Process simulator-based optimization of biorefinery downstream processes under the Generalized Disjunctive Programming framework

Michele Corbetta[+], Ignacio E. Grossmann[‡], Flavio Manenti[+*]

[+] Politecnico di Milano, Dipartimento di Chimica, Materiali e Ingegneria Chimica *"Giulio Natta",* Piazza Leonardo da Vinci 32, 20133 Milano, Italy.

[‡] Carnegie Mellon University, Department of Chemical Engineering, 5000 Forbes Avenue, 15213 Pittsburgh, PA, USA.

[*]To whom correspondence should be addressed

Phone: +39 02 2399 3273

Fax: +39 02 2399 3280

E-mail: flavio.manenti@polimi.it

Process simulator-based optimization of biorefinery downstream processes under the Generalized Disjunctive Programming framework

Michele Corbetta, Ignacio E. Grossmann, Flavio Manenti

Abstract

Downstream processing of biofuels and bio-based chemicals represents a challenging problem for process synthesis and optimization, due to the intrinsic nonideal thermodynamics of the liquid mixtures derived from the (bio)chemical conversion of biomass. In this work, we propose a new interface between the process simulator SimSci PRO/II and the optimization environment of GAMS for the structural and parameter optimization of this type of flowsheets with rigorous and detailed models. The optimization problem is formulated within the Generalized Disjunctive Programming framework and the solution of the reformulated MINLP problem is approached with a decomposition strategy based on the Outer Approximation algorithm, where NLP subproblems are solved with the BzzMath derivative free optimizer, and MILP master problems are solved with CPLEX/GAMS. Several validation examples are proposed spanning from the economic optimization of two different distillation columns, the dewatering task of diluted bio-mixtures, up to the distillation sequencing with simultaneous mixed-integer design of each distillation column for a quaternary mixture in the presence of azeotropes.

Keywords: Optimization, Generalized Disjunctive Programming, MINLP, Downstream, Dewatering, Distillation, Process Simulator, Oxygenated Chemicals, Azeotropes, UNIQUAC, PRO/II, GAMS, Derivative Free Optimization, BzzMath.

1. Introduction

The renewed interest in the field of distillation has been recently promoted by the consistent research on biomass conversion technologies to biofuels and bio-based chemicals. These technologies are based on (bio)chemical reactors that produce highly diluted aqueous solutions. The downstream processing of those mixtures usually involves distillation, leading to high operating costs motivated by the high heat of vaporization of water (G. Q. Chen, 2009; Xiu & Zeng, 2008). For this reason, attempts to optimize and thermally integrate the purification step (Ahmetovic et al., 2010; Dias et al., 2009; Karuppiah et al., 2008) result in a relevant lowering of the production costs that reduces the economic gap with respect to cheaper fossil-based products (Hermann & Patel, 2007; Sauer et al., 2008). In addition, the optimization of this type of downstream processes, as opposed to hydrocarbon distillation, involves highly nonideal liquid mixtures that demand rigorous thermodynamic models.

In this context, process simulators offer a reliable and rigorous modeling environment that rely on extensive thermodynamic properties databanks and tailored distillation algorithms, in contrast with equation-oriented optimization tools that are usually based on shortcut models for the unit operations and for the estimation of physical and thermodynamic properties (Navarro-Amoros et al., 2013). Unfortunately, it has been demonstrated that the optimization tools available within commercial

simulation packages are not as effective and flexible as it would be required (Biegler, 1985) due to the high nonlinearity of the equation systems, and to the impossibility to optimize structural (integer) decision variables. This was the motivation for several authors to develop ad-hoc interfaces for the process simulator-based optimization with MINLP optimization algorithms. Two main strategies have been proposed; the one based on the augmented penalty/equality relaxation outer-approximation (AP/ER/OA) deterministic algorithm (Viswanathan & Grossmann, 1990), and the ones based on metaheuristic methods, such as the evolutionary algorithms (Gross & Roosen, 1998).

Starting from the deterministic approach, (Harsh et al., 1989) developed an MINLP algorithm for the retrofit of chemical plants with fixed topology based on the FLOWTRAN process simulator, and they applied it to the ammonia synthesis process. (Diaz & Bandoni, 1996) derived an MINLP approach to optimize the structure and the parameters of a real ethylene plant in operation, interfacing a specific simulation code. (Caballero et al., 2005) proposed an optimization algorithm for the rigorous design of single distillation columns using Aspen HYSYS. (Brunet et al., 2012) applied the same methodology to assist decision makers in the design of environmentally conscious ammonia–water absorption machines for cooling and refrigeration. (Navarro-Amoros et al., 2014) proposed a new algorithm for the structural optimization of process superstructures within the Generalized Disjunctive Programming (GDP) framework. Finally, (Garcia et al., 2014) proposed a hybrid simulation-multiobjective optimization approach that optimizes the production cost and minimizes the associated environmental impacts of isobutane alkylation. The simultaneous process optimization and heat integration approach has been also addressed by coupling process simulators with external equation systems (Y. Chen et al., 2015; Navarro-Amoros et al., 2013).

On the other hand, several authors have proposed optimization algorithms based on evolutionary methods in order to overcome some difficulties that arise from the use of deterministic nonlinear programming solvers with real-world complex problems. For instance, (Gross & Roosen, 1998) addressed the simultaneous structural and parameter optimization in process synthesis coupling Aspen Plus with evolutionary methods. Similarly, an optimization framework is proposed by (Leboreiro & Acevedo, 2004) for the synthesis and design of complex distillation sequences, based on a modified genetic algorithm coupled with a sequential process simulator, succeeding in problems where deterministic mathematical algorithms had failed. (Vazquez-Castillo et al., 2009) address the optimization of intensified distillation systems for quaternary distillations with a multiobjective genetic algorithm coupled to the Aspen Plus process simulator. Subsequently, (Gutierrez-Antonio & Briones-Ramirez, 2009) implemented a multiobjective genetic algorithm coupled with Aspen Plus to obtain the Pareto front of Petlyuk sequences. (Bravo-Bravo et al., 2010) proposed a novel extractive dividing wall distillation column, which has been designed using a constrained stochastic multiobjective optimization technique, based on the use of GA algorithms. Finally, (Eslick & Miller, 2011) developed a modular framework for multi-objective analysis aimed at minimizing freshwater consumption and levelized cost of electricity for the retrofit of a hypothetical 550 MW subcritical pulverized coal power plant with an MEA-based carbon capture and compression system.

In this work, a new interface between the process simulator SimSci PRO/II and GAMS is proposed for the structural and parameter optimization of downstream processes based on the OA algorithm and on the use of a Derivative Free Optimizer (DFO). The optimization tool is applied to several case studies, including the distillation sequencing with simultaneous mixed-integer optimal design of each distillation column for a quaternary mixture in presence of azeotropes. The paper is structured as follows. Section 2 defines the problem statement, including the main modeling assumptions and the required inputs for the superstructure optimization. Then, the process superstructure (Section 3) and the modeling framework (Section 4) are described. Section 5 addresses the optimization algorithm,

while Section 6 reports the solution strategy and discusses implementation issues. Finally, Section 7 provides a selection of application examples in the field of bio-based chemicals downstream processing.

2. Problem statement

The aim of this work is to propose a new algorithm for the topological optimization of complex process superstructures based on rigorous thermodynamic models, with special emphasis on distillation downstream processes in the biorefining area. Specifically, the distillation sequencing problem with simultaneous design of number of trays and feed tray location is addressed. Both continuous (e.g. split ratio, reflux ratio, pressure) and integer (e.g. number of trays, feed trays, equipment existence) decision variables are optimized under the Generalized Disjunctive Programming (GDP) framework (Grossmann & Trespalacios, 2013), using the process modeling environment of SimSci PRO/II, and the optimization environment of GAMS.

The process modeling is based on the assumptions of the process simulator. Particularly, the most representative equipment is the distillation column, which is described as a cascade of countercurrent vapour-liquid phase equilibrium stages, with a constant pressure drop per stage, a constant High Equivalent to a Theoretical Plate (HETP) for structured packing internals, and a kettle-type reboiler.

The optimization algorithm basically requires a superstructure, the propositional logic to define the topology of the superstructure, selected flowsheets implemented in PRO/II, a set of bounded continuous and integer decision variables, nonlinear (eventually implicit) constraints (e.g. purity and safety constraints), and an economic objective function.

3. Process superstructure

The optimization procedure starts from the definition of the process superstructure. The most general superstructure that can be handled by this algorithm is based on the interconnection of permanent units with elementary conditional unit&trays modules. While permanent units are present in each possible optimal flowsheet originated from the superstructure, the elementary conditional unit&trays modules are introduced in the superstructure to describe the conditional units (or conditional sections with more than one unit) that are not necessarily present in the final optimized flowsheet. Conditional trays are introduced within the conditional unit&trays module for the rectification and stripping sections of the distillation columns eventually present (Figure 1). The GDP conditional tray representation is adopted to define feed stage and number of stages of the distillation column (Barttfield et al., 2004).

It is worth to note that the superstructure is never completely implemented as a unique process simulator flowsheet. Rather the model could be depicted as a collection of different possible black-box simulations, which are defined by disjunctions, in contrast with fully equation-oriented models. For this reason, only permanent units and selected conditional units are solved at each call of the process simulator. In this way, no splitters are required for conditional units, and zero-flow units are avoided.

A relevant case of this kind of superstructure arise from the solution of distillation sequencing problems. At this level, it is possible to represent the sequencing with either a State Task Network (STN) or a State Equipment Network (SEN) (Yeomans & Grossmann, 1999). Figure 2 reports the

two different superstructures for the distillation of a ternary mixture. It is possible to highlight that the SEN requires a smaller number of columns but it introduces recycles. Nevertheless, since the modeling is accomplished by the process simulator with a logic-based definition of the input file that considers only selected units, it does not matter if either STN or SEN superstructure is adopted.

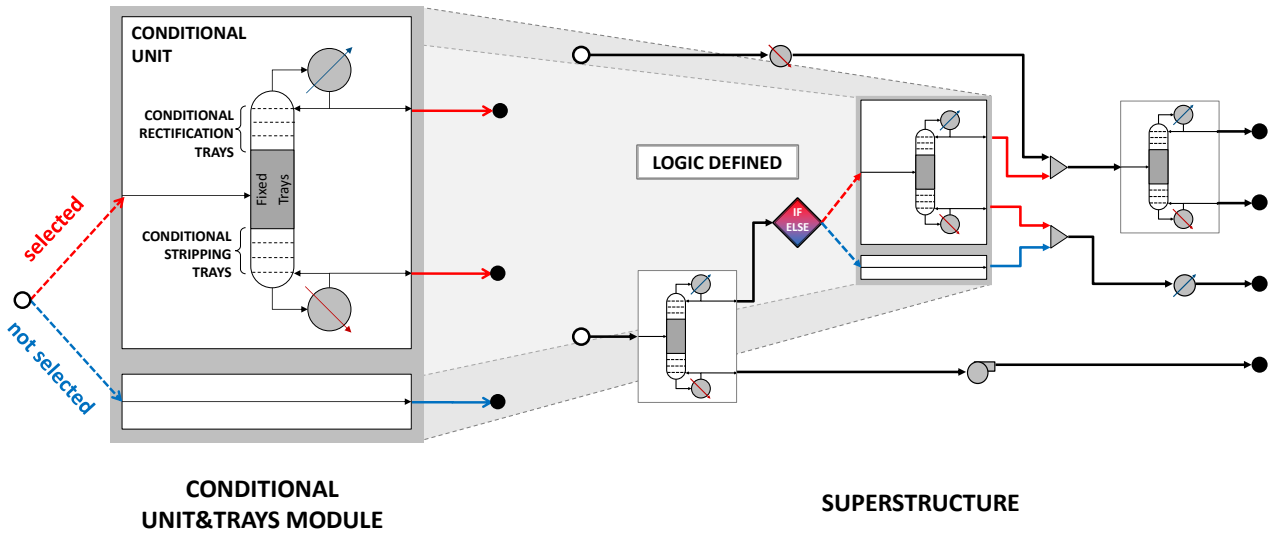
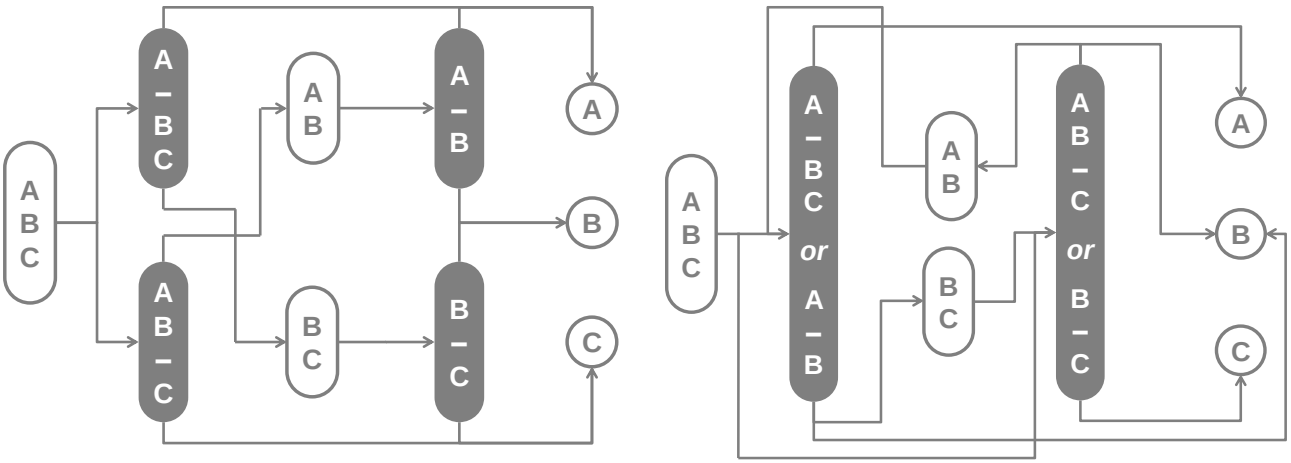


Figure 1: Representations of the conditional unit&trays elementary module and of a typical superstructure.



(a) (b)

Figure 2: (a) State Task Network and (b) State Equipment Network representations for the sharp distillation of a ternary mixture.

4. Modeling

The detailed modeling is achieved with a process simulator (SimSci PRO/II) taking advantage of thermodynamic databanks for the estimation of physical properties and ad-hoc algorithms for the solution of nonlinear systems derived from distillation columns and other unit operations. Within this process modeling environment there is also the flexibility to introduce custom modeling components, which can be required in case of nonconventional unit operations, as discussed elsewhere (Corbetta et al., 2014).

4.1 Thermodynamic modeling

When the target is to address the synthesis of biorefinery downstream processes, it is worth mentioning that predictive thermodynamic models, such as UNIFAC, frequently fail. This is due to the highly complex nature of the interactions between oxygenated chemicals in the aqueous phase. These complex liquid mixtures can be obtained, for instance, in the form of fermentation broth withdrawn from a bioreactor or from the outlet of catalytic deoxygenation converters, and they present some common characteristics. Usually they are diluted organic aqueous solutions, in which water can represent up to 80-90 wt.%. Moreover, they are made up of a large amount of components that belong to the same chemical class (e.g. ketones, alcohols, polyols), with a frequent occurrence of homogeneous azeotropes and pinch points in the corresponding VLE equilibrium diagrams. Finally, a heavy cut composed by soluble solids is usually present due to incomplete biomass conversion, presence of an inert lignin fraction, and/or production of high molecular weight components by side reactions. Consequently, developing reliable VL(L)E thermodynamic models based on nonlinear regression of (UNIQUAC or NRTL) binary interaction parameters on experimental data (Pirola et al., 2014) plays a major role in correctly predicting the distillation behavior, as will be highlighted in Section 7.

4.2 Cost functions

Once the process simulator is set up with a proper thermodynamic model, the convergence of a flowsheet provides the value of the implicit variables for the evaluation of the economic objective function and for measuring the violation of nonlinear constraints. Specifically, the techno-economic assessment is approached with nonlinear cost functions (Douglas, 1988) and rigorous sizing models embedded within the process simulator (e.g. tray and packing hydraulics). For distillation columns, the minimization of the cost objective function is performed by considering both annualized capital costs (CAPEX) and operating cost (OPEX), which added together determine the Total Annualized Costs (TAC) as reported in Eq.(1).

$$TAC = \frac{C_{inv}}{payback\ time} + C_{op} = \left(\frac{C_{col} + C_{internals} + C_{reb} + C_{cond}}{payback\ time} \right) + C_{steam} + C_{cw} + C_{entrainer} \quad (1)$$

Operating costs include utility costs (steam, cooling water and eventually entrainers), while column investment costs include trays or packing, column vessel, condenser and reboiler installation and purchase costs, which depend on the value of the decision optimization variables and on the size of the equipment. For CAPEX evaluation, the following function, Eq.(2), is adopted, where constants c_1, c_2, e_1 and e_2 depend on the equipment type, L_1 and L_2 are relevant size, Fc is a parameter depending on the fabrication material and operating pressure, while *M&S* is the Marshall & Swift economic index.

$$C_{equipment}[\$] = c_1 \left(\frac{M\ \&\ S}{280} \right) L_1^{e_1} L_2^{e_2} (c_2 + \mathrm{F}_{C,equiment}) \quad (2)$$

4.3 Generalized Disjunctive Programming formulation

The optimization problem is formulated within the Generalized Disjunctive Programming framework (Grossmann & Trespalacios, 2013), in which also implicit variables ($\mathbf{x_I}$) are assumed, along with continuous decision variables ($\mathbf{x}$) and Boolean decision variables ($\mathbf{Y}$), as outlined in Eq.(3). These implicit variables are evaluated by the process simulator, which is treated as a black-box and is represented by the implicit function (f_I) that links implicit variables with decision variables. Nonlinear constraints (g) can be both global (set G) and conditional (set G_k). The mathematical programming formulation involves the definition of a set K of disjunctions, corresponding to conditional unit&trays

modules, for which additional constraints, equations, and cost contributions are defined. Each disjunction involves only two terms, corresponding to the selection or not of a conditional unit (Y_k), with embedded disjunctions for the conditional trays (Y_j) belonging to each elementary module (set of CT_k conditional rectification and stripping trays of column k). The economic objective function includes these contributions along with a general function ($f(\mathbf{x}, \mathbf{x_I})$) of the decision variables. Finally, the propositional logic defines the units' interconnectivity. Topology logic ($\Omega_{topology}(Y_k) = true$) defines the interconnection between conditional units, while the three subsequent sets of logic constraints are required to ensure that no conditional trays are selected if the conditional unit is not selected, at most only one Boolean variable should be true for the rectification (RCT_k) and for the stripping (SCT_k) conditional trays of each distillation column (either permanent or conditional).

$$
\begin{aligned}
&\min Z = \sum_{k \in K} TAC_k + f(\mathbf{x}, \mathbf{x_I}) \\
&\text{s.t. } \mathbf{x_I} = [\mathbf{x_I}' \ \mathbf{x_I}''] \\
&\quad \mathbf{x_I}' = f_I'(\mathbf{x}) \\
&\quad g_n(\mathbf{x}, \mathbf{x_I}) \le 0 \qquad n \in G \\
&\quad \begin{bmatrix} Y_k \\ \begin{bmatrix} Y_j \\ TAC_j = \gamma_j \end{bmatrix} \vee \begin{bmatrix} \neg Y_j \\ TAC_j = 0 \end{bmatrix} \\ x_{I,k}'' = f_I''(\mathbf{x}) \\ g_n(\mathbf{x}, \mathbf{x_I}) \le 0 \\ TAC_k = \sum_{j \in CT_k} \gamma_j(\mathrm{x}, \mathrm{y}, \mathrm{x_I}) \end{bmatrix} \vee \begin{bmatrix} \neg Y_k \\ x_{I,k}'' = 0 \\ TAC_k = 0 \end{bmatrix} \qquad j \in CT_k,\ n \in G_k,\ k \in K \\
&\quad \Omega_{topology}(Y_k) = true \qquad (3) \\
&\quad \neg Y_k \Rightarrow \neg Y_j \qquad k \in K,\ j \in CT_k \\
&\quad \dot{\vee} Y_i \qquad i \in RCT_k,\ k \in K \\
&\quad \dot{\vee} Y_i \qquad i \in SCT_k,\ k \in K \\
&\quad \mathbf{x^L} \le \mathbf{x} \le \mathbf{x^U} \\
&\quad \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x_I} \in \mathbb{R}^{n_I},\ \mathbf{c} \in \mathbb{R}^m,\ \mathbf{Y} \in \{true; false\}^m
\end{aligned}
$$

5. Optimization algorithm

The optimization problem is solved with a decomposition strategy based on the Logic-Based Outer Approximation (LBOA) algorithm (Turkay & Grossmann, 1996). The algorithm involves NLP subproblems that arise from fixed values at the Boolean variables Y in Eq.(3), and MILP master problems that correspond to a linear approximation at the GDP in Eq.(3). NLP subproblems are solved with a Derivative Free Optimizer (C++/BzzMath), and MILP master problems are solved with Branch & Cut methods (GAMS/CPLEX). The corresponding block diagram of this algorithm is shown in Figure 3.

The LBOA algorithm requires a first guess for both the continuous and integer optimization variables. An initial NLP subproblem gives the initial values of the continuous decision variables ($\mathbf{x_0}$). In case of conditional units, i.e. when the feasible flowsheet derived from the superstructure does not include necessarily all the units, a set covering problem is considered from which several NLP subproblems are solved to provide linearization points for all the process units (Turkay & Grossmann, 1996). Subsequently, a main loop iterates between the solution of the MILP master problem and the NLP subproblem until the exit condition is satisfied. The MILP master problem updates the value of the integer decision variables ($\mathbf{y_{MILP}}$) and provides a first guess ($\mathbf{x_{MILP}}$) to the NLP subproblem, which in turn, updates the value of the continuous optimization variables ($\mathbf{x_{NLP}}$). The loop is terminated as soon as both the NLP worsening condition and the crossing of lower and upper bounds occur.

The C++ programming environment provides the communication platform between PRO/II and GAMS, by writing text input files and retrieving results from text output files of the simulator and of the optimizer, respectively. Auxiliary C++ functions allows to translate the GDP with fixed Boolean variables Y_k to a specific topology that corresponds to an MINLP with feed tray and number of trays as variables to be determined, and to evaluate the economic objective function and the violation of nonlinear constraints.

It is important to emphasize that NLP subproblems involve only selected units, thus avoiding convergence difficulties due to zero flows.

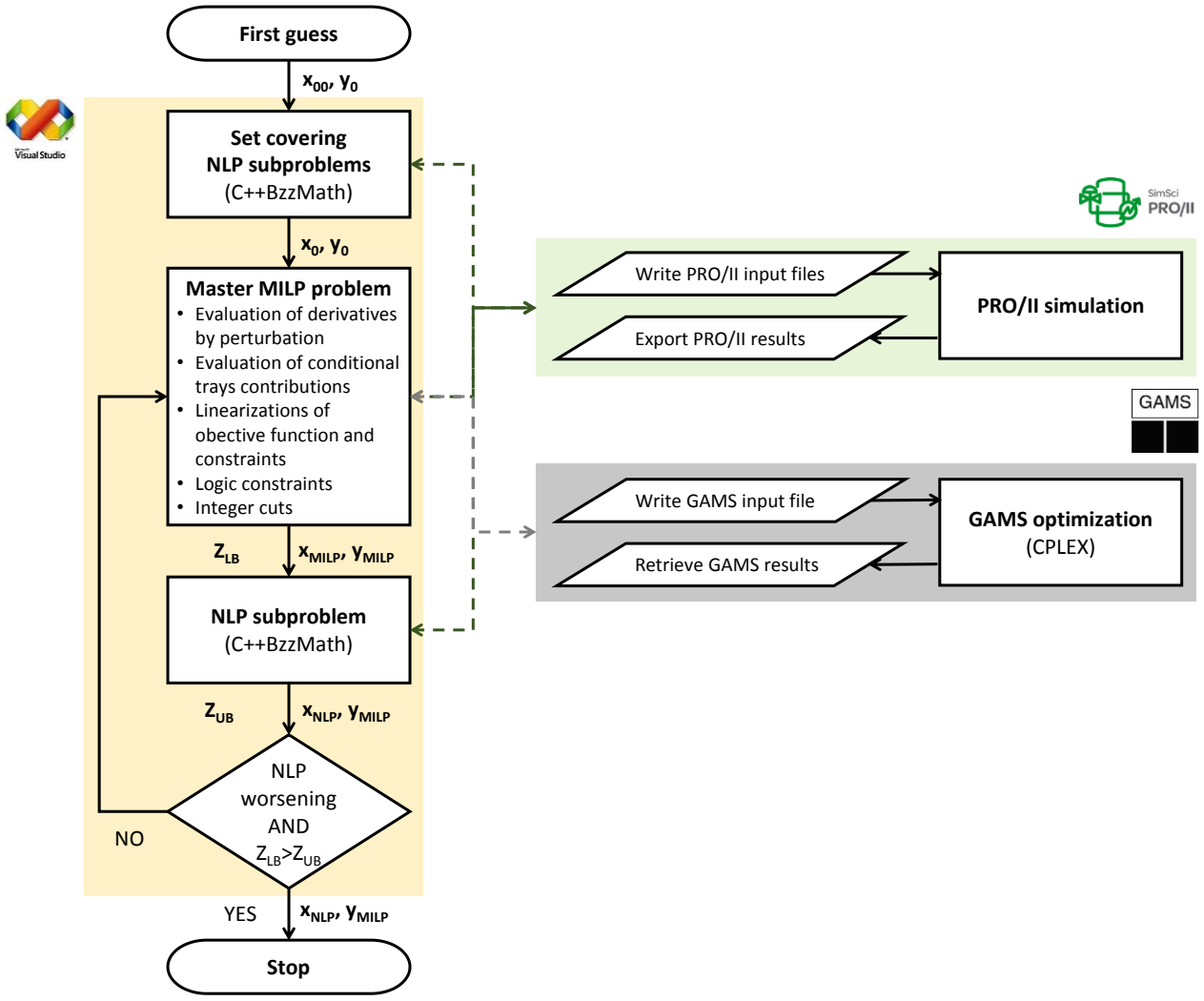


Figure 3: Optimization algorithm block diagram.

5.1 NLP subproblem

NLP subproblems are solved with the Derivative Free Optimizer belonging to the BzzMath numerical library (Buzzi-Ferraris & Manenti, 2012), which is available online at *http://super.chem.polimi.it/*. The `BzzMinimizationRobust` is based on a modified Nelder-Mead Simplex direct search method (OPTNOV variant), which has proved to handle problems with highly nonlinear, nondifferentiable and discontinuos functions, and problems with narrow valleys (Buzzi-Ferraris, 1967; G. Buzzi-Ferraris & F. Manenti, 2010).

Nonlinear constraints of the NLP subproblems (in the form $g\left(\mathbf{x},\mathbf{x_I}\right)\le 0$) are handled with a penalty function proportional to the constraint violation that is added to the objective function, allowing to start with an infeasible first guess. A Sequential Unconstrained Minimization Technique (Correia et al., 2010) is adopted by progressively increasing the penalty weight (w_{g_m}) in order to limit difficulties with narrow valleys.

The n-th subproblem, corresponding to the $\mathbf{y_n}$ integer variables, is reported in Eq.(4), where the first set of constraints derives from the convex-hull reformulation of the disjunctions (Grossmann & Trespalacios, 2013). These constraints force the decision variables of the non-selected units ($y_k = 0$) to be zero.

$$\begin{aligned}
&\min_{\mathbf{x}} Z_n^{UB}\left(\mathbf{y}_n\right) = \sum_{k\in K} TAC_k\left(\mathbf{y}_n\right) + f\left(\mathbf{x},\mathbf{x_I}\right) + \sum_{i\in G\cup G_k} w_{g_i}\max\left(0; g_i(\mathbf{x},\mathbf{x_I})\right)\\
&\text{s.t.}\quad y_k\cdot {x_k}^L \le x_k \le y_k\cdot {x_k}^U \qquad k\in K\\
&\qquad \mathbf{x}\in\mathbb{R}^n
\end{aligned} \tag{4}$$

5.2 MILP master problem

The MILP master problem, reported in Eq.(5), is a linearization of the original nonlinear problem, which includes accumulated linearizations of the objective function and of the constraints, logic constraints representing the interconnectivity among process units within the superstructure, integer cuts, and convex-hull constraints derived from the reformulation of the disjunctions.

$$\begin{aligned}
&\min_{\mathbf{x},\mathbf{y}} Z_{OA}^{LB} = \alpha + w\cdot\left[\sum_{l\in L}\left(slack_f^l + slack_g^l\right)\right]\\
&s.t.\ f_{obj}(\mathbf{x}^l,\mathbf{y}^l,\mathbf{x}_\mathbf{I}^l) + \sum_{i\in iX}\frac{\partial f_{obj}}{\partial x_i}(\mathbf{x}^l,\mathbf{y}^l,\mathbf{x}_\mathbf{I}^l)(\mathrm{x}_i-\mathrm{x}_i^l) + \sum_{t\in CT}\gamma_i y_i \le slack_f^l + \alpha \qquad l\in L\\
&\quad g_n(\mathbf{x}^l,\mathbf{y}^l,\mathbf{x}_\mathbf{I}^l) + \sum_{i\in iX}\frac{\partial g_n}{\partial x_i}(\mathbf{x}^l,\mathbf{y}^l,\mathbf{x}_\mathbf{I}^l)(\mathrm{x}_i-\mathrm{x}_i^l) + \sum_{t\in CT}\eta_t y_t \le slack_{g_i}^l \qquad n\in G,\ l\in L\\
&\quad \mathrm{y}_k\cdot g_n(\mathbf{x}^l,\mathbf{y}^l,\mathbf{x}_\mathbf{I}^l) + \sum_{i\in iX}\frac{\partial g_n}{\partial x_i}(\mathbf{x}^l,\mathbf{y}^l,\mathbf{x}_\mathbf{I}^l)(\mathrm{x}_i-\mathrm{x}_i^l) + \sum_{t\in CT}\eta_t y_t \le slack_{g_i}^l \qquad n\in G_k,\ k\in K,\ l\in L\\
&\quad \mathbf{Ay}\le\mathbf{b}\\
&\quad \sum_{i\in B}\mathrm{y}_i - \sum_{i\in N}\mathrm{y}_i \le |B| - 1\\
&\quad y_k\cdot {x_k}^L \le x_k \le y_k\cdot {x_k}^U \qquad k\in K\\
&\quad \mathbf{x}^\mathbf{L}\le\mathbf{x}\le\mathbf{x}^\mathbf{U}\\
&\quad \mathbf{x}\in\mathbb{R}^n,\ \mathbf{y}\in\{0;1\}^m,\ \alpha\in\mathbb{R},\ slack_f^l\in\mathbb{R}^+,\ slack_g^l\in\mathbb{R}^+ \qquad l\in L
\end{aligned} \tag{5}$$

Accumulated linearizations are obtained using as linearization points $(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l)$ those that are obtained from all the previous NLP subproblems, derivatives with respect to the continuous decision variables, and delta contributions (γ_i, η_i) associated to each conditional tray. It should be noted that the set L defines all the previous main iterations, including the ones of the set covering problem. Moreover, slack variables are added to the linearization cuts to handle feasible region and objective function nonconvexities. They allow to find possible better solutions in the surrounding of the linearized nonconvex feasible region (Viswanathan & Grossmann, 1990). It should be noted that the integer constraints in the form $\mathbf{Ay} \leq \mathbf{b}$, are the translation of the propositional logic in Eq.(3) that defines the superstructure topology in the GDP representation (Grossmann & Trespalacios, 2013).

Derivatives of objective function and constraints are found to be very important for the effectiveness of the master problem. In fact, an inaccurate linearization of the objective function can cut-off optimal solutions, while inaccurate linearizations of constraints can cut-off feasible region areas (potentially excluding the optimal solution). Using a process simulator does not allow one to directly access analytical derivatives. For this reason, the finite difference based on perturbations has been adopted. From an analytical point of view, perturbations should be as small as possible to ensure good derivative estimates. However, it is crucial to mention that decreasing the value of the perturbation results in increasing the error due to the noise of the NLP solvers embedded within process simulators. As a result, there is a trade-off between accuracy and noise for small values of the perturbation, and depending on the decision variable and on the linearization point, there is an optimal range of perturbation. For this reason, a perturbation test is performed for each continuous decision variable at each MILP main iteration. The derivative is estimated starting from a larger perturbation; then the perturbation is halved and this procedure is iterated until the relative change of the derivative is below a certain tolerance. Typical values of the perturbations are in the range of 10^{-3} to 10^{-4}.

Delta contributions of a conditional tray for the objective function (γ_i) and for nonlinear constraints (η_i) are computed by converging a flowsheet that differs from the i-th linearization point only for the number of trays of the distillation section to which the conditional tray belongs, as outlined in Figure 4.

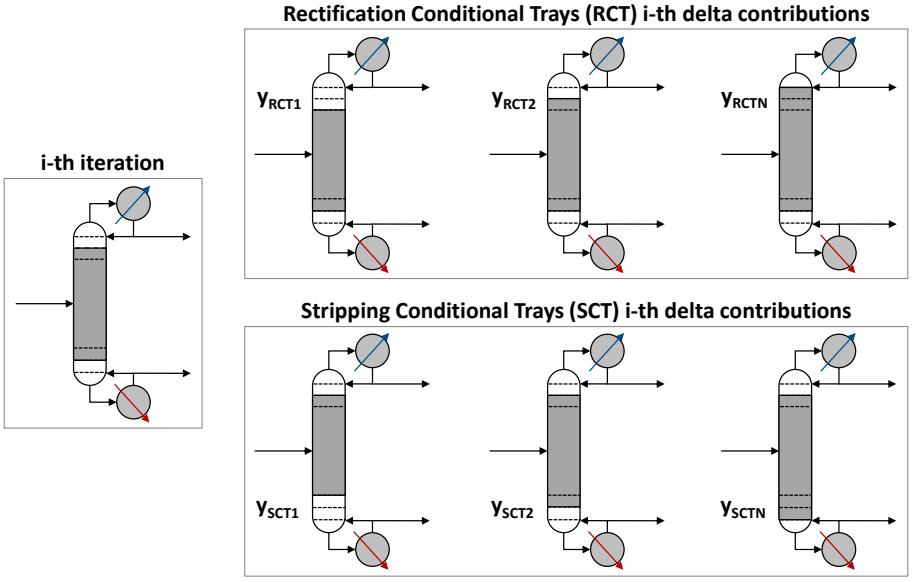


Figure 4: Procedure to evaluate delta contributions of rectification and stripping conditional trays (Caballero et al., 2005).

6. Solution strategy and remarks

The success of this optimization algorithm is strongly influenced by a careful formulation of the problem and by the selection of proper parameters and settings.

The problem should be formulated selecting continuous and discrete decision variables that lead to a fast and easy convergence of the process simulator flowsheet with all the values in the range between lower and upper bounds. This is true especially for distillation columns, for which two specifications should be provided, along with the number of trays and feed tray. The two continuous decision variables are usually imposed as reflux ratio and bottom to feed ratio, or, in case of distillation sequencing, they can be the light key component recovery at the overhead and the heavy key component recovery at the bottom. It is sometimes useful to insert a few fixed trays in the rectification and in the stripping sections. This means that the permanent stages are not only the feed tray, the condenser and the reboiler, but also a suitable small number of trays above and below the feed, based on the separation requirements that can be checked at the end of the optimization procedure. This strategy helps the convergence of the column and avoids the solution of columns with a number of trays that is too small (such as only 3 trays), thus reducing the number of evaluation for the conditional trays delta contributions and speeding up the solution of the MILP master problem.

On the other hand, there are two main sets of parameters to consider: the ones related to the optimization algorithm, and the ones related to the SimSci PRO/II process simulator settings. Optimization parameters include perturbation step size and penalty weights for the NLP and MILP problems. Perturbation step size is selected based on the aforementioned adaptive step size strategy. Penalty weights for the violation of nonlinear constraints can be selected with the same order of magnitude of the objective function and can be successively increased with a SUMT procedure (Correia et al., 2010), while penalty weights for slack variables are found not to have a significant impact on the solution of the optimization problem. SimSci PRO/II simulation settings, in turn, should be carefully tuned, depending on the distillation column that is to be solved. At first, a suitable distillation algorithm should be selected. In this work, the CHEMDIST algorithm (Bondy, 1991) has been selected for the VLLE heteroextractive distillation column case study, while the INPUT-OUTPUT algorithm (Russell, 1983) for the other conventional columns. To ensure the convergence of the columns, three key settings are mandatory, i.e. initial estimates model, damping factor, and homotopy. The initialization model provides temperature estimates and either vapor or liquid molar flow rate estimates to initiate the iterative calculations. The selection of a suitable method equally depends on the distillation system. The CHEMICAL initialization model is adopted for complex thermodynamic systems and it is based on a multi-flash technique to bring the profiles closer to the final solution before the column algorithm takes over. Another option is to use the CONVENTIONAL initialization model that it is based on the Fenske shortcut model; it is less CPU intensive and it is recommended when nonidealities are less strong. On the other side, the damping factor should be reduced from 1 to a value below 0.5 for oscillating systems, while homotopy could be adopted if the previous strategies fail.

Typical problem size should not exceed 10 continuous decision variables in the NLP subproblems, due to the intrinsic limit of DFO solvers (Rios & Sahinidis, 2013), but can reach several hundreds of discrete decision variables in the MILP master problems. The overall number of continuous decision variables in the MILP could be much greater than 10, because in the NLP subproblems only selected conditional units are accounted for, along with their decision variables. The ratio between computing times for the NLP and MILP problems is proportional to the ratio between continuous and discrete

decision variables, and the MILP computing time is largely due to the evaluation of conditional trays delta contributions for objective function and constraints.

7. Application examples

The optimization algorithm is applied to four different case studies, dealing with the purification of aqueous solutions of oxygenated chemicals derived from biorefining. These flowsheets essentially include distillation columns with a highly nonideal liquid phase, considered with the UNIQUAC liquid activity coefficient model. The first case study introduces this methodology to a single distillation column with one homogeneous azeotropes; the second case study involves a two-liquid phase heteroextractive distillation column with an heterogeneous azeotrope; the third case study includes the dewatering section of a downstream process with thermally coupled multieffect distillation columns; finally, the last example addresses the distillation sequencing of a bio-mixture.

For each application example, scope, thermodynamic modeling, problem formulation and results are provided, along with remarks and comments. The significant importance of developing reliable VL(L)E thermodynamic models based on nonlinear regression on experimental data is highlighted in the following section.

7.1 Single distillation column

The first case study takes in consideration a single multicomponent distillation column that realizes the cut between 1,2-propylene glycol (1,2-PG) and ethylene glycol (EG), which are mixed with other co-products obtained from the hydroprocessing of lignocellulosic sugars. The feed is composed by 40 wt.% EG, 50 wt.% 1,2-PG and 10 wt.% of mixed heavier oxygenated chemicals that form also an homogeneous azeotrope with EG. The distillation column with structured packing internals is operated at atmospheric pressure with a total condenser and a kettle reboiler. The aim of the optimization is to determine the optimal number of trays, feed tray location, and the value of the two specifications (reflux ratio and bottom flowrate) that minimize TAC, with purity constraints at the top and at the bottom for 1,2-PG and EG, respectively.

The first step for an effective optimization of bio-mixtures is the definition of a reliable thermodynamic model. Binary interaction parameters and predictive models included in process simulators for this kind of components are most of the times not reliable. For this reason, UNIQUAC binary interaction parameters have been estimated by nonlinear regression with the `BzzNonlinearRegression` class of the BzzMath library (Buzzi-Ferraris & Manenti, 2009; G. Buzzi-Ferraris & F. Manenti, 2010; Pirola et al., 2014), using published phase equilibrium experimental data (Yang et al., 2014; Zhang et al., 2013; Zhong et al., 2014). Only two representative VLE Txy plots are reported in Figure 5 for the sake of conciseness.

The first guess for the decision variables has been derived by preliminary short-cut evaluations and involves a column with 123 stages (HETP = 0.250 m) with the feed at the 92nd stage. The evolution of the objective function over the main iterations is reported for both the MILP lower bound and the NLP upper bound in Figure 6a. The optimal configuration is found at the third iteration (red dashed line), with 135 stages and the feed at the 98th stage, because there is the simultaneous crossing between lower and upper bounds and the worsening of the NLP. Table 1 summarizes the most important indicators of the case study 1. It is worth noting that the number of discrete decision variables (40) derives from the number of conditional trays, 20 for the rectification and 20 for the stripping sections.

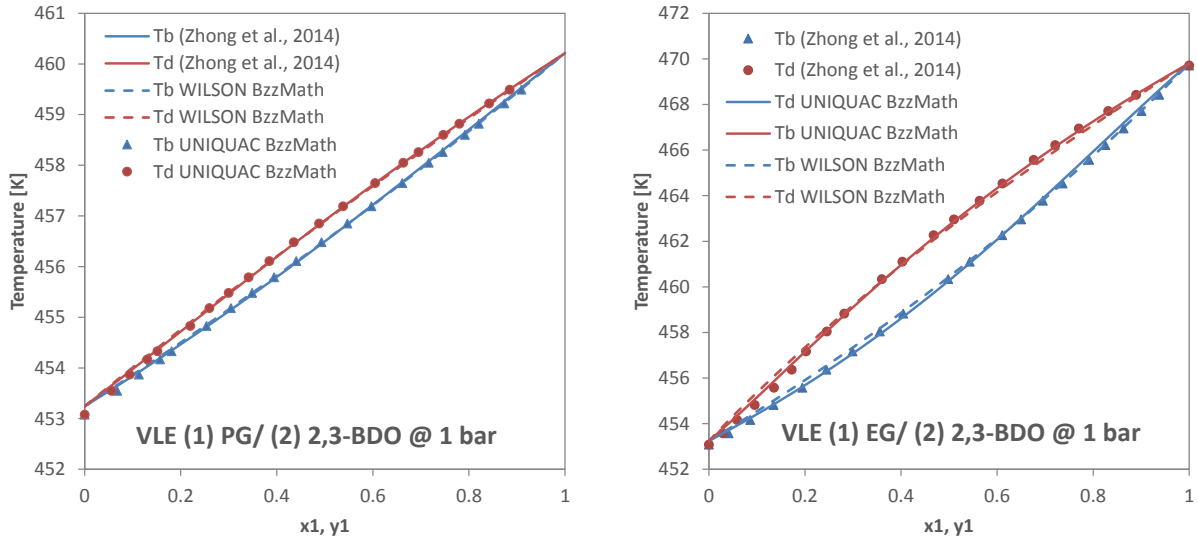


Figure 5: Txy VLE plots of 1,2-PG and EG with 2,3-butanediol at 1 bar. Experimental data (symbols) and model predictions (lines).

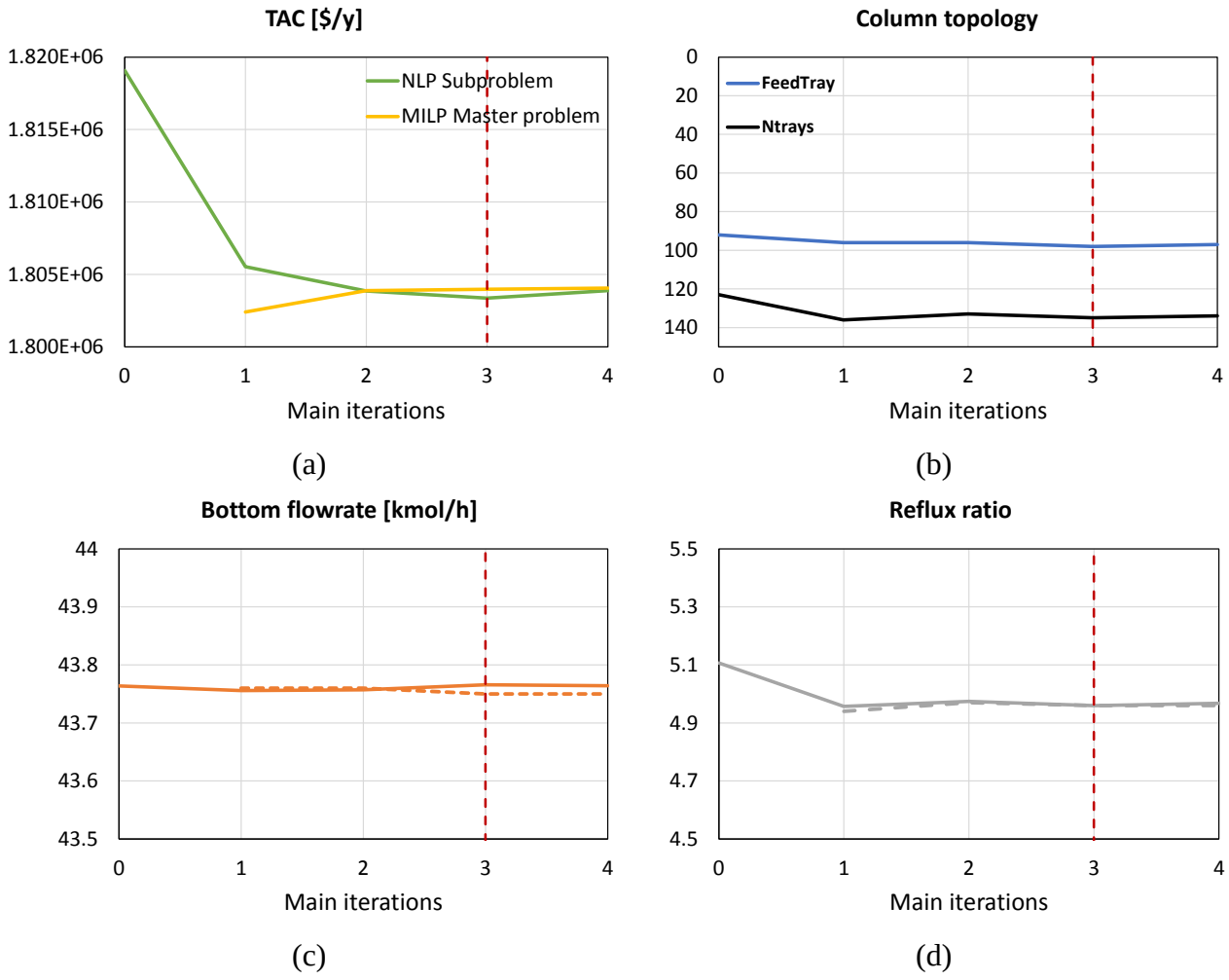


Figure 6: Evolution of objective function (a), structural decision variables (b) and continuous decision variables (c-d) of case study 1 as a function of the OA main iterations. The dashed line reports the first guess given by the MILP master problem.

Table 1: Summary of case study 1.

Number of continuous decision variables	2	
Number of discrete decision variables	40	
Number of nonlinear constraints	2	
Objective function of the first guess	3,986,700	$/y
Objective function of the optimal solution	1,803,370	$/y

The importance of adopting a rigorous optimization algorithm, such as the OA for dealing with mixed-integer problems is emphasized by comparing the optimal solution obtained with a heuristic "rounding" method. This method simply considers the number of stages and feed stage as continuous variables, rounding the value of the variable to the closest integer value. In this way, only one NLP problem is solved optimizing simultaneously the two specifications and the two structural decision variables. Implementing the rounding method a ≈1.5 % larger value of the cost objective function is obtained (1.83E6 $/y), even if this is a simple case study and the number of stages is large (pseudo-continuous variable). The optimal feed stage and number of stages are 83 and 126 (vs. 98 and 135), respectively (Figure 7). Larger discrepancies are found in the case of complex configurations and columns with a smaller numbers of trays, as shown in the following example (Subsection 7.2).

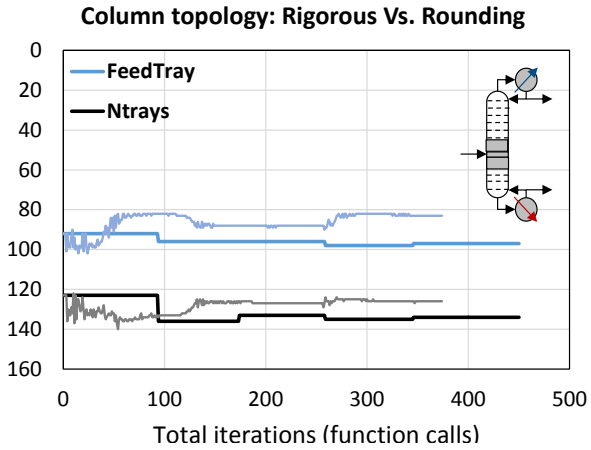


Figure 7: Comparison between the column topology obtained with the rigorous OA algorithm (thick lines) and the rounding algorithm (thin lines) as a function of the NLP function calls.

The rigorous OA algorithm requires 4 NLP subproblems while the rounding method requires only one NLP problem. Nevertheless, this last NLP problem is much larger (4 decision variables instead of 2), and the OA NLP subproblems have a good first guess given by the MILP master problems that is close to the optimal solution. For these two main reasons, both methods require about 400 function calls. The CPU time is about 30 minutes, and it is directly related to the number of black-box calls, where each call requires about 3.5 seconds. The final MILP master problem (4^{th} iteration) is solved instantaneously and involves 18 equations, 56 continuous variables, and 40 binary variables.

7.2 Heteroextractive distillation column

The second case study involves the purification of acetic acid from water. Downstream operations in the cellulose acetate industry and in the terephthalic and isophthalic acids production processes involve the separation of water (W) and acetic acid (AcAc) mixtures. Considering the vapor-liquid equilibrium of the W/AcAc binary system, a tangent pinch on the pure water end makes the separation of these two components difficult. In order to overcome this technical issue, previous experimental

and modeling studies (Pirola et al., 2014) showed that p-xylene (pX) is a suitable entrainer to operate a heterogeneous extractive distillation. The miscibility gap between W and pX is exploited to separate the distillate stream in a decanter where an aqueous and an organic phase are obtained. The organic phase is composed of nearly pure pX that is recycled back as entrainer to the top section of the column, while the aqueous stream with traces of AcAc is withdrawn and sent to wastewater treatment. From the bottom of the heteroextractive column, a stream composed by AcAc, and the excess of the extractive entrainer is recycled back to the upstream synthesis process.

Reliable thermodynamic models, based on the UNIQUAC liquid activity model, have been developed elsewhere (Pirola et al., 2014), and are hereinafter adopted to perform the economic optimization of the heteroextractive column (Figure 8).

The distillation column is simulated with the CHEMDIST VLLE algorithm, which is based on the full Newton-Raphson method and is able to deal with both single and double liquid phase mixtures, choosing in each case the appropriate set of interaction parameters. It should be stressed that the use of an incorrect distillation model, such as the conventional VLE distillation model, would result in a completely incorrect simulation of the system, leading to a distorted optimization result. The minimization of the cost objective function (TAC) is subject to two purity specification constraints, one imposed as a design specification for the convergence of the column (W in the bottom < 3 % w/w), and the other applied as an implicit constraint (AcAc in the top aqueous stream ≤ 5000 ppm). The aim of the optimization is to determine the optimal number of trays, feed tray location, entrainer feed tray location, the value of the specification (reflux ratio), pX entrainer flowrate, and entrainer temperature.

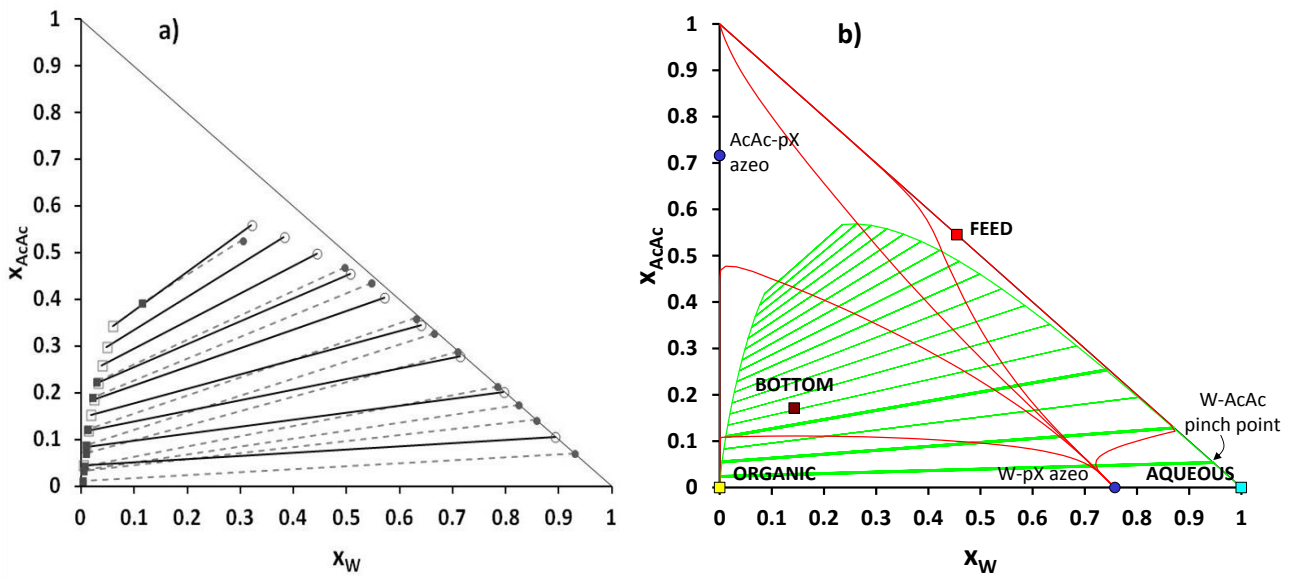


Figure 8: W/AcAc/pX VLLE. (a) Experimental data (dotted line) (Murogova et al., 1971) and model predictions (solid line) with estimated binary parameters. (b) Distillation column stream composition for the optimized configuration.

The most important indicators of the case study 2 are summarized in Table 2, while optimization results are reported in Figure 9. The results show that the optimal configuration involves a 7 stage column with most of the stages being in between the feed and the entrainer's feed locations. The entrainer is fed to the 2nd stage (1st stage is the condenser), while the W/AcAc mixture is fed to the 6th stage (7th stage is the reboiler). The reflux ratio is very small and the entrainer stream is just below the water boiling point and with a molar flowrate almost double of the W/AcAc feed stream. Total

annualized costs are largely due to column purchase and installation costs (≈ 70 %), and reboiler and steam costs (≈ 20 %).

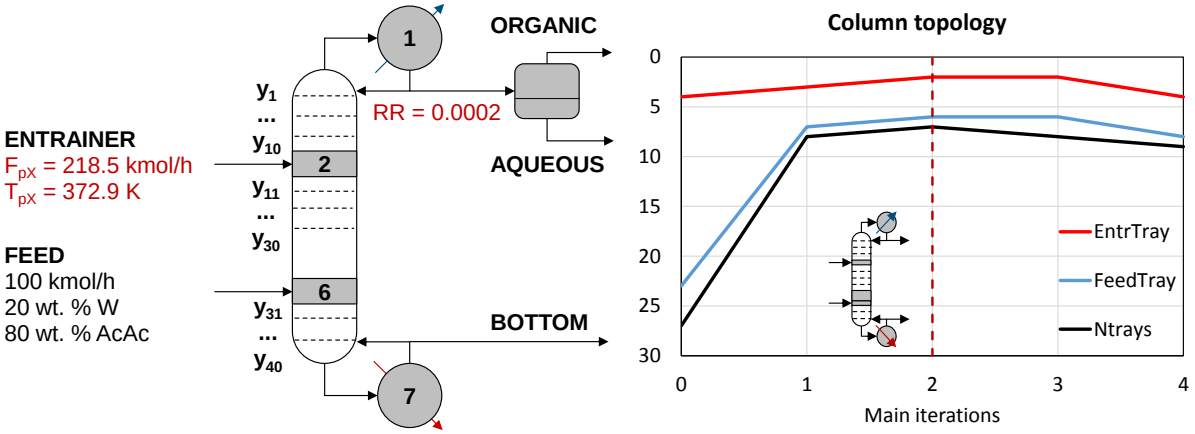


Figure 9: Optimized configuration (a), and structural decision variables (b) of case study 2 as a function of the OA main iterations.

Table 2: Summary of case study 2.

Number of continuous decision variables	3	
Number of discrete decision variables	40	
Number of nonlinear constraints	1	
Objective function of the first guess	507,660	$/y
Objective function of the optimal solution	115,170	$/y

A comparison between the rigorous AO and the rounding algorithms is also addressed here. In this case, the first guess for the column topology is voluntarily far from the optimal solution (27 vs. 7 stages) and the rounding method is not able to find a good solution (89.9 % larger costs with respect to the OA algorithm). The rounding algorithm's optimal solution is stuck to 24 total stages after about ¼ of function calls of the rigorous algorithm, while the OA algorithm finds the optimal value of 7 stages.

7.3 Dewatering process section

As mentioned before, the dewatering processing of bio-mixtures is an energy intensive step that can significantly contribute to the total costs of the downstream operation. Thermal integration, and more specifically multieffect distillation (Ahmetovic et al., 2010; Karuppiah et al., 2008), can substantially reduce heat duties, thus decreasing total costs.

The dewatering task of a diluted aqueous stream (85 wt.% of water) of mixed bio-based glycols is addressed here. Two different process technologies are assessed (Figure 10). Namely, simple distillation at atmospheric pressure (column LP1) and multieffect distillation with two heat integrated distillation columns operating at different pressures (columns LP2 and HP2). The optimization targets are to select the proper technology, and to optimize the selected columns in terms of both structural and operating parameters. Continuous decision variables include reflux ratio and B/F of the selected columns, split ratio between HP2 and LP2 columns (for disjunction 2), and condenser pressure of the HP2 column (for disjunction 2). Integer decision variables include the number of trays and feed tray of the selected columns, and the decision for the selection of the technology 1 (atmospheric

distillation) or technology 2 (multieffect distillation). A glycols recovery specification and a water purity specification are considered as implicit global constraints, while a conditional constraint related to the feasibility for the heat integration between HP2 and LP2 columns applies only to the second disjunction. This last constraint forces the temperature of the HP2 condenser to be greater than the temperature of the LP2 reboiler plus the minimum approach temperature (EMAT = 5 °C) as shown in Eq.(6).

$$
\begin{aligned}
& Y_1 \dot{\vee} Y_2 \\
& \begin{bmatrix} Y_1 \\ \begin{bmatrix} Y_i \\ TAC_{1i} = \gamma_i \end{bmatrix} \vee \begin{bmatrix} \neg Y_i \\ TAC_{1i} = 0 \end{bmatrix} \\ TAC_1 = \sum_{i \in CT1} \gamma_i(\mathrm{x}, \mathrm{y}, \mathrm{x}_1) \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ TAC_1 = 0 \end{bmatrix} \quad i \in CT_1 \\
& \begin{bmatrix} Y_2 \\ \begin{bmatrix} Y_i \\ TAC_{2i} = \gamma_i \end{bmatrix} \vee \begin{bmatrix} \neg Y_i \\ TAC_{2i} = 0 \end{bmatrix} \\ T_C^{HP}(\mathrm{x}, \mathrm{y}, \mathrm{x}_1) \geq T_R^{LP}(\mathrm{x}, \mathrm{y}, \mathrm{x}_1) + \mathrm{EMAT} \\ TAC_2 = \sum_{i \in CT2} \gamma_i(\mathrm{x}, \mathrm{y}, \mathrm{x}_1) - CQ_{EX} \end{bmatrix} \vee \begin{bmatrix} \neg Y_2 \\ TAC_2 = 0 \end{bmatrix} \quad i \in CT_2
\end{aligned}
\tag{6}
$$

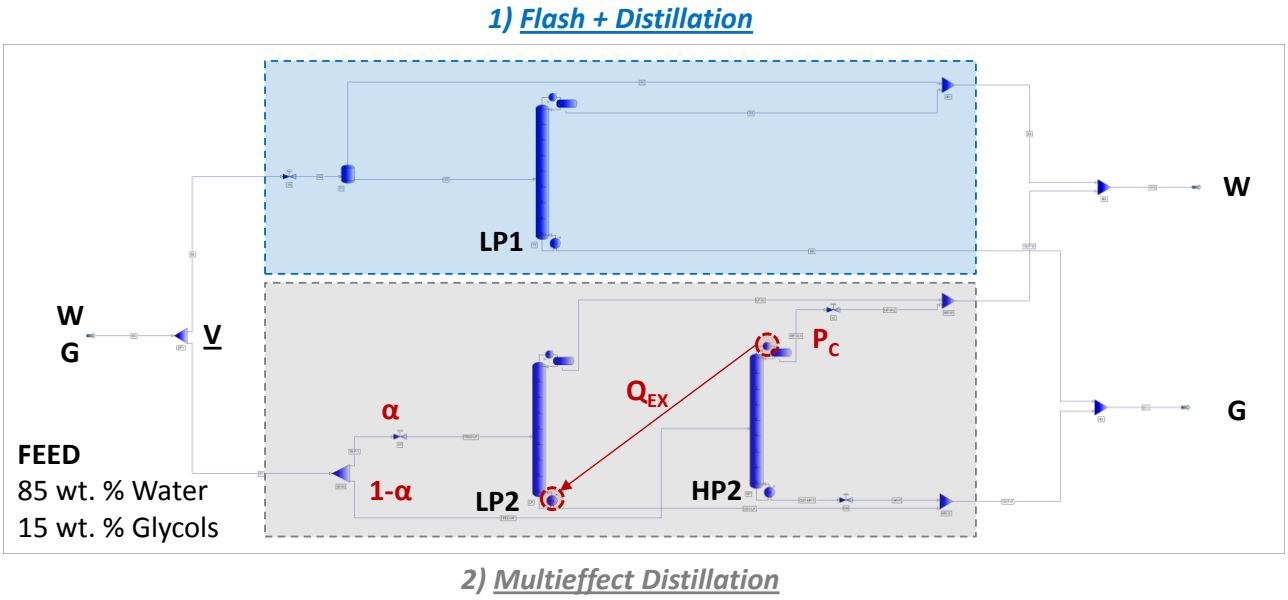


Figure 10: Superstructure of the dewatering task of case study 3.

The results summarized in Figure 11 and Table 3 show that the multieffect distillation is the most favorable technology. Even in the initial (non optimized) set covering problem, the total cost associated with the single atmospheric distillation is ≈ 5 % higher than the one associated with the heat integrated double configuration, while the optimized solution further reduces costs to about 30 %. Interestingly, the two heat integrated distillation columns have similar topology and the same number of optimal stages. The ratio between rectification to stripping sections is about 3:1 for the LP2 column and 2:1 for HP2 column, with small reflux ratios. The split ratio between LP2 and HP2

is close to 50 %, while the HP2 condenser pressure is in the order of 10 bar. The heat integration between HP2 condenser and LP2 reboiler is greater than the 99.9 % ($Q_R^{LP} \approx Q_C^{HP}$) of the condenser duty, and it allows to save steam utility costs (CQ_{EX}) of 122,289 $/y, while TAC of LP2 and HP2 are 169,813 and 163,997 $/y respectively.

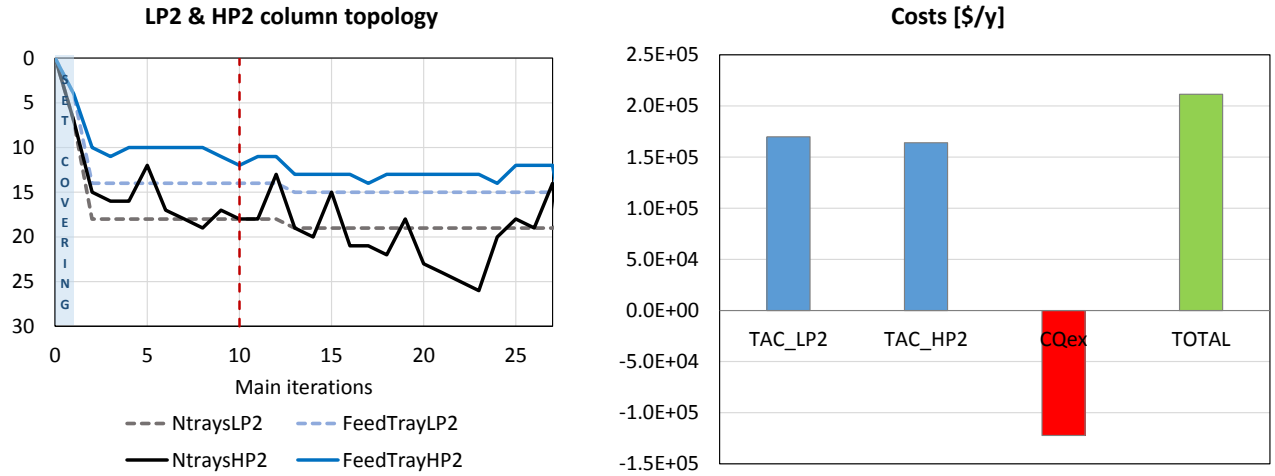


Figure 11: Case study 3, dewatering section. (a) Column topology as a function of the OA main iterations. (b) Cost distribution of the optimal solution.

Table 3: Summary of case study 3.

Number of continuous decision variables	8	
Number of discrete decision variables	62	
Number of nonlinear constraints	3	
Objective function of the best first guess among the two set covering configurations	300,492	$/y
Objective function of the optimal solution	211,521	$/y

7.4 Distillation sequencing

This subsection deals with the rigorous design of the distillation sequencing for a highly nonideal quaternary mixture of mixed bio-based glycols and butanediols (BDOs) with a homogeneous minimum boiling azeotrope between EG and 1,2-BDO (Yang et al., 2014).

The feed mixture is composed by 50 wt.% 1,2-PG, 40 wt.% EG, and 10 wt.% of mixed butanediols (BDOs) at the bubble point. The design is achieved by considering a superstructure for the quaternary mixture (in which one pseudo-component, C, is the azeotrope) with 10 atmospheric distillation columns (Figure 12), and optimizing both the selection of the columns and their conditional trays, as well as two continuous decision variables to specify the operation of each active column. In this case, the reflux ratio and the bottom total molar flowrate are selected as continuous decision variables to help the convergence of the columns by providing exact estimates for reflux ratio and product flowrates. Four implicit inequality constraints are considered for the purity specifications (mole fractions) and for the mass recoveries of the two main products (i.e. EG and 1,2-PG).

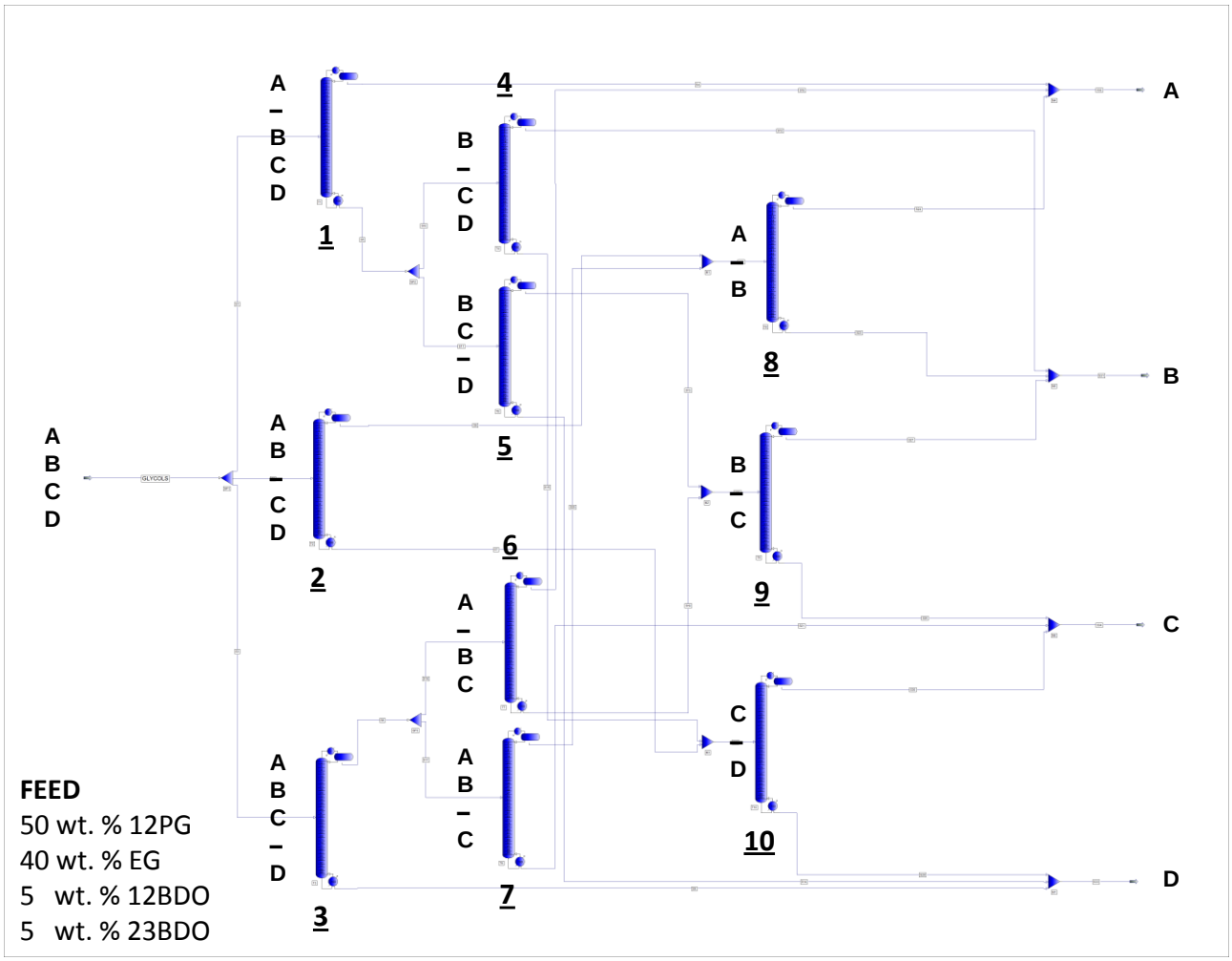


Figure 12: Superstructure of the distillation sequencing for a quaternary mixture (case study 4).

The propositional logic defining the superstructure is reported in Eq.(7) for the sake of completeness, and results in the decision among 5 possible separation process layouts.

$$
\begin{aligned}
&Y_1 \dot{\vee} Y_2 \dot{\vee} Y_3 \\
&Y_4 \dot{\vee} Y_5 \\
&Y_6 \dot{\vee} Y_7 \\
&Y_8 \dot{\vee} Y_9 \dot{\vee} Y_{10} \\
&Y_1 \Rightarrow Y_4 \vee Y_5 \\
&Y_2 \Rightarrow Y_8 \wedge Y_{10} \\
&Y_3 \Rightarrow Y_6 \vee Y_7 \\
&Y_4 \Rightarrow Y_{10} \\
&Y_5 \Rightarrow Y_9 \\
&Y_6 \Rightarrow Y_9 \\
&Y_7 \Rightarrow Y_8
\end{aligned}
\tag{7}
$$

An additional cut, Eq.(8), is added to the constraints of the MILP master problems to enforce the selection of only three distillation column, which is the minimum number required for the separation of four components (NC - 1).

$$\sum_{i \in COL} y_i = 3 \tag{8}$$

The results are summarized in Figure 13 and Table 4. The selected configuration is the sequence of columns 2/8/10 that first realize the central cut between 1,2-PG and EG and then purify the two main products from butanediols (BDOs). Considering the set covering problem, keeping the same total number of trays, the configurations 1/4/10, 1/5/9, 3/6/9, 3/7/8 have 24 %, 96 %, 67 % and 64 % higher extra total annualized costs. The second best configuration is the direct sequence (1/4/10), while the worst configuration is the 1/5/9 that requires higher reflux ratios to avoid violating the purity and recovery constraints. For the optimized columns topology, the cost distribution among the selected columns is 54 % for column 2, 29 % for column 8 and 18 % for column 10. The OPEX/CAPEX ratio is in the range of 0.90-1.25.

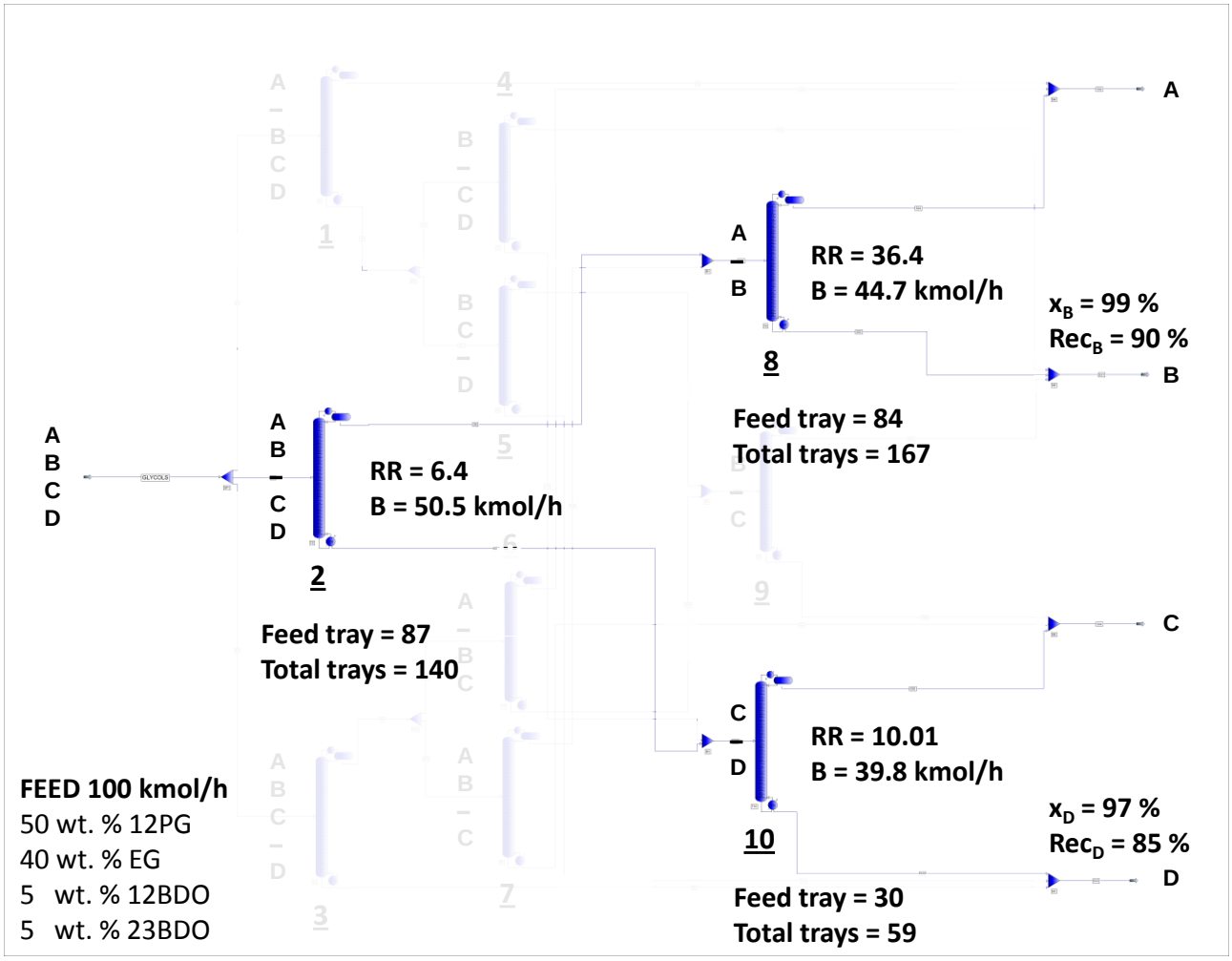


Figure 13: Case study 4. Optimal distillation sequence and decision variables.

Table 4: Summary of case study 4.

Number of continuous decision variables	20	
Number of discrete decision variables	210	
Number of nonlinear constraints	4	
Objective function of the best first guess among all set covering configurations	3,249,840	$/y
Objective function of the optimal solution	3,133,220	$/y

Conclusions

This paper has presented a new methodology for the optimal synthesis of downstream processes based on the rigorous models embedded within the process simulator SimSci PRO/II, and on a deterministic Mixed-Integer Nonlinear Programming algorithm (Outer Approximation), implemented in C++ and GAMS. The effectiveness of this procedure was demonstrated with four different case studies in the field of biorefining.

It is important to note that the parametric optimization can involve only a limited number of continuous decision variables due to the intrinsic limitation of DFO solvers for large problems. Nevertheless, the restrictions on the scale of the combinatorial problem, related with the structural optimization are less limiting.

Future work will address in more detail the sequencing with thermal integration and the comparative study of different DFO solvers for the solution of NLP subproblems.

Acknowledgments

Biochemtex S.p.a. is gratefully acknowledged for partially founding this research, while Prof. Caballero, Prof. Sahinidis, and Prof. Pirola are gratefully acknowledged for their suggestions and fruitful discussions.

References

Ahmetovic, E., Martin, M., & Grossmann, I. E. (2010). Optimization of Energy and Water Consumption in Corn-Based Ethanol Plants. Industrial & Engineering Chemistry Research, 49, 7972-7982.

Barttfield, M., Aguirre, P. A., & Grossmann, I. E. (2004). A decomposition method for synthesizing complex column configurations using tray-by-tray GDP models. Computers & Chemical Engineering, 28, 2165-2188.

Biegler, L. T. (1985). Improved Infeasible Path Optimization for Sequential Modular Simulators .1. The Interface. Computers & Chemical Engineering, 9, 245-256.

Bondy, R. W. (1991). A New Distillation Algorithm for Non-Ideal System. In AIChE Annual Meeting.

Bravo-Bravo, C., Segovia-Hernandez, J. G., Gutierrez-Antonio, C., Duran, A. L., Bonilla-Petriciolet, A., & Briones-Ramirez, A. (2010). Extractive Dividing Wall Column: Design and Optimization. Industrial & Engineering Chemistry Research, 49, 3672-3688.

Brunet, R., Reyes-Labarta, J. A., Guillen-Gosalbez, G., Jimenez, L., & Boer, D. (2012). Combined simulation-optimization methodology for the design of environmental conscious absorption systems. Computers & Chemical Engineering, 46, 205-216.

Buzzi-Ferraris, G. (1967). Ottimizzazione di funzioni a più variabili. Nota I. Variabili non vincolate. Ing. Chim. It., 3, 101.

Buzzi-Ferraris, G., & Manenti, F. (2009). Kinetic models analysis. Chemical Engineering Science, 64, 1061-1074.

Buzzi-Ferraris, G., & Manenti, F. (2010). A combination of parallel computing and object-oriented programming to improve optimizer robustness and efficiency. Computer Aided Chemical Engineering, 28, 337-342.

Buzzi-Ferraris, G., & Manenti, F. (2010). Interpolation and regression models for the chemical engineer: Solving numerical problems.

Buzzi-Ferraris, G., & Manenti, F. (2012). BzzMath: Library Overview and Recent Advances in Numerical Methods. Computer-Aided Chemical Engineering, 30, 1312-1316.

Caballero, J. A., Milan-Yanez, D., & Grossmann, I. E. (2005). Rigorous design of distillation columns: Integration of disjunctive programming and process simulators. Industrial & Engineering Chemistry Research, 44, 6760-6775.

Chen, G. Q. (2009). A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. Chemical Society Reviews, 38, 2434-2446.

Chen, Y., Eslick, J. C., Grossmann, I. E., & Miller, D. C. (2015). Simultaneous Process Optimization and Heat Integration Based on Rigorous Process Simulations. Computers and Chemical Engineering.

Corbetta, M., Manenti, F., Pirola, C., Tsodikov, M. V., & Chistyakov, A. V. (2014). Aromatization of propane: Techno-economic analysis by multiscale "kinetics-to-process" simulation. Computers & Chemical Engineering, 71, 457-466.

Correia, A., Matias, J., Mestre, P., & Serôdio, C. (2010). Direct-search penalty/barrier methods. Proceedings of The World Congress on Engineering 2010, 3, 1729-1734.

Dias, M. O. S., Ensinas, A. V., Nebra, S. A., Maciel, R., Rossell, C. E. V., & Maciel, M. R. W. (2009). Production of bioethanol and other bio-based materials from sugarcane bagasse: Integration to conventional bioethanol production process. Chemical Engineering Research & Design, 87, 1206-1216.

Diaz, M. S., & Bandoni, J. A. (1996). A mixed integer optimization strategy for a large scale chemical plant in operation. Computers & Chemical Engineering, 20, 531-545.

Douglas, J. M. (1988). Conceptual design of chemical processes. New York.

Eslick, J. C., & Miller, D. C. (2011). A multi-objective analysis for the retrofit of a pulverized coal power plant with a CO2 capture and compression process. Computers & Chemical Engineering, 35, 1488-1500.

Garcia, N., Fernandez-Torres, M. J., & Caballero, J. A. (2014). Simultaneous environmental and economic process synthesis of isobutane alkylation. Journal of Cleaner Production, 81, 270-280.

Gross, B., & Roosen, P. (1998). Total process optimization in chemical engineering with evolutionary algorithms. Computers & Chemical Engineering, 22, S229-S236.

Grossmann, I. E., & Trespalacios, F. (2013). Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming. Aiche Journal, 59, 3276-3295.

Gutierrez-Antonio, C., & Briones-Ramirez, A. (2009). Pareto front of ideal Petlyuk sequences using a multiobjective genetic algorithm with constraints. Computers & Chemical Engineering, 33, 454-464.

Harsh, M. G., Saderne, P., & Biegler, L. T. (1989). A Mixed Integer Flowsheet Optimization Strategy for Process Retrofits - the Debottlenecking Problem. Computers & Chemical Engineering, 13, 947-957.

Hermann, B. G., & Patel, M. (2007). Today's and tomorrow's bio-based bulk chemicals from white biotechnology - A techno-economic analysis. Applied Biochemistry and Biotechnology, 136, 361-388.

Karuppiah, R., Peschel, A., Grossmann, I. E., Martin, M., Martinson, W., & Zullo, L. (2008). Energy optimization for the design of corn-based ethanol plants. Aiche Journal, 54, 1499-1525.

Leboreiro, J., & Acevedo, J. (2004). Processes synthesis and design of distillation sequences using modular simulators: a genetic algorithm framework. Computers & Chemical Engineering, 28, 1223-1236.

Murogova, R. A., Tudorovskaya, G. L., Pleskach, N. I., Safonova, N. A., Gridin, I. D., & Serafimov, L. A. (1971). Dampf fluessing gleichgewichtim system wasser-essigsaeure-p-xylolbei 760 mm Hg (Vol. 46). Leningrad.

Navarro-Amoros, M. A., Caballero, J. A., Ruiz-Femenia, R., & Grossmann, I. E. (2013). An alternative disjunctive optimization model for heat integration with variable temperatures. Computers & Chemical Engineering, 56, 12-26.

Navarro-Amoros, M. A., Ruiz-Femenia, R., & Caballero, J. A. (2014). Integration of modular process simulators under the Generalized Disjunctive Programming framework for the structural flowsheet optimization. Computers & Chemical Engineering, 67, 13-25.

Pirola, C., Galli, F., Manenti, F., Corbetta, M., & Bianchi, C. L. (2014). Simulation and Related Experimental Validation of Acetic Acid/Water Distillation Using p-Xylene as Entrainer. Industrial & Engineering Chemistry Research, 53, 18063-18070.

Rios, L. M., & Sahinidis, N. V. (2013). Derivative-free optimization: a review of algorithms and comparison of software implementations. Journal of Global Optimization, 56, 1247-1293.

Russell, R. A. (1983). A Flexible and Reliable Method Solves Single-Tower and Crude-Distillation-Column Problems. Chemical Engineering, 90, 53-59.

Sauer, M., Porro, D., Mattanovich, D., & Branduardi, P. (2008). Microbial production of organic acids: expanding the markets. Trends in Biotechnology, 26, 100-108.
Turkay, M., & Grossmann, I. E. (1996). Logic-based MINLP algorithms for the optimal synthesis of process networks. Computers & Chemical Engineering, 20, 959-978.
Vazquez-Castillo, J. A., Venegas-Sanchez, J. A., Segovia-Hernandez, J. G., Hernandez-Escoto, H., Hernandez, S., Gutierrez-Antonio, C., & Briones-Ramirez, A. (2009). Design and optimization, using genetic algorithms, of intensified distillation systems for a class of quaternary mixtures. Computers & Chemical Engineering, 33, 1841-1850.
Viswanathan, J., & Grossmann, I. E. (1990). A Combined Penalty-Function and Outer-Approximation Method for Minlp Optimization. Computers & Chemical Engineering, 14, 769-782.
Xiu, Z. L., & Zeng, A. P. (2008). Present state and perspective of downstream processing of biologically produced 1,3-propanediol and 2,3-butanediol. Applied Microbiology and Biotechnology, 78, 917-926.
Yang, Z., Xia, S. Q., Shang, Q. Y., Yan, F. Y., & Ma, P. S. (2014). Isobaric Vapor Liquid Equilibrium for the Binary System (Ethane-1,2-diol + Butan-1,2-diol) at (20, 30, and 40) kPa. Journal of Chemical and Engineering Data, 59, 825-831.
Yeomans, H., & Grossmann, I. E. (1999). A systematic modeling framework of superstructure optimization in process synthesis. Computers & Chemical Engineering, 23, 709-731.
Zhang, L. H., Wu, W. H., Sun, Y. L., Li, L. Q., Jiang, B., Li, X. G., Yang, N., & Ding, H. (2013). Isobaric Vapor-Liquid Equilibria for the Binary Mixtures Composed of Ethylene Glycol, 1,2-Propylene Glycol, 1,2-Butanediol, and 1,3-Butanediol at 10.00 kPa. Journal of Chemical and Engineering Data, 58, 1308-1315.
Zhong, Y., Wu, Y. Y., Zhu, J. W., Chen, K., Wu, B., & Ji, L. J. (2014). Thermodynamics in Separation for the Ternary System 1,2-Ethanediol+1,2-Propanediol+2,3-Butanediol. Industrial & Engineering Chemistry Research, 53, 12143-12148.