We first identify the key stages in bio-separation processes, by investigating general process synthesis heuristics, e.g., removing the most plentiful and easiest-to-remove impurities first (Harrison, et al., 2003); general bio-separation principles (Harrison, et al., 2003; Lightfoot & Moscariello, 2004; Huang, et al., 2008); and insights obtained from industrial separation processes of specific products, such as lactic acid, citric acid, polyhydroxyalkanoate (PHA), 3-Hydroxypropionic acid (3HP), ethanol, and β -galactosidase (Burton, 1937; Inskeep, et al., 1952; Ling, et al., 1997; Horowitz & Brennan, 2001; Wasewar, 2005; Jacquiel, et al., 2008; Viiavakumar, et al., 2008; Koller, et al., 2013; SuperPro, v9.0). Consequently, four key separation stages are identified:

- **Stage 1** - cell treatment, where cells are harvested and then disrupted to release intracellular products (present if the product is intracellular; bypassed if the product is extracellular);
- **Stage 2** – product phase isolation, where the phase that contains the product is isolated;
- **Stage 3** - concentration and purification, where water and impurities are removed; and
- **Stage 4** – refinement, where the product is further refined.

Figure 2. Overview of the superstructure-based framework for bio-separation network synthesis.

Based on the four stages, we first perform a stage-wise analysis of general bio-separation processes. Then, for each stage, we systematically develop a stage-superstructure. Subsequently, all of the stage-superstructures are integrated to generate a general superstructure that accounts for all types of products. To address specific instances (e.g. when the product is extracellular, soluble, volatile, etc.), a reduction method is proposed, leading to a reduced superstructure that excludes non-applicable configurations. Finally, we formulate and solve an optimization model to identify the optimal process. These steps are summarized in Figure 2. Note that we employ a systematic method to generate the general superstructure. Therefore, when dealing with a specific case, the generation step can be bypassed (unless the general superstructure needs to be modified), and start from the reduction step.

Note that we do not explicitly account for the separation of co-products. This will be the topic of future work. However, the current framework is sufficient in some cases: e.g., if the product is volatile and the co-product is non-volatile, then a distillation unit used for separating the product (top stream) from water (bottom stream) will also remove the co-product.

4. Stage-superstructure generation

4.1 Stage-wise analysis

For each stage, we identify the following (see Figure 3):

- (1) Key physical properties of the product (referred to as “properties” hereafter), such as solubility and density, and the corresponding attributes (specifications of the properties), e.g. soluble/insoluble in water as a specification of solubility;
- (2) Key modules, which are superstructures generated for different combinations of product attributes; the combination of these modules in each stage leads to a complete stage-superstructure;
- (3) Technologies that can accomplish the separation task in the stage.

Figure 3. Stage-wise analysis of general bio-separation processes. (A)-(D): key product properties, attributes, modules and technologies for the corresponding stages. Key product properties (and the corresponding attributes) include: localization (extracellular-EX/intracellular-IN), solubility in water (soluble-SOL/insoluble-NSL), density compared to water (light-LT/heavy-HV), volatility with respect to water (volatile-VOL/non-volatile-NVL), state of matter (liquid-LQD/solid-SLD), and category (commodity chemicals-CMD/single-component specialty chemicals and fine chemicals-SPC). Technologies include flocculation (Flc), sedimentation (Sdm), flotation (Flt), centrifugation (Cnt), membrane (Mbr), filtration (Ftt), cell disruption (Cdr), cell debris differential digestion (Ddg), distillation (Dst), extraction (Ext), precipitation (Prc), solubilization (Slb), adsorption (Ads), chromatography (Chr), bleaching (Blc), crystallization (Crs), and Drying (Dry).

Relevant product properties (and the corresponding attributes) include localization (extracellular-EX/intracellular-IN), solubility in water (soluble-SOL/insoluble-NSL), density compared to water (light-LT/heavy-HV), volatility compared to water (volatile-VOL/non-volatile-NVL), state of matter at normal conditions (liquid-LQD/solid-SLD), and category (commodity chemicals-CMD/single-component specialty chemicals and fine chemicals-SPC). For simplicity, hereafter, we refer to product attributes using their abbreviations. The attributes can be used in combination to describe a product, e.g., a product that is insoluble and light is denoted by NSL LT. They can also describe modules, e.g., a module generated for an NSL LT product is also denoted by NSL LT. Note that the properties, attributes and modules relate to the product (a pure component), but not the streams containing the product (a mixture of components). For example, a SLD product can be dissolved in water and thus the product-containing stream is a liquid.

4.1.1 Stage 1 – cell treatment

In Stage 1, an EX product (secreted to the outside of the cells) requires no processing, while an IN product (located inside the cells) undergoes cell harvesting and cell disruption (see Figure 3A). Cell harvesting increases cell concentration by removing extracellular liquid (mainly water). Next, the harvested cells (assumed to be insoluble in water and have density higher than water) are disrupted to release the intracellular product; i.e., the cells are converted to the product, cell debris (assumed to be insoluble in water and have density higher than water), and other soluble components.

Accordingly, the key property determining the separation process in this stage is the localization of the product (IN/EX). Hence, only two key modules exist in Stage 1: the IN module leading to cell harvesting followed by cell disruption, and the EX module leading to no processing (i.e. bypass). Typical technologies include sedimentation, flotation, centrifugation, membrane, filtration, and cell disruption. Flocculation can also be added in the beginning to pre-treat the stream for more efficient harvesting.

4.1.2 Stage 2 – product phase isolation

In Stage 2, the phase containing the product is isolated (see Figure 3B). Distribution of the product is in general determined by three properties: localization (EX/IN), solubility in water (SOL/ NSL), and relative density with respect to water (LT/ HV). The treatment in Stage 1 guarantees that the product has been released out of the cells. Thus, the product distribution is determined by the combination of SOL/NSL and HV/LT attributes. Specifically, the product is dissolved in the water phase if it is SOL (no matter HV or LT); it floats to the top (given enough time) if the product is NSL LT; it settles to the bottom together with the cells (or debris) and other insoluble impurities if the product is NSL HV. Hence, the key modules in this stage are SOL, NSL LT, and NSL HV. For the NSL HV case, if the product is LQD, then it can be directly separated from the cells or debris using membrane or filtration. If the product is SLD, then it can be dissolved with a solubilizing agent, and thus the product is transferred into the water phase, leaving mainly the cells or debris at the bottom. Alternatively, if the product is IN SLD, differential digestion can be used to dissolve cell debris (Jacquel, et al., 2008; Koller, et al., 2013), thus leaving mainly the product at the bottom for easier separation. Technologies applicable for phase separation in this stage include flocculation, sedimentation, flotation, centrifugation, membrane, filtration, differential digestion, and solubilization.

4.1.3 Stage 3 – concentration and purification

In Stage 3, the product concentration is increased (large amount of water is removed) and impurities are removed, by utilizing the differences between the product and the other components in terms of volatility, molecular size, diffusivity, solubility in solvents, ability to precipitate, etc (see Figure 3C). The key properties in this stage are product state at normal condition (LQD/ SLD), solubility (SOL/NSL), and relative volatility with respect to water (VOL/NVL). The product state and solubility properties determine whether precipitation is a viable option: precipitation is only viable if the product is SLD SOL. Volatility determines the outlet port from which the product leaves a distillation column (either used directly, or following extraction to recover the solvents). Accordingly, we examine the SLD SOL LT and SLD SOL HV modules to account for the use of precipitation, and the VOL and NVL modules to account for the use of distillation and extraction. Relevant technologies include distillation, membrane, extraction, precipitation, and phase separation technologies (sedimentation, flotation, centrifugation, membrane, and filtration) required after precipitation. Also,

solubilization may be needed after precipitation to convert the product back to its original dissolved form. The use of sedimentation, flotation, filtration and centrifugation in this stage is excluded because these technologies have been employed already in the first two stages. On the other hand, due to the variety of membrane types (e.g. micro-, ultra- and nano-filtrations), membrane technologies are still included in this stage.

4.1.4 Stage 4 – refinement

In Stage 4, trace impurities are further removed and refinement operations are performed to satisfy special final product stream specifications, such as colorlessness, complete dryness, and crystalline form (see Figure 3D). The key properties in this stage are state of matter (LQD/SLD) and category (CMD/ SPC). Hence, the key modules in this stage are LQD CMD, SLD CMD, LQD SPC and SLD SPC. Typical technologies in this stage include adsorption, chromatography, bleaching, crystallization, drying, solubilization (to dissolve solid particles, thus allowing a solid-free stream required by the adsorption, chromatography, bleaching, and crystallization feed conditions), and sedimentation and filtration required after crystallization (since the crystal size is usually large, membrane, centrifugation and flotation are excluded). Chromatography and crystallization are usually applicable to SPC chemicals but not to CMD chemicals. In addition, crystallization and drying require the product to be SLD.

4.2 Module-superstructure generation

Based on the identified key modules in each stage, we first generate the module-superstructures using the systematic approach developed by Wu et al (see Figure 1). Specifically, we first generate a fully connected superstructure for each module based on the technologies in the corresponding stage. We assume that each unit in the fully connected superstructure can complete its assigned task, and thus we do not consider the same type of units used in series or in parallel for a single task. Then we apply the four connectivity rules to the fully connected superstructure to obtain a simplified superstructure (Wu, et al., 2016). For an illustration of the four rules see Figure 1B. An example of the simplified superstructure is shown in Figure 4A, for the IN module in Stage 1, but only flocculation, sedimentation, membrane and cell disruption are included for the sake of demonstration. Also, to simplify the presentation, unit numbers are omitted in Figure 4 and hereafter, except in the reduced superstructures to be discussed in Sections 7 and 9.

Figure 4. Simplification of the superstructure for the IN module in Stage 1, with reduced number of technologies as an example. (A) The superstructure obtained after applying four connectivity rules on the fully connected superstructure; (B) further simplification based on the six additional connectivity rules; (C) final simplified superstructure. Colored boxes represent units. Circles that are connected to units represent their ports; the port type (inlet/outlet) can be identified based on the direction of the connected streams. In (A), C^M and C^F represent minimal and feasible component sets, respectively, for the corresponding ports p1, p2, etc. In (B), the removed connections are marked with dashed lines, and the numbers of the corresponding rules applied are labeled accordingly.

Furthermore, since we are interested in a preliminary assessment, we apply the following six additional connectivity rules which are based on simplifications.

Rule 1: The outlet ports of a unit should not be connected to the unit's own inlet ports. The outlet ports of a separation unit are either product lean or product rich, and thus feeding the separated streams back to the inlet port is counter-productive. For further separation, the outlet ports should be connected to other separation units. For instance, as shown in Figure 4A, the concentration of the cells/flocculated cells in the sedimentation unit (Sdm) outlet port p7 is higher than that in the inlet port p5. Therefore, p7 should not be connected to p5. Furthermore, units that function to pre-treat streams for more efficient downstream separation (which we refer to as “pre-treatment units”), e.g. flocculation units used before phase separation, are assumed to be highly efficient, and thus recycling outgoing streams back to their own inlet ports is also considered unnecessary.

Note that in principle, each of the purged streams (see Figure 4A) can be connected to a sink unit. Similarly, source units can be used to provide additional input streams (e.g. the flocculent stream entering the flocculation unit). However, for simplicity, we use a single dummy unit that functions as a “reservoir” (“Rsv”), not shown in the superstructures, with one inlet port and one outlet port. The reservoir provides additional input streams and receive all waste streams.

Rule 2: If multiple technologies are used in series for the same task (e.g. sedimentation and centrifugation for cell harvesting), then more expensive ones (usually with higher performance such

as separation resolution and throughput) should be used after cheaper ones (e.g. centrifugation is used after sedimentation; chromatography is used after adsorption). If the cost and performance of the technologies are similar or difficult to assess (e.g., because different driving forces are used), then we sequence them based on industrial best practices, if available.

Rule 3: Filtration and membrane should not be used in series. This is because the use of filtration or membrane, as well as the type of membrane used, is pre-determined based on the characteristics (e.g. molecular/particle size) of the components to be separated.

Rule 4: Typically, only one of the outlet ports of a unit is rich in the product, therefore, only the product-rich outlet port should be connected to the inlet ports of other units, while the other outlet ports are connected to the reservoir inlet port (i.e. the streams from these ports are purged), except for the following cases:

- a) If the product is dissolved in water in the inlet port of a sedimentation, flotation or centrifugation unit (i.e. the product-rich stream will exit the unit through the outlet port for the middle water phase), then the product-lean stream (in the outlet port for the bottom phase) can also be fed to other units for further separation (see example in Section 5.2). This is because the bottom phase may still have a substantial amount of dissolved product.
- b) If a distillation unit is used to recycle solvent from an upstream extraction unit, then one outlet port of the distillation unit is connected to other units for further purification of the product-containing stream, while the other is connected to the inlet port of the extraction unit for solvent recycling.

Rule 5: Pre-treatment technologies should be placed at the beginning of each stage, to facilitate separation in the current stage, or the end, to facilitate separation in the next stage.

Rule 6: Connections that do not facilitate any processing tasks in the stage should be removed. For example, the direct connection from the flocculation unit to the cell disruption unit in Stage 1 facilitates neither the harvesting of cells, nor the disruption of the cells. Thus, it should be removed.

The implementation of these rules is demonstrated in Figure 4B and Figure 4C. Note that mixer-splitters can be used to simplify the superstructure when multiple units are fully-interconnected (Wu, et al., 2016). A mixer-splitter (represented by a diamond as shown in Figure 4C) has one inlet port and one outlet port, which is used to mix and split streams at the same time. For example, it is used after the flocculation unit in Figure 4C to represent the by-pass of the flocculation unit in a simpler fashion, instead of connecting the outlet ports of the source and flocculation units directly to the inlet ports of the sedimentation and membrane units. In addition, we name the ports connecting the upstream and downstream stages in the superstructures as “upstream/ downstream stage-

connecting ports”. For example, as shown in Figure 4A, the downstream stage-connecting port p13 at the bottom of the superstructure represents the connection to Stage 2.

4.3 Combination of module-superstructures

Next, we combine module-superstructures to generate stage-superstructures. The specific steps are as follows.

Step 1: Using the stage-connecting ports as anchoring points, align in parallel all the module-superstructures (see example in Sections 5.1-5.4).

Step 2: Further simplify the combined superstructures for Stages 3 and 4, by reducing the number of units of the same type, while ensuring that all the module-superstructures are still uniquely embedded in the simplified stage-superstructures. The specifics are discussed in Sections 5.3 and 5.4.

The combined superstructures include both units and modules. Modules are represented by dashed rounded rectangles. Also, we use dashed rectangles to indicate the product attributes to which the units enclosed by the dashed rectangles are applicable. Specific examples can be found in Sections 5.1-5.4.

To improve graphical representation of the superstructures, placement and numbering of unit ports should be arranged intuitively as follows:

(1) On each unit, we place outlet ports with lower density (or higher volatility) streams above those with higher density (or lower volatility) streams, so the function of a unit can be easily identified. For example, if the stream leaving the bottom outlet port of a membrane unit is purged (while the one leaving the top outlet port is fed to other units), then this membrane unit functions to remove the concentrate stream containing e.g., cells.

(2) An extraction unit is assumed to be a counter-current exchange system, where the solvent stream enters at the bottom port and leaves at the top port, and the product-containing stream enters at the other top port and leaves at the other bottom port.

(3) For each unit, we assign smaller numbers to ports with lower density (or higher volatility) streams, and assign larger ones to ports with higher density (or lower volatility) streams.

(4) For each unit, we number the inlet port receiving the product-containing stream as 1, and the other inlet ports as 2,3...

Specific placement and numbering of each unit port can be found in Figure B1 and Table B1 in Appendix B.

5. Stage-superstructures

5.1 Stage 1 Superstructure

After applying the methods discussed in Sections 4.2 and 4.3, the module-superstructures and the stage-superstructure for Stage 1 are shown in Figure 5A and Figure 5B, respectively. Flocculent in flocculation units only binds to cells and cell debris (no influence on the product), and is assumed to be completely consumed in the units. Different types of cell disruption are available (such as bead milling and homogenization), however, for the purpose of preliminary assessment, it is assumed that the specific type is pre-determined.

Figure 5. Module- and stage- superstructures for Stage 1. **(A)** Module-superstructures; **(B)** stage-superstructure, obtained by combining the module-superstructures in (A). Dashed rectangles indicate the attributes (in this case, EX) to which the enclosed units (in this case, none) are applicable. Dashed rounded rectangles indicate modules generated for specific combinations of attributes (in this case, IN). The same presentation is used in the subsequent figures.

5.2 Stage 2 Superstructure

The module-superstructures and the stage-superstructure for Stage 2 are shown in Figure 6A and Figure 6B, respectively.

For the sedimentation, flotation and centrifugation in the SOL module in Figure 6A, the outlet ports containing undesired particles (cells or cell debris) are also connected to other units for further separation (see Rule 4a in Section 4.2), because the corresponding stream may still contain a substantial amount of dissolved product (due to the limitation of concentrating factors).

For the NSL LT module, sedimentation and centrifugation with three outlet ports are used (API, 1989; Smith & Harriott, 1993; Ambler, 1988; SuperPro, v9.0). One outlet port (at the top of the units) is for the top phase containing the LT product, one (in the middle) is for the water phase, and the other (at the bottom) is for the bottom phase containing cells/debris. Note that flocculation is not used prior to sedimentation and centrifugation, because we assume that no cells/debris end up in the top phase, and no product ends up in the bottom phase. Thus, addition of a flocculent does not facilitate the separation of the product. Also, flotation is not considered because it will cause the cells/debris to rise to the top phase, which is counterproductive. Membrane and filtration remove cells/debris, and

thus the product-rich streams (which are cell/debris-free) from the sedimentation and centrifugation units are not fed to the membrane and filtration units. Also note that filtration can only be used when the product is LQD, because the cells/debris cannot be separated from an otherwise NSL SLD product.

For the NSL HV module, flocculation is not used prior to differential digestion, because cells/debris are to be dissolved in the digestion unit, not separated. Also, it is assumed that only the product is dissolved in solubilization units.

Figure 6. Module- and stage- superstructures for Stage 2. (A) Module-superstructures; (B) stage-superstructure, obtained by combining the module-superstructures in (A).

5.3 Stage 3 Superstructure

The module-superstructures for Stage 3 are shown in Figure 7A. It is assumed that the addition of precipitant leads to the precipitation of the product only. The combined stage-superstructure obtained after applying Step 1 in Section 4.3 is shown in Figure 7B. The final simplified Stage 3 superstructure is obtained after applying Step 2 in Section 4.3 (see Figure 7C). For example, the four membrane units shown in Figure 7B serve the same separation task, namely, isolating the product in the concentrate stream and removing the other components (mainly water) through the permeate stream. Therefore, these membrane units are combined into one (which is applicable to all product attributes), and all the connections to/from the original units are combined accordingly. Thus, as shown in Figure 7C, the inlet port of the combined membrane unit is connected to five outlet ports:

the upstream stage-connecting port, the outlet ports of the distillation units directly used for VOL and NVL products, and the outlet ports of the distillation units (for VOL and NVL products) that are used for solvent recovery after extraction. Note that the four distillation units cannot be combined, because their separation tasks are different. For instance, the distillation unit directly used for VOL products has the product-rich stream at the top while the one for the NVL products has the product-rich stream at the bottom. Accordingly, both distillation units are kept in the final Stage 3 superstructure, and they are applicable to VOL and NVL products, respectively. Note that all module-superstructures (Figure 7A) are still uniquely embedded in the simplified stage-superstructure (Figure 7C). To verify, in Figure 7C, if we eliminate all the inapplicable units for a given module (VOL, NVL, SLD SOL LT, or SLD SOL HV) and the connections to/from them, then the resulting superstructure is the same as the corresponding module-superstructure in Figure 7A.

5.4 Stage 4 Superstructure

The module-superstructures are shown in Figure 8A; the combined stage-superstructure obtained after applying Step 1 in Section 4.3 is shown in Figure 8B; the final simplified superstructure for Stage 4 is shown in Figure 8C. Solubilization is required for SLD products because adsorption, chromatography, bleaching and crystallization require solid-free feed streams. It is assumed that the type of crystallization (evaporation/cooling) is pre-selected. For simplicity, we always assume that evaporative crystallization is selected in the discussions hereafter, and thus crystallization units have two outlet ports (see Figure 8). However, if cooling is selected, then the crystallization units should have only one outlet port. A mixer-splitter is used in Figure 8C to concisely represent the different processes for SLD products (using solubilization) and LQD products (by-passing solubilization).

6. General superstructure

To generate the general superstructure, we combine all the stage-superstructures (Figure 5B, Figure 6B, Figure 7C and Figure 8C), by connecting them through the stage-connecting ports, and then include additional cross-stage connections. The final superstructure is shown in Figure 9. The specific cross-stage connections are as follows.

First, since extraction and distillation can handle solid particles (up to certain concentrations), the stream exiting Stage 1 can be directly fed to the extraction or distillation unit in Stage 3 if the product is VOL (thus bypassing Stage 2). Also, a by-pass stream is added to Stage 3.

Second, after processed in the solubilization and phase separation units in the NSL HV module in Stage 2 (see Figure 6), the NSL SLD product is dissolved and thus precipitation can be directly used afterwards in Stage 3. Accordingly, the SLD SOL HV module in Figure 7 is also applicable to SLD NSL HV products that are solubilized in Stage 2, and thus we rename the module as “SLD HV” in Figure 9.

This module only admits upstream connections from the SOL module and the left part of the NSL HV module (without differential digestion treatment) in Stage 2 (see Figure 9). Similarly, the SLD SOL LT module in Stage 3 is only connected to the SOL module in Stage 2. The output stream from the left part of the NSL HV module (without differential digestion treatment) can also bypass Stage 3 and the solubilization unit in Stage 4 entirely, because it has been pre-treated for the refinement in Stage 4.

Figure 7. Module- and stage- superstructures for Stage 3. (A) Module-superstructures; (B) combined stage-superstructure with all the module-superstructures aligned in parallel; (C) simplified stage-superstructure.

Figure 8. Module- and stage- superstructures for Stage 4. **(A)** Module-superstructures; **(B)** combined stage-superstructure with all the module-superstructures aligned in parallel; **(C)** simplified stage-superstructure. The Blc unit has only one inlet port and one outlet port, because we assume that the stream will become color-free once it passes through a Blc unit.

Third, the stream without the solubilization treatment at the end of the SLD SOL LT and SLD HV modules in Stage 3 can be directly fed to the mixer-splitter before drying in Stage 4. The streams after solubilization at the exit of these modules can bypass the solubilization unit at the beginning of Stage 4.

Note that the general superstructure can also be modified to account for customized modeling. Examples of such customization include: (1) inclusion of more technologies in each stage, such as multiple types of cell disruption in Stage 1; (2) inclusion of different connectivity rules to allow more or fewer connections in the module-superstructures; and (3) inclusion of more or fewer cross-stage connections.

Figure 9. General bio-separation superstructure.

7. Reduction method

The general superstructure discussed in Section 6 embeds millions of bio-separation configurations. Modeling and solving for such a large superstructure is both impractical and unnecessary when product-specific information is available. For example, if the product is known to be EX SOL, then all the units and modules only applicable to IN or NSL products in the general superstructure can be removed. We propose the following steps to generate a simpler superstructure (“reduced superstructure”) which yet includes all relevant configurations.

Step 1: Remove all the units and modules that are not applicable to the product attributes. For example, if the product is EX, then the IN module in Stage 1 should be removed.

Step 2: Remove all the units and modules that are not available or suitable. For example, if no extraction solvent is known for the product in question, then extraction units should be removed.

Step 3: Remove additional units, modules, or streams based on case-specific considerations. For example, if the product is EX SOL with an extremely low titer, and based on experiments we know that sedimentation is too slow and expensive even after flocculation, then sedimentation can be removed. Similarly, if we know that the concentration of solid particles (e.g. cells) after Stage 1 is greater than the handling limit of distillation, then the bypass stream from Stage 1 to the distillation unit in Stage 3 should be removed.

Step 4: Remove irrelevant units and streams based on final product stream specifications, including colorlessness, dryness and crystalline form. Specifically, if the final product stream is required to be colorless, then all the streams bypassing the bleaching unit should be removed. On the other hand, if colorlessness is not required, then the bleaching unit in Stage 4 should be removed. If the final product stream is required to be completely dry, or be in its crystalline form, then all the streams bypassing the drying unit or the crystallization unit should be removed. However, if no such requirements are specified, they should not be removed because, unlike the bleaching units, drying and crystallization can be used to concentrate the product-containing stream and remove impurities.

Step 5: Choose between filtration and membrane (if filtration is not removed from the previous steps), based on the characteristics (e.g. molecular weight or particle size) of the components to be separated. The specific type of membrane (e.g. micro-, ultra- and nano-filtration) can also be pre-determined accordingly.

Step 6: After the above five steps, we apply the following two rules iteratively to remove all the remaining irrelevant connections: (i) if a unit is removed, then all of its ports and the streams connected to its ports should be removed; (ii) if a port is not connected by any streams, then the unit to which the port belongs to should be removed. For example, if crystallization in Stage 4 is removed

in the previous steps, then based on Rule (i), all of its ports and the connected streams should be removed. Further, the downstream membrane and sedimentation should be removed based on Rule (ii).

To illustrate the proposed reduction method, we consider a product, which is EX NSL LT NVL LQD CMD. The final product stream is required to be colorless. Membrane is selected instead of filtration in Stage 2. After applying steps 1, 4, 5 and 6, the reduced superstructure is shown in Figure 10.

The superstructure thus obtained still contains redundant mixer-splitters. A final “polishing” step is performed to eliminate them, leading to the final reduced superstructure shown in Figure 11, where all units are numbered accordingly.

Figure 10. Example superstructure reduction for a EX NSL LT NVL LQD CMD product, and the final product stream is required to be colorless. The dimmed parts are removed after applying steps 1, 4, 5 and 6.

Figure 11. Final reduced superstructure for the example in Figure 10.

8. Modeling approach

We adopt the modeling approach proposed by Wu et al. (Wu, et al., 2016), which is designed to work coherently with the UPCS representation. In general, we first define the basic sets (unit types, unit numbers, port types, port numbers, and components) and three key multidimensional sets (units, ports, and streams), as well as multiple subsets to facilitate modeling (see details in Appendices A and C, and Supplementary Material). Then we employ standard short-cut unit models using port variables and unit internal variables. The port and stream models are simplified in the current work because stream conditioning is incorporated into the unit models. Moreover, we adopt the “single-stream” assumption (Kocis & Grossmann, 1989; Wu, et al., 2016) for outlet ports, which allows the activation of at most one outgoing stream; in the optimal solution at most one stream will be selected. The benefit of using single-stream outlet ports is that the nonlinear equations involving split fraction variables at an outlet port (which is treated as a splitter) are removed. Also note that we use binary attributes for superstructure generation, but continuous property values when necessary; e.g., although a product is categorized as VOL or NVL, its relative volatility is considered. Furthermore, for components that cannot be strictly classified, we can include more modules in the superstructure (e.g., for both VOL and NVL products).

Implementing the above general modeling approach, we formulate the general bio-separation optimization model as a MINLP problem, as shown in Appendix C and Supplementary Material. The model includes unit models for all the technologies shown in the general superstructure in Figure 9. Due to the modular nature of our modeling approach, modeling for the reduced superstructure is performed simply by (1) removing irrelevant unit models, and (2) adding case-specific constraints, including logic constraints inferred by the final product stream specifications, and other case-specific

considerations (e.g., no more than two units can be activated in series). Also, since this framework is used for preliminary assessment, we neither consider multiple units of the same type used in parallel due to unit size restrictions, nor do we set a minimum size for an activated unit. Instead, we calculate the equivalent total size required to complete its separation task.

The unit models are formulated based on multiple sources: membrane and filtration (McGregor, 1986; McCabe, et al., 2005); sedimentation and centrifugation (McCabe, et al., 2005; Ambler, 1988; API, 1989); extraction (Belter, et al., 1988; Smith R, 2005); distillation (Doherty & Malone, 2001; Biegler, et al., 1997; Assael & Magilliotou, 1988; Seader, et al., 2011); other units (Belter, et al., 1988; McCabe, et al., 2005; SuperPro, v9.0). A summary of these models is given in Table 1, while the specific models are presented in Supplementary Material.

Table 1. Number of variables and equations of each unit model. “c” represents the number of components in the corresponding unit.

Unit	Number of var. ^{a,c}	Number of eqn. ^{b,c}	Unit	Number of var. ^{a,c}	Number of eqn. ^{b,c}
Adsorption/chromatography	$3c+1$	$2c+1$	Extraction	$7c-2$	$6c-4$
Bleaching	$2c+1$	$c+1$	Membrane/filtration	$3c+7$	$2c+6$
Cell disruption	$2c+1$	$c+1$	Flocculation	$3c+1$	$2c+1$
Crystallization (cooling)	$2c+2$	$c+2$	Flotation ^f	$3c+c'+1$	$2c+c'-1$
Crystallization (evaporation) ^d	$3c+c'+2$	$c+c'$	Precipitation	$3c+1$	$2c+1$
Differential digestion	$3c+1$	$2c+1$	Sedimentation/centrifugation ^f	$4c+c'+1$	$3c+c'-2$
Distillation ^e	$4c+c'+9$	$3c+c'+9$	Solubilization	$3c+1$	$2c+1$
Drying ^d	$3c+c'+2$	$2c+2$			

(a) Var.=variables;

(b) Eqn.=equations;

(c) The counts exclude variables and equations associated with costs (see Appendix C) of units; Equations C2, C3 and C4 for the port models in Appendix C are also included when mass fraction variables $X_{p,c}$ are used in the unit models (see Supplementary Material);

(d) c': number of components more volatile than the product;

(e) c': number of components less volatile than the light key, and more volatile than the heavy key;

(f) c': number of insoluble components in the unit.

9. Case studies

In this section, we present three case studies. Each case deals with a different set of product attributes, technology availability, case-specific considerations, and final product stream specifications. The common objective is to minimize the total cost for purifying a dilute initial product stream to a final product stream containing 500 kg/hour product, with at least 90 wt% purity. The initial stream contains ~5 wt% product, and 5 wt% cells if the product is EX; ~10 wt% cells (with 50 wt% intracellular product) if the product is IN. It also contains other impurities, which may be removed to satisfy the minimum purity requirement. Impurities in a different phase than the product (e.g., insoluble impurities and a soluble product) can be removed by phase separation technologies;

impurities in the same phase as the product can be removed by technologies such as extraction, distillation, adsorption and precipitation, which utilize differences in physical properties other than solubility. In terms of the choice between filtration and membrane, we assume that a specific type of membrane is selected. The superstructure model includes: equations for ports and streams, general logic constraints, cost equations, and initial and final product stream specifications (see Appendix C), and unit models (see Supplementary Material). We consider only models of units that exist in the reduced superstructure. Case-specific logic constraints and parameter values adopted for these cases can be found in Supplementary Material. The optimization model is generated using GAMS 24.4.6, and solved with BARON (Tawarmalani & Sahinidis, 2005), in a PC with an Intel (R) Core (TM) i5-3570 CPU @ 3.40 GHz.

9.1. Case 1 - EX NSL LT NVL LQD CMD product

We continue with the example discussed previously in Section 7 (Figure 11), where the product is EX NSL LT NVL LQD CMD, and the final product stream is required to be colorless. After reduction (see Figure 10), the reduced superstructure includes 9 units (see Figure 12A).

Figure 12. Case 1. (A) Reduced superstructure, and the optimal process (the corresponding ports and streams are marked red, and the streams are thickened); (B) cost distribution in the optimal solution. The product is EX NSL LT NVL LQD CMD, and the final product stream is required to be colorless

The model has a total of 726 variables, 44 binary variables and 804 constraints. The time required to solve to 5 % gap (optcr) is 5 s, while 112 s are required to solve to optimality. The optimal process is shown in Figure 12A. The solution involves (Cnt,1) with a sedimentation-equivalent area (“sigma factor”) of 1383 m², and (Blc,1) with volume 0.56 m³. The objective value (total annualized cost) is 1.6 million \$/year, resulting in a unit cost of 0.44 \$/kg product (pure product basis), with 36 % of the cost incurring in Stage 2, and 64 % in Stage 4. The specific cost distribution, including annualized capital cost (ACC, \$/year) and operating cost (OC, \$/year) of the selected units, is shown in Figure 12B. The product is NSL LT and thus will naturally form a separate phase from water and cells.

Therefore, in this case (based on the specific parameter values we assume for membrane rejection coefficient, product settling velocity, etc.), centrifugation outperforms membrane in Stage 2. Since the product settling velocity is relatively small, centrifugation is preferred to sedimentation due to lower capital cost. Stage 3 and the adsorption unit in Stage 4 are bypassed because the product concentration requirement is met after centrifugation. However, the optimal process may change given a different set of parameter values such as rejection coefficient, settling velocity, solvent partition coefficient, and equipment costs. The same note applies to all examples.

9.2. Case 2 - IN NSL HV NVL SLD SPC product

We study a IN NSL HV NVL SLD SPC product. The final product stream is required to be dry (containing at most 1 wt% water). Extraction is excluded because no suitable solvent is known for the product. Additional case-specific considerations are that centrifugation, distillation and precipitation should not be considered; sedimentation should not be used after crystallization; and the refinement stage cannot be bypassed. After reduction (see Figure 13A), the reduced superstructure includes 21 units.

Figure 13. Case 2. (A) Reduced superstructure and the optimal process (the corresponding ports and streams are marked red, and the streams are also thickened); (B) cost distribution in the optimal solution. The product is IN NSL HV NVL SLD SPC, and the final product stream is required to be dry (containing at most 1 wt% water). Extraction is excluded because no suitable solvent is known for the product. Additional case-specific considerations are that centrifugation, distillation and precipitation should not be considered; sedimentation should not be used after crystallization; and the refinement stage cannot be bypassed.

The model has a total of 2548 variables, 95 binary variables and 2868 constraints. The time required to solve to 5 % gap is 739 s, while 2120 s are required to solve to optimality. The optimal process is shown in Figure 13A. The solution involves (Flc,1) with volume 2.0 m³, (Mbr,1) with area 768 m²,

(Cdr,1) with volume 0.3 m³, (Flc,2) with volume 0.8 m³, (Slb,1) with volume 0.8 m³, (Mbr,2) with area 650 m², (Crs,1) with volume 5.4 m³, and (Dry,1) with volume 3.6 m³. The objective value (total annualized cost) is 55 million \$/year, resulting in a unit cost of 12.6 \$/kg product (pure product basis). The costs incurred in Stage 1, Stage 2 and Stage 4 are 28 %, 61 % and 11 % of the total cost, respectively. The specific cost distribution is shown in Figure 13B. In Stage 1, flocculation is first used to increase the cell particle size for a more efficient phase separation. Membrane is preferred to sedimentation and flotation because the cells have low settling velocity while the membrane rejection coefficient is relatively large. In stage 2, solubilization has lower cost than differential digestion because the product is not only separated from the cell debris but also solubilized after Stage 2, and thus the solubilization unit (Slb,2) at the beginning of Stage 4 can be bypassed to reduce cost. Otherwise, (Slb,2) has to be activated if (Ddg,1) is activated.

9.3. Case 3 - EX SOL VOL LQD CMD product

We study a product that is EX SOL VOL LQD CMD, with no special final product stream requirement. Flocculation, extraction, sedimentation and flotation are excluded based on technology availability and case-specific considerations. After reduction, the reduced superstructure includes 5 units (see Figure 14A). The model has a total of 383 variables, 25 binary variables and 414 constraints. The time required to solve to 5 % gap is 281 s, while 437 s are required to solve to optimality. The optimal process is shown in Figure 14A. The solution involves only (Dst,2) in Stage 3 (volume=18.9 m³). The objective value (total annualized cost) is 1.6 million \$/year, resulting in a unit cost of 0.4 \$/kg product (pure product basis). The specific cost distribution is shown in Figure 14B. A direct distillation is preferred (bypassing Stage 2) because the product is VOL and thus there is no need to first separate the product from cells using the phase separation in Stage 2 and then perform the concentration step in Stage 3. Also, the final product specification is met after Stage 3, and thus Stage 4 is bypassed.

Figure 14. Case 3. (A) Reduced superstructure and the optimal process (the corresponding ports and streams are marked red, and the streams are also thickened); (B) cost distribution in the optimal solution. The product is EX SOL VOL LQD CMD, with no special final product stream requirement. Flocculation, extraction, sedimentation and flotation are excluded based on technology availability and case-specific considerations.

10. Conclusions

In this work, we developed a superstructure-based framework for bio-separation network synthesis. We first identified four stages: cell treatment, product phase isolation, concentration and purification, and refinement. Then, for all four stages, we systematically developed stage-superstructures, which, together with cross-stage connections, leads to the generation of the general bio-separation superstructure. We further developed a superstructure reduction method to solve case-specific problems, based on product attributes, technology availability, case-specific considerations, and final product stream specifications. An MINLP optimization model, including short-cut models for all technologies considered in the framework was formulated. The proposed framework can be used to quickly evaluate separation processes for different bioreactor effluent streams and different products, thereby aiding the identification of chemicals that can be produced effectively using bioconversions.

Supplementary Material

Supplementary material for this work is available on line.

Acknowledgements

This work was funded by National Science Foundation through the Emerging Frontiers in Research and Innovation program (EFRI-1240268), and the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494).

Appendix A – Component set specification

We first present a method to specify the component set ($\mathbf{C}$) for the reduced superstructure. In general, if the product is EX, then the initial product stream contains the product, water, cells, impurities that are soluble in water, and impurities that are insoluble in water. If the product is IN, then the initial product stream contains water and cells (see columns “Basic” in Table A1). We refer to the set composed of these elements as the “basic component set” ($\mathbf{C}^{\text{Basic}}$). The existence of flocculation, cell disruption, precipitation, extraction, solubilization, crystallization, and differential digestion technologies in the reduced superstructure will lead to the introduction of new elements (see Table A1). For example, in cell disruption, the cells are converted to product (“prodt”), insoluble debris (“debris”) and other soluble impurities (“solimp”). Thus, if cell disruption is included in the reduced superstructure, then “prodt”, “solimp” and “debris” should be added to the component set. Mathematically, it is expressed as follows:

$$\mathbf{C} = \mathbf{C}^{\text{Basic}} \cup \mathbf{C}^{\text{Flc}} \cup \mathbf{C}^{\text{Cdr}} \cup \dots \quad (\text{A1})$$

where $\mathbf{C}^{\text{Flc}}$, $\mathbf{C}^{\text{Cdr}}$... represent sets of components specific for technologies Flt, Cdr, etc. Note that the components in Table A1 can be further modified, e.g. “solimp” can be replaced by “proteins”, “pigments”, etc.

Table A1. Component set. Prodt=product, solimp=impurities soluble in water, insolimp=impurities insoluble in water, floct=flocculent, cellfc=flocculated cells, debrisfc= flocculated debris, prodtprc=precipitated product, prcpt=precipitant, extsolv=solvent used for extraction, prodtsol=solubilized product, solbr=solubilizing agent, prodtcrs=crystallized product, digr=digesting agent for the cell debris, digdebris=digested debris. Components for the example case in Figure 11 are marked bold.

	$\mathbf{C}^{\text{Basic}}$	$\mathbf{C}^{\text{Flc}}$	$\mathbf{C}^{\text{Cdr}}$	$\mathbf{C}^{\text{Prc}}$	$\mathbf{C}^{\text{Ext}}$	$\mathbf{C}^{\text{Slb}}$	$\mathbf{C}^{\text{Crs}}$	$\mathbf{C}^{\text{Ddg}}$
EX	prodt, water, cell, solimp, insolimp	floct, cellfc	NA	prodtprc, prcpt	extsolv	prodtsol, solbr	prodtcrs	NA
IN	water, cell	floct, cellfc, debrisfc	prodt, solimp, debris	prodtprc, prcpt	extsolv	prodtsol, solbr	prodtcrs	digr, digdebris

Appendix B – Port numbering

The specific numbering of ports is shown in Table B1. Note that the insoluble components in flotation units float to the top, and thus the corresponding port is assigned a smaller number.

Table B1: Unit port numbering.

unit type <i>ut</i>	port type <i>pt</i>	port No. <i>pn</i>	port usage	unit type <i>ut</i>	port type <i>pt</i>	port No. <i>pn</i>	port usage
Cdr	in	1	feed stream	Sdm/Cnt	in	1	feed stream
	out	1	disrupted cells			1	light phase
Blc	in	1	feed stream		out	2	water
	out	1	bleached stream			3	heavy phase
Flt	in	1	feed stream	Flc	in	1	feed stream
	out	1	insoluble compoents		out	2	flocculent
Mbr/Ftt		2	water			1	flocs
	in	1	feed stream	Slb	in	1	feed stream
	out	1	permeate		out	2	solvent
		2	concentrate			1	solubilized product
Dst	in	1	feed stream	Ddg	in	1	feed stream
	out	1	top		out	2	digesting agent
		2	bottom			1	digested debris
Ads/Chr	in	1	feed stream	Prc	in	1	feed stream
	out	1	non-product		out	2	precipitant
		2	product			1	precipitated product
Crs	in	1	feed stream	Ext	in	1	feed stream
	out	1	vapor/liquid		out	2	solvent
		2	crystal		out	1	extract
Dry	in	1	feed stream			2	raffinate
	out	1	vapor	Src	out	1	initial product stream
		2	solid	Snk	in	1	final product stream

Figure B1: New labels for unit ports. The boxes represent technologies. The “a,b” label in each box indicates the number of inlet ports (“a”) and outlet ports (“b”) of the corresponding technology. The corresponding technologies with the specified number of ports are labeled below the boxes. “pn=in/out,m” labels around the ports mean that the inlet/outlet port with identification number m is renamed as “pn” for the corresponding unit type. For example, the third box on the left has one inlet port (“in,1”) and one outlet port (“out,1”), which corresponds to Cdr and Blc. The ports are renamed “p1” and “p2”, respectively. If crystallization is performed by cooling instead of evaporation, then Crs has one outlet port.

To facilitate reading of the unit models, new port labels (p1, p2...) are assigned, as shown in Figure B1.

Appendix C –Bio-separation model components

We present general equations for ports and streams, general logic constraints, cost equations, and initial and final product stream specifications. Only membrane, filtration and extraction unit models are included here. The other specific unit models are presented in Supplementary Material, including models for sedimentation, centrifugation, distillation, flocculation, flotation, cell disruption, differential digestion, precipitation, solubilization, adsorption, chromatography, bleaching, crystallization, and drying.

Indices and sets

$ut \in \mathbf{UT}$	=	unit types
$un \in \mathbf{UN}$	=	unit numbers
$u \in \mathbf{U} \subset \mathbf{UT} \times \mathbf{UN}$	=	units
$pt \in \mathbf{PT}$	=	port types, i.e. in/out
$pn \in \mathbf{PN}$	=	port numbers
$p \in \mathbf{P} \subset \mathbf{U} \times \mathbf{PT} \times \mathbf{PN}$	=	ports
$s \in \mathbf{S} \subset \mathbf{P} \times \mathbf{P}$	=	streams defined by two ports
$c \in \mathbf{C}$	=	components
$l \in \mathbf{L}$	=	utility types, i.e. cooling/heating/electricity/labor

Subsets

$\mathbf{U}_{ut}/\mathbf{U}_s$	=	units of type ut / connected to stream s
$\mathbf{P}_{u,pt}$	=	ports belonging to unit u and of type pt
$\mathbf{S}_{pt}$	=	streams connected to type pt ports
$\mathbf{S}_p$	=	streams connected to port p
$\mathbf{C}_u^{\text{SOL}}/\mathbf{C}_u^{\text{SOLV}}$	=	solute/solvent components in unit $u \in \mathbf{U}_{\text{Ext}}$

The model equations are presented in their native form below. They are reformulated in the modeling implementation stage, so that e.g., no variables are in the denominators.

Ports & Streams

We use the following variables:

F_s/F_p	=	mass flow rate [kg/hr] of stream s /in port p
$F_{s,c}/F_{p,c}$	=	mass flow rate [kg/hr] of component c in stream s /port p
$X_{p,c}$	=	mass fraction [wt%] of component c in port p

Mass balance in ports:

$$\sum_{s \in \mathbf{S}_p} F_{s,c} = F_{p,c}, \quad p, c \quad (\text{C1})$$

Component & total flow relation in ports:

$$\sum_c F_{p,c} = F_p, \quad p \quad (C2)$$

Mass fraction definition in ports:

$$F_{p,c} = F_p X_{p,c}, \quad p, c = c_1, \dots, c_{|C|-1} \quad (C3)$$

Mass fractions summation to one:

$$\sum_c X_{p,c} = 1, \quad p \quad (C4)$$

Note that Equations C3 and C4 are only formulated for ports that have $X_{p,c}$ variables in the corresponding unit models. For example, $X_{p2,c}$ (mass fraction of component c in the permeate port p2) is used in the membrane & filtration models to define rejection coefficient, while $X_{p1,c}$ (mass fraction of component c in the feed port p1) does not appear in any equations. Thus, Equations C3 and C4 are formulated for port p2, but not formulated for port p1.

Component & total flow relation in streams:

$$\sum_c F_{s,c} = F_s, \quad s \quad (C5)$$

Logic constraints

For logic constraints, we use the following parameters and variables:

$$\begin{aligned} f^{\text{up}} &= \text{global flow rate upper bound [kg/hr]} \\ Y_u/Y_s &= 1 \text{ if unit } u/\text{stream } s \text{ is activated, otherwise } 0 \end{aligned}$$

Single-stream outlet port assumption:

$$Y_u = \sum_{s \in \mathbf{S}_p} Y_s, \quad u, p \in \mathbf{P}_{u,\text{out}} \quad (C6)$$

General logic constraints:

$$Y_u \leq \sum_{s \in \mathbf{S}_p} Y_s, \quad u, p \in \mathbf{P}_{u,\text{in}} \quad (C7)$$

$$Y_s \leq Y_u, \quad s \in \mathbf{S}_{\text{in}}, u \in \mathbf{U}_s$$

Zero-flow enforcement for deactivated streams:

$$F_{sc} \leq f^{\text{up}} Y_s, \quad c, s \in \mathbf{S} \quad (C8)$$

Extraction

For $u \in \mathbf{U}_{\text{Ext}}$, we use the following parameters:

$$\begin{aligned} \rho_c &= \text{density [kg/m}^3\text{] of component } c \\ \tau_u &= \text{residence time [hr] of unit } u \end{aligned}$$

$$\begin{aligned}
h &= \text{minimum fraction of fresh solvent in port p2} \\
\varepsilon_{cc'} &= \text{component } c \text{ solubility [kg/m}^3\text{] in } c' \\
b_{u,c} &= \text{partition coefficient of component } c \text{ in unit } u \\
e_{u,c}^{\text{lo}}/e_{u,c}^{\text{up}} &= \text{extraction coefficient lower/upper bound of component } c \text{ in unit } u
\end{aligned}$$

and the following variables:

$$\begin{aligned}
E_{u,c} &= \text{extraction coefficient of component } c \text{ in unit } u \\
M_u^{\text{ext}} &= \text{number of stages of unit } u \\
A_u &= \text{size of unit } u
\end{aligned}$$

Extraction coefficient definition:

$$E_{u,c} = \frac{b_{u,c} F_{p2,c'}}{F_{p1,\text{water}}}, \quad c \in \mathbf{C}_u^{\text{SOL}}, c' \in \mathbf{C}_u^{\text{SOLV}} \quad (\text{C9})$$

Pure solvent feed assumption:

$$F_{\text{rsv},1,\text{out},1,p2,c} = 0, \quad c \notin \mathbf{C}_u^{\text{SOLV}} \quad (\text{C10})$$

Minimum fresh solvent fraction:

$$F_{\text{rsv},1,\text{out},1,p2,c} \geq h F_{p2,c}, \quad c \in \mathbf{C}_u^{\text{SOLV}} \quad (\text{C11})$$

Extraction coefficient bounds:

$$e_{u,c}^{\text{lo}} \leq E_{u,c} \leq e_{u,c}^{\text{up}} \quad (\text{C12})$$

Concentrating effect relation:

$$X_{p1,c} = \left(\frac{E_{u,c} M_u^{\text{ext}} + 1}{E_{u,c} - 1} \right) X_{p4,c}, \quad c \in \mathbf{C}_u^{\text{SOL}} \quad (\text{C13})$$

Water distribution to solvent phase:

$$F_{p3,\text{water}} = \varepsilon_{\text{water},c} \frac{F_{p3,c}}{\rho_c}, \quad c \in \mathbf{C}_u^{\text{SOLV}} \quad (\text{C14})$$

Solvent distribution to water phase:

$$F_{p4,c} = \varepsilon_{c,\text{water}} \frac{F_{p4,\text{water}}}{\rho_{\text{water}}}, \quad c \in \mathbf{C}_u^{\text{SOLV}} \quad (\text{C15})$$

Unit volume:

$$A_u = \tau_u M_u^{\text{ext}} \left(\sum_c \frac{F_{p4,c}}{\rho_c} + \sum_{c'} \frac{F_{p3,c'}}{\rho_{c'}} \right) \quad (\text{C16})$$

Membrane & filtration

For $u \in \mathbf{U}_{\text{Mbr}} \cup \mathbf{U}_{\text{Ftt}}$, we use the following parameters:

$$\begin{aligned} \gamma_{u,c} &= \text{rejection coefficient of component } c \text{ in unit } u \\ \omega_u &= \text{flux [m/hr] of unit } u \\ n_u^{\text{lo}}/n_u^{\text{up}} &= \text{lower/upper bound of concentrating factor of unit } u \end{aligned}$$

and the following variables:

$$N_u = \text{concentrating factor of unit } u$$

Concentrating factor definition:

$$N_u = \frac{F_{p1}}{F_{p3}} \quad (\text{C17})$$

Concentrating factor bounds:

$$n_u^{\text{lo}} \leq N_u \leq n_u^{\text{up}} \quad (\text{C18})$$

Rejection coefficient definition:

$$\gamma_{u,c} = 1 - \frac{X_{p2,c}}{X_{p3,c}}, \quad c \in \{c | \gamma_{uc} > 0\} \quad (\text{C19})$$

Non-rejecting components distribution:

$$\begin{aligned} \frac{F_{p3,c}}{\sum_{c'' \in \{c | \gamma_{u,c''} = 0\}} F_{p3,c''}} &= \frac{F_{p1,c}}{\sum_{c' \in \{c' | \gamma_{u,c'} = 0\}} F_{p1,c'}}, \\ c &\in \{c | \gamma_{u,c} = 0\}, c = c_1, \dots, c_{|C|-1} \end{aligned} \quad (\text{C20})$$

Unit area:

$$A_u = \frac{N_u - 1}{\omega_u} \sum_c \frac{F_{p1,c}}{\rho_c} \quad (\text{C21})$$

Costs

For cost calculations, we use the following parameters:

$$\begin{aligned} o &= \text{operating hours [hr] per year} \\ \delta_u^C / \delta_u^S &= \text{reference cost [\$]/size of unit } u \\ \sigma^{\text{pc_cc}} &= \text{conversion factor of purchase cost to total capital cost} \\ \sigma^{\text{mat_oc}} / \sigma^{\text{util_oc}} &= \text{contribution factor of material/utility in operating cost} \\ \sigma^{\text{labr_oc}} / \sigma^{\text{pc_oc}} / \sigma^{\text{cons_oc}} &= \text{contribution factor of labor/purchase/consumable in operating cost} \\ \kappa_c / \kappa_l &= \text{purchase price [$/kg] of component } c / \text{price [$/KWh or $/hr] of utility } l \\ v_u^{\text{elec}} / v_u^{\text{labr}} &= \text{electricity [KWh/m}^2\text{] / labor [hr / (unit \cdot \text{operating hr})] requirement of unit } u \end{aligned}$$

$$\begin{aligned}
sz_u^{\max} &= \text{maximum size of unit } u \text{ an operator can handle, only used for} \\
&\quad \text{labor cost calculation, not for equipment sizing} \\
cs_u^{\text{price}} &= \text{consumable price [$/ (size \cdot \text{operating hr})] of unit } u
\end{aligned}$$

and the following variables:

$$\begin{aligned}
C_u^{\text{pc}} &= \text{purchase cost [\$] of unit } u \\
C_u^{\text{acc}}/C_u^{\text{labr}}/C_u^{\text{cons}} &= \text{capital/labor/consumable annualized cost [$/yr] of unit } u \\
C_u^{\text{util}}/C_u^{\text{oc}}/C_u^{\text{atc}} &= \text{utility/operating/total annualized cost [$/yr] of unit } u \\
U &= \text{total annual cost [\$]}
\end{aligned}$$

Operating costs usually involve components such as overhead, maintenance, lab cost, insurance, administration, etc (Vasudevan & Ulrich, 2004; Peters, et al., 2005), which are estimated as a percentage of other costs (fixed capital cost, labor cost, utility cost, etc.). Thus, we express the total operating cost as a linear combination of these other costs, using the σ coefficients.

See Figure 3 caption for abbreviations of unit types, such as Dst (distillation), Cnt (centrifugation) and Mbr (membrane).

Unit purchase cost:

$$C_u^{\text{pc}} = \delta_u^{\text{c}} \left(\frac{A_u}{\delta_u^{\text{s}}} \right)^{0.6}, \quad u \quad (\text{C22})$$

Unit annualized capital cost (ACC):

$$C_u^{\text{acc}} = \sigma^{\text{pc-cc}} \beta C_u^{\text{pc}}, \quad u \quad (\text{C23})$$

Dst utility cost:

$$C_u^{\text{util}} = oQ_u \sum_l \kappa_l \quad (\text{C24})$$

Cnt & Cdr utility cost:

$$C_u^{\text{util}} = o\kappa_{\text{elec}} v_u^{\text{elec}} A_u \quad (\text{C25})$$

Crs (cooling) utility cost:

$$C_u^{\text{util}} = oQ_u \kappa_{\text{cool}} \quad (\text{C26})$$

Crs (evaporation) & Dry utility cost:

$$C_u^{\text{util}} = oQ_u \kappa_{\text{steam}} \quad (\text{C27})$$

Mbr, Ftt, Ads, Chr & Blc consumable cost:

$$C_u^{\text{cons}} = ocs_u^{\text{price}} A_u \quad (\text{C28})$$

Flc, Slb, Ddg & Ext material cost:

$$C_u^{\text{mat}} = o \sum_c \kappa_c F_{s,c}, \quad s \in \mathbf{S}_{\text{Rsv},1,\text{out},1,u,\text{in},2} \quad (\text{C29})$$

Unit Labor cost:

$$C_u^{\text{labr}} = o \kappa_{\text{labr}} v_u^{\text{labr}} \frac{A_u}{SZ_u^{\text{max}}} \quad (\text{C30})$$

Cdr, Dst, Crs, Cnt & Dry operating cost:

$$C_u^{\text{oc}} = \sigma^{\text{util_oc}} C_u^{\text{util}} + \sigma^{\text{labr_oc}} C_u^{\text{labr}} + \sigma^{\text{pc_oc}} C_u^{\text{pc}} \quad (\text{C31})$$

Mbr, Ftt, Ads, Chr & Blc operating cost:

$$C_u^{\text{oc}} = \sigma^{\text{cons_oc}} C_u^{\text{cons}} + \sigma^{\text{labr_oc}} C_u^{\text{labr}} + \sigma^{\text{pc_oc}} C_u^{\text{pc}} \quad (\text{C32})$$

Flc, Slb, Ddg & Ext operating cost:

$$C_u^{\text{oc}} = \sigma^{\text{mat_oc}} C_u^{\text{mat}} + \sigma^{\text{labr_oc}} C_u^{\text{labr}} + \sigma^{\text{pc_oc}} C_u^{\text{pc}} \quad (\text{C33})$$

Sdm & Flt operating cost:

$$C_u^{\text{oc}} = \sigma^{\text{labr_oc}} C_u^{\text{labr}} + \sigma^{\text{pc_oc}} C_u^{\text{pc}} \quad (\text{C34})$$

Unit annual total cost:

$$C_u^{\text{atc}} = C_u^{\text{acc}} + C_u^{\text{oc}} \quad \forall u \quad (\text{C35})$$

Src,1 specifications

For the source unit $u \in \mathbf{U}_{\text{Src},1}$, we use the following parameters:

$$x_c^{\text{src}} = \text{source mass fraction [wt\%] of component } c$$

Source flow composition specification:

$$X_{p1,c} = x_c^{\text{src}} \quad \forall c \quad (\text{C36})$$

Snk,1 specifications

For the sink unit $u \in \mathbf{U}_{\text{Snk},1}$, we use the following parameters:

$$\begin{aligned} z &= \text{final product purity [wt\%]} \\ \zeta &= \text{minimum overall product recovery [wt\%]} \\ f^{\text{snk}} &= \text{sink product flow rate [kg/hr]} \\ \mu^{\text{prodt}} &= \text{mass fraction of product in cells} \end{aligned}$$

Note that the minimum recovery value ζ should be small, only to provide a valid (but not necessarily tight) lower bound.

Final product flow specification:

$$\sum_{c \in \mathbb{C}^{\text{prodt}}} F_{p1,c} = f^{\text{snk}} \quad (\text{C37})$$

Final product stream purity specification:

$$\sum_{c \in \mathbb{C}^{\text{prodt}}} F_{p1,c} \geq z F_{p1} \quad (\text{C38})$$

Overall product recovery specification:

$$\text{For EX product: } \sum_{c \in \mathbb{C}^{\text{prodt}}} F_{p1,c} \geq \zeta F_{\text{src},1,\text{out},1,\text{prodt}} \quad (\text{C39})$$

$$\text{For IN product: } \sum_{c \in \mathbb{C}^{\text{prodt}}} F_{p1,c} \geq \zeta \mu^{\text{prodt}} F_{\text{src},1,\text{out},1,\text{cell}}$$

Global bounds

$$F_{s,c} \leq f^{\text{up}} \quad (\text{C40})$$

$$X_{p,c} \leq 1 \quad (\text{C41})$$

Note that f^{up} is a valid (but not necessarily tight) upper bound (see the calculation in Supplementary Material).

References

- Agarwal, A. & Biegler, L., 2013. A trust-region framework for constrained optimization using reduced order modeling. *Optim & Eng*, 14(1), pp. 3-35.
- Ambler, C., 1988. *Handbook of separation techniques for chemical engineers (2nd edition)*. New York: MacGraw-Hill Education.
- Angermayr, S. A., Gorchs Rovira, A. & Hellingwerf, K. J., 2015. Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends Biotechnol*, Volume 33, pp. 352-361.
- API, 1989. *Monographs on refinery environmental control– management of water discharges– design and operation of oil-water separators*. Washington: American Petroleum Institute.
- Arora, J. S., Elwakeil, O. A., Chahande, A. I. & Hsieh, C. C., 1995. Global optimization methods for engineering applications: a review. *Structural optimization*, 9(3), pp. 137-159.
- Assael, M. & Magilliotou, M., 1988. *Introduction to unit operations*. Greece: Aristotle University Publications.
- Bao, B. et al., 2011. A shortcut method for the preliminary synthesis of process-technology pathways: an optimization approach and application for the conceptual design of integrated biorefineries. *Comp & Chem Eng*, 35(8), pp. 1374-1383.
- Barnicki, S. D. & Siirola, J. J., 2004. Process synthesis prospective. *Computers & Chemical Engineering*, 28(4), pp. 441-446.
- Belter, P., Cussler, E. & Hu, W., 1988. *Bio-separations- downstream processing for biotechnology*. New York: John Wiley & Sons.
- Biegler, L., Grossmann, I. & Westerberg, A., 1997. *Systematic methods of chemical process design*. New Jersey: Prentice Hall PTR.
- Biegler, L., Lang, Y. & Lin, W., 2014. Multi-scale optimization for process systems engineering. *Computers & Chemical Engineering*, Volume 60, pp. 17-30.
- Bornscheuer, U. T. & Nielsen, A. T., 2015. Editorial overview: chemical biotechnology: interdisciplinary concepts for modern biotechnological production of biochemicals and biofuels. *Current Opinion in Biotechnology*, Volume 35, pp. 133-134.
- Boukouvala, F. & Ierapetritou, M. G., 2013. Surrogate-based optimization of expensive flowsheet modeling for continuous pharmaceutical manufacturing. *J Pharm Innov*, Volume 8, pp. 131-145.
- Boukouvala, F., Misener, R. & Floudas, C. A., 2016. Global Optimization Advances in Mixed-Integer Nonlinear Programming, MINLP, and Constrained Derivative-Free Optimization, CDFO. *European Journal of Operational Research*, 252(3), pp. 701-727.
- Brandt, K. & Schembecker, G., 2016. Production Rate-Dependent Key Performance Indicators for a Systematic Design of Biochemical Downstream Processes. *Chemical Engineering & Technology*, 39(2), pp. 354-364.

- Burer, S. & Letchford, A. N., 2012. Non-convex mixed-integer nonlinear programming: a survey. *Surveys in Operations Research and Management Science*, 17(2), pp. 97-106.
- Burton, L., 1937. Byproducts of milk. *Food Ind*, Volume 9, pp. 571-575.
- Caballero, J. A. & Grossmann, I., 2001. Generalized disjunctive programming model for the optimal synthesis of thermally linked distillation columns. *Ind Eng Chem Res*, 40(10), pp. 2260-2274.
- Caballero, J. & Grossmann, I., 2008. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE J*, 54(10), pp. 2633-2650.
- Caballero, J. & Grossmann, I., 2008. Rigorous flowsheet optimization using process simulators and surrogate models. *Computer Aided Chemical Engineering*, Volume 25, p. 551.
- Clark, R., Cameron, J., Root, T. & Pflieger, B., 2014. Insights into the industrial growth of cyanobacteria from a model of the carbon-concentrating mechanism. *AIChE J*, 60(4), pp. 1269-1277.
- Cozad, A., Sahinidis, N. V. & Miller, D. C., 2014. Learning surrogate models for simulation-based optimization. *AIChE J*, 60(6), pp. 2211-2227.
- Cremaschi, S., 2015. A perspective on process synthesis: Challenges and prospects. *Computers & Chemical Engineering*, Volume 81, pp. 130-137.
- Doherty, M. & Malone, M., 2001. *Conceptual Design of Distillation Systems*. New York: McGraw-Hill.
- Douglas, J., 1988. *Conceptual design of chemical processes*. New York: McGraw-Hill.
- Douglas, J. M., 1985. A hierarchical decision procedure for process synthesis. *AIChE J*, 31(3), pp. 353-362.
- Dowling, A. & Biegler, L., 2015. A framework for efficient large scale equation-oriented flowsheet optimization. *Comp Chem Eng*, Volume 72, pp. 3-20.
- Eason, J. & Cremaschi, S., 2014. Adaptive sequential sampling for surrogate model generation with artificial neural networks. *Comput Chem Eng*, Volume 68, pp. 220-232.
- Eden, M., Jorgensen, S., Gani, R. & El-Halwagi, M., 2004. A novel framework for simultaneous separation process and product design. *Chem Eng Process*, 43(5), pp. 595-608.
- Farkas, T., Rev, E. & Lelkes, Z., 2005. Process flowsheet superstructures: structural multiplicity and redundancy part I: basic GDP and MINLP representations. *Comput Chem Eng*, 29(10), pp. 2180-2197.
- Floudas, C. A. & Chrysanthos, E. G., 2009. A review of recent advances in global optimization. *Journal of Global Optimization*, 45(1), pp. 3-38.
- Friedler, F., Tarjan, K., Huang, Y. W. & Fan, L. T., 1992. Graph-theoretic approach to process synthesis - axioms and theorems. *Chem Eng Sci*, 47(8), pp. 1973-1988.
- Friedler, F., Tarjan, K., Huang, Y. W. & Fan, L. T., 1993. Graph-theoretic approach to process synthesis - polynomial algorithm for maximal structure generation. *Comput Chem Eng*, 17(9), pp. 929-942.
- Gavrilescu, M. & Chisti, Y., 2005. Biotechnology- a sustainable alternative for chemical industry. *Biotechnology advances*, Volume 23, pp. 471-499.

- Gonzalez-Delgado, A., Kafarov, V. & El-Halwagi, M., 2015. Development of a topology of microalgae-based biorefinery: process synthesis and optimization using a combined forward–backward screening and superstructure approach. *Clean Technologies and En. Clean Tech Env Pol*, 17(8), pp. 2213-2228.
- Grossmann, I. E. & Daichendt, M. M., 1996. New trends in optimization-based approaches to process synthesis. *Computers & chemical engineering*, 20(6), pp. 665-683.
- Grossmann, I. E. & Guillen-Gosalbez, G., 2010. Scope for the application of mathematical programming techniques in the synthesis and planning of sustainable processes. *Computers & Chemical Engineering. Computers & Chemical Engineering*, 34(9), pp. 1365-1376.
- Hailemariam, L. & Venkatasubramanian, V., 2010. Purdue ontology for pharmaceutical engineering: Part II. Applications. *J Pharm Innov*, 5(4), pp. 139-146.
- Harrison, R. G., Todd, P. W., Rudge, S. R. & Petrides, D., 2003. *Bioseparations Science and Engineering*. Oxford, UK: Oxford University Press.
- Henao, C. & Maravelias, C., 2011. Process superstructure optimization using surrogate models. *AIChE J*, 57(5), pp. 1216-1232.
- Horowitz, D. & Brennan, E., 2001. *Methods for separation and purification of biopolymers*. US, Patent No. 20010006802.
- Horst, R. & Pardalos, P. M., 2013. *Handbook of global optimization*. s.l.:Springer Science & Business Media.
- Huang, H. J., Ramaswamy, S., Tschirner, U. W. & Ramarao, B. V., 2008. A review of separation technologies in current and future biorefineries. *Separation and purification technology*, Volume 62, pp. 1-21.
- Inskeep, G., Taylor, G. & Breitzke, W., 1952. Lactic acid from corn sugar. *Ind Eng Chem*, 44(9), pp. 1955-1966.
- Jacquel, N. et al., 2008. Isolation and purification of bacterial poly (3-hydroxyalkanoates). *Biochem Eng J*, Volume 39, pp. 15-27.
- Julleson, D., David, F., Pfleger, B. & Nielsen, J., 2015. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnol Adv*, 33(7), pp. 1395-1402.
- Kelloway, A. & Daoutidis, P., 2013. Process synthesis of biorefineries: optimization of biomass conversion to fuels and chemicals. *Industrial & Engineering Chemistry Research*, Volume 53, pp. 5261-5273.
- Kim, J., Sen, M. & Maravelias, C., 2013. An optimization-based assessment framework for biomass-to-fuel conversion strategies. *Energy & Environmental Science*, 6(4), pp. 1093-1104.
- King, C., Gantz, D. & Barnes, F., 1972. Systematic evolutionary process synthesis. *Ind Eng Chem Proc Des Dev*, Volume 11, pp. 271-283.
- Kiss, A. A., Grievink, J. & Rito-Palomares, M., 2015. A systems engineering perspective on process integration in industrial biotechnology. *J Chem Technol Biotechnol*, Volume 90, pp. 349-355.

- Knoop, H. & Steuer, R., 2015. A computational analysis of stoichiometric constraints and trade-offs in cyanobacterial biofuel production. *Front Bioeng Biotechnol*, 3(47).
- Kocis, G. R. & Grossmann, I., 1989. A Modeling and decomposition strategy for the MINLP optimization of process flowsheets. *Comput Chem Eng*, 13(7), pp. 797-819.
- Kokossis, A. C., Yang, A., Tsakalova, M. & Lin, T. C., 2010. A systems platform for the optimal synthesis of biomass based manufacturing systems. *Comp Aid Chem Eng*, Volume 28, pp. 1105-1110.
- Kokossis, A. & Yang, A., 2010. On the use of systems technologies and a systematic approach for the synthesis and the design of future biorefineries. *Computer Chemical Engineering*, Volume 34, pp. 1397-1405.
- Koller, M., Niebelschutz, H. & Braunegg, G., 2013. Strategies for recovery and purification of poly-[(R)-3-hydroxyalkanoates] (PHA) biopolyesters from surrounding biomass. *Eng Life Sci*, Volume 13, pp. 549-562.
- Kondili, E., Pantelides, C. C. & Sargent, R. W., 1993. A general algorithm for short-term scheduling of batch operations-I. MILP formulation. *Comp & Chem Eng*, Volume 17, pp. 211-227.
- Kong, L. et al., 2016. A superstructure-based framework for simultaneous process synthesis, heat integration, and utility plant design. *Computers & Chemical Engineering*, Volume 91, pp. 68-84.
- Kravanja, Z., 2010. Challenges in sustainable integrated process synthesis and the capabilities of an MINLP process synthesizer MipSyn. *Computers & Chemical Engineering*, 34(11), pp. 1831-1848.
- Kravanja, Z. & Grossmann, I. E., 1990. PROSYN—an MINLP process synthesizer. *Computers & chemical engineering*, 14(12), pp. 1363-1378.
- Kravanja, Z. & Grossmann, I. E., 1993. Prosyn—an automated topology and parameter process synthesizer. *Computers & chemical engineering*, Volume 17, pp. 87-94.
- Lakshmanan, A. & Biegler, L., 1996. Synthesis of optimal chemical reactor networks. *Ind Eng Chem Res.*, 35(4), pp. 1344-1353.
- Liao, J. C., Mi, L., Pontrelli, S. & Luo, S., 2016. Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, 14(5), pp. 288-304.
- Lightfoot, E. N. & Moscariello, J. S., 2004. Bioseparations. *Biotechnology and bioengineering*, 87(3), pp. 259-273.
- Ling, Y., Williams, D., Thomas, C. & Middelberg, A., 1997. Recovery of poly-3-hydroxybutyrate from recombinant Escherichia coli by homogenization and centrifugation. *Biotech Techn*, 11(6), pp. 409-412.
- Liu, P., Georgiadis, M. C. & Pistikopoulos, E. N., 2010. Advances in energy systems engineering. *Ind Eng Chem Res*, 50(9), pp. 4915-4926.
- Li, X. & Kraslawski, A., 2004. Conceptual process synthesis: past and current trends. *Chemical Engineering and Processing: Process Intensification*, 43(5), pp. 583-594.

- Marquardt, W., Kossack, S. & Kraemer, K., 2008. A framework for the systematic design of hybrid separation processes. *Chinese J Chem Eng*, 16(3), pp. 333-342.
- Mascia, S. et al., 2013. End-to-End Continuous Manufacturing of pharmaceuticals: integrated synthesis, purification ,and final dosage formation. *Angewandte Chemie International Edition*, Volume 52, pp. 12359-12363.
- McCabe, W., Smith, J. & Harriott, P., 2005. *Unit operations of chemical engineering*. New York: McGraw-Hill Education.
- McGregor, W., 1986. *Membrane separations in biotechnology*. New York: Marcel Dekker Inc.
- Nallasivam, U. et al., 2013. Global optimization of multicomponent distillation configurations: 1. Need for a reliable global optimization algorithm. *AIChE J*, 59(3), pp. 971-981.
- Nishida, N., Stephanopoulos, G. & Westerberg, A., 1981. A review of process synthesis. *AIChE J*, Volume 27, pp. 321-351.
- Noble, R. D. & Agrawal, R., 2005. Separations research needs for the 21st century. *Ind & Eng Chem Res*, 44(9), pp. 2887-2892.
- Novak, Z., Kravanja, Z. & Grossmann, I. E., 1993. Simultaneous optimization model for multicomponent separation. *Computers and Chemical Engineering*, Volume 18, pp. S125-S129.
- Peters, M., Timmerhaus, K. & West, R., 2005. *Plant design and economics for chemical engineers*. 5th ed. New York: McGraw Hill.
- Pollak, P., 2011. *Fine chemicals: the industry and the business*. 2nd ed. New Jersey: John Wiley & Sons.
- Rizwan, M., Lee, J. H. & Gani, R., 2013. Optimal processing pathway for the production of biodiesel from microalgal biomass: a superstructure based approach. *Comp & Chem Eng*, Volume 58, pp. 305-314.
- Rogers, A. & Ierapetritou, M., 2015. Feasibility and flexibility analysis of black-box processes Part 2: surrogate-based flexibility analysis. *Chem Eng Sci*, Volume 137, pp. 1005-1013.
- Schaber, S. D. et al., 2011. Economic analysis of integrated continous and batch pharmaceutical manufacturing: a case study. *Industiral & Engineering Chemistry Research*, Volume 50, pp. 10083-10092.
- Schweiger, C. & Floudas, C., 1999. Optimization framework for the synthesis of chemical reactor networks. *Ind Eng Chem Res*, 38(3), pp. 744-766.
- Seader, J., Henley, E. & Roper, D., 2011. *Separation process principles: chemical and biochemical operations*. 3rd ed. New York: John Wiley & Sons.
- Sirola, J., 1996. Strategic process synthesis: advances in the hierarchical approach. *Comp Chem Eng*, Volume 20, pp. 1637-1643.
- Smith R, 2005. *Chemical Process design and integration*. New York: John Wiley & Sons.
- Smith, E., 1996. *On the optimal design of continuous processes*, London: Diss. Imperial College London.
- Smith, J. C. & Harriott, P., 1993. *Unit Operations of Chemical Engineering*. 5th ed. s.l.:McGraw-Hill.

- Smith, R., 2015. Chemical Process Design. In: Kirk-Othmer, ed. *Kirk-Othmer Encyclopedia of Chemical Technology*. New Jersey: John Wiley & Sons, Inc, pp. 1-29.
- Smith, R. & Linnhoff, B., 1988. The design of separators in the context of overall processes. *Chem Eng Res Des*, Volume 66, pp. 195-228.
- Stephanopoulos, G. & Westerberg, A., 1975. Synthesis of optimal process flowsheets by an infeasible decomposition technique in the presence of functional non-convexities. *Canadian J Chem Eng*, 53(5), pp. 551-555.
- SuperPro, v9.0. NJ, USA: Intelligen.
- Suresh, P. et al., 2010. OntoMODEL: ontological mathematical modeling knowledge management in pharmaceutical product development, 1: conceptual framework. *Ind & Eng Chem Res*, 49(17), pp. 7758-7767.
- Tawarmalani, M. & Sahinidis, N., 2005. A polyhedral branch-and-cut approach to global optimization. *Math Program*, 103(2), pp. 225-249.
- Tawarmalani, M. & Sahinidis, N. V., 2004. Global optimization of mixed-integer nonlinear programs: a theoretical and computational study. *Math Program*, Volume 99, pp. 563-591.
- Tee, T. W., Chowdhury, A., Maranas, C. D. & Shanks, J. V., 2014. Systems metabolic engineering design: Fatty acid production as an emerging case study. *Biotechnol Bioeng*, Volume 111, pp. 849-857.
- Trespalacios, F. & Grossmann, I., 2014. Review of Mixed-Integer Nonlinear and Generalized Disjunctive Programming Methods. *Chemie Ingenieur Technik*, 86(7), pp. 991-1012.
- Tula, A. K., Eden, M. R. & Gani, R., 2015. Process synthesis, design and analysis using a process-group contribution method. *Computers & Chemical Engineering*, Volume 81, pp. 245-259.
- Vasudevan, P. & Ulrich, G., 2004. *Chemical engineering process design and economics: a practical guide*. 2nd ed. New Hampshire: Process Publishing.
- Vausa, J. & Marquardt, W., 2000. Shortcut design methods for hybrid membrane/distillation processes for the separation of nonideal multicomponent mixtures. *Ind Eng Chem Res*, 39(6), pp. 1658-1672.
- Venkatasubramanian, V. et al., 2006. Ontological informatics infrastructure for pharmaceutical product development and manufacturing. *Computers and Chemical Engineering*, Volume 30, pp. 1482-1496.
- Viiavakumar, J., Aravindan, R. & Viruthagiri, T., 2008. Recent trends in the production, purification and application of lactic acid. *Chem Biochem Eng Quart*, Volume 22, pp. 45-264.
- Wasewar, K., 2005. Separation of lactic acid: recent advances. *Chem Biochem Eng Q*, 19(2), pp. 159-172.
- Westerberg, A. W., 2004. A retrospective on design and process synthesis. *Computers & Chemical Engineering*, 28(4), pp. 447-458.
- Wilson, S. A. & Roberts, S. C., 2014. Metabolic engineering approaches for production of biochemicals in food and medicinal plants. *Current Opinion in Biotechnology*, Volume 26, pp. 174-182.

- Wu, W., Henao, C. A. & Maravelias, C. T., 2016. A Superstructure Representation, Generation and Modeling Framework for Chemical Process Synthesis. *AIChE J*, p. 10.1002/aic.15300.
- Yee, T. F. & Grossmann, I., 1990. Simultaneous-optimization models for heat integration. 2. heat-exchanger network synthesis. *Comput Chem Eng.*, 14(10), pp. 1165-1184.
- Yeomans, H. & Grossmann, I., 1999. A systematic modeling framework of superstructure optimization in process synthesis. *Comput Chem Eng*, 23(6), pp. 709-731.
- Yuan, Z., Chen, B. & Gani, R., 2013. Applications of process synthesis: moving from conventional chemical processes towards biorefinery processes. *Computer Chemical Engineering*, Volume 49, pp. 217-229.
- Yu, M., Miller, D. & Biegler, L., 2015. Dynamic reduced order models for simulating bubbling fluidized bed adsorbers. *Ind Eng Chem Res*, 54(27), pp. 6959-6974.
- Zhang, X., Tervo, C. J. & Reed, J. L., 2016. Metabolic Assessment of E. coli as a Biofactory for Commercial Products. *Metabolic Engineering*, Volume 35, pp. 64-74.
- Zhang, Y. & Sahinidis, N., 2012. Uncertainty quantification in CO₂ sequestration using surrogate models from polynomial chaos expansion. *Industrial & Engineering Chemistry Research*, 52(9), pp. 3121-3132.