Improvement of the Optimal Design Procedure Using Randomized Algorithm and Process Simulators

Julian Cabrera-Ruiz[1*], Shinji Hasebe[2], and J. Rafael Alcantara-Avila[2]

[1]Chemical Engineering Department, University of Guanajuato, Guanajuato Campus, Noria Alta s/n, Guanajuato, Guanajuato, 36050, Mexico

[2]Deparment of Chemical Engineering, Kyoto University, Katsura Campus, Kyoto, 615-8510, Japan

Abstract Though a global optimization procedure using a randomized algorithm and a commercial process simulator is relatively easy to implement for complex design problems (i.e., intensified design processes), a dominant problem is their heavy computation load. As the process simulation is repeatedly executed to calculate the objective function, it is inevitable to spend long computation time to derive the optimal solution. Also, the randomized algorithms consider the treatment of all variables as continuous. Thus, the reduction of the number of iterations is crucial for such optimization procedures that include integer variables. In this work, an estimation procedure of the objective function having integer design variables is proposed. In the proposed procedure, the values of the objective function at the nodes of hyper-triangle that includes the suggested next search point are used to estimate the objective function, at the same time normalization of the design optimization variables is recommended. The procedure was implemented on the simulated annealing stochastic algorithm with a trivial case of a binary mixture in order to know the optimal solution and compare the traditional optimizations procedures and the proposed one. The proposed procedure show improvement not only for reducing the number iterations, but also for an increase of accuracy of finding the optimal solution.

1 Introduction

With the new tendencies of the intensified processes, many designs had been proposed. The new intensified processes design involves a big challenge to seek optimal design. Many strategies have been developed and applied for the optimization problem. Some works applied gradient-based approaches (i.e., deterministic) and others the randomization approaches (i.e., stochastic) (Cavazzuti, 2013). The advantages and disadvantages have been widely discussed. Both approaches have tackled the global optimization problems for intensified processes (Floudas and Gounaris, 2009; Ricardez-Sandoval *et al*., 2009; Yuan *et al*., 2012; Segovia-Hernández *et al*., 2015). In particular, randomization approaches are increasing their popularity due to their relatively easy implementation in combination with software frameworks (Cabrera-Ruiz *et al*., 2012; Vazquez-Castillo *et al*., 2009; Gitizadeh, *et al*., 2013; Santaella *et al*., 2014). As both approaches show advantages in the disadvantages with respect to each other, hybrid optimization has been suggested as a way to complement each other (Grossmann and Westerberg, 2000). Many works have proposed hybrid methods, either gradient-based (Floudas and Gounaris, 2009; Goryachih and Rachinskaya, 2017) and randomized methods (Sayyaadi *et al*., 2010; Gitizadeh *et al*., 2013) or hybrid between different method approach

(Renders and Flasse, 1996; Popović *et al*., 2014). Munawar and Gudy (2005) have pointed out that the "advantage" about to take all values as continuous can be a great "disadvantage" in terms of efficiency. Despite the large numbers of works in randomized optimization, only a few works had focused on the main issue to increase the efficiency of the design processes problems. Nevertheless, it is necessary to improve the efficiency of these algorithms, because of their use in solving optimization problems have been increased in order to find new intensified process.

This work focuses on improving randomized algorithms used in the design of chemical engineering processes that became highly popular due to the new software frameworks that allow optimizing the process using commercial simulator process (e.g., Aspen Plus). The way to handle the randomized algorithms is based in their original development through the solution of well known mathematical problems; however the optimization criteria of these algorithms cannot be the best for the design processes problems because of the several restrictions managed for the process, most of then physical than mathematical thought.

2 Methodology

---

* Corresponding author: j.cabreraruiz@ugto.mx

This section widely explains the main issues that considerably affect the efficiency of randomized algorithms used for the processes optimization; It is caused by excessive, and unnecessary iterations. The main factors have been tackled for: purity restrictions, the range value differences between variables, and the discrete variables handled as continuous.

2.1 Definition of the problem

Formally for the randomized algorithms, it is necessary to establish the range of the variables as large as are possible to allow the algorithm to move freely through to whole values region. This can be true for most of the mathematical problems, but as it is known, mathematical models of chemical processes show high stiffness caused by the model complexity itself, and the combination of the discrete and continuous values, and their restrictions, respectively.

Figure 1 shows two discrete values N_1 and N_2, where only the green dots correspond to real feasible points, and the rest of the shading area is unfeasible. In this work, the following notation is used: $N_{i,j}$ represents that element j is within the bounds of variable i where $N_{i,j} < N_{i,j+1}$.
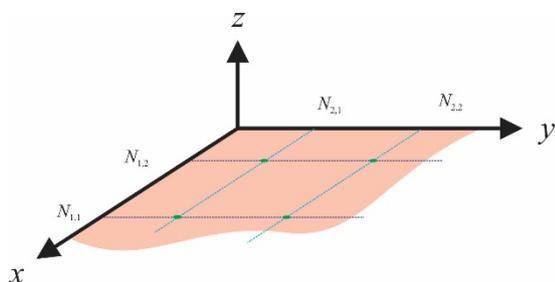

Figure 1. Discrete variables and the unfeasible area

Nevertheless, typical randomized algorithms take these discrete values as continuous as shown in Figure 2, where $N_{i,j} < N_{i,j+1}$ and i is the discrete variable and j the feasible element of the variable; therefore, any real value can be selected ($N_{1,R}$, $N_{2,R}$). Then, these values are enforced to take integer values by rounding them up (ceil), rounding them down (floor), or a regular rounding. In any case, a plain surface solution will be generated for all the values circumscribing the feasible points. Among the four feasible combinations, the point ($N_{1,1}$, $N_{2,2}$) has the lowest value of the objective function z, and it will never be taken because rounding up or down steps will evaluate the objective function at points ($N_{1,2}$, $N_{2,2}$) or ($N_{1,1}$, $N_{2,1}$), respectively. Only the area corresponding to the green surface will take the point ($N_{1,1}$, $N_{2,2}$). Eventually, searching the best point can spend a lot of computing time moving through this area, and only if the optimization algorithm searches other points, thus another area, unnecessarily calculations on the previous area with all its unfeasible points resulted in time lost. The situation is the same for any of the three rounding options, however despite can be considered a better option to choose a regular rounding, this option it is not very appropriated to move around the limits.

Also, derivated of the freedom of the variables many times all variables design had chosen optimization variables. Nevertheless we must consider that these variables have a direct influence to the restriction; furthermore, the behavior of these variables are mostly convex (e.g., reflux ratio and distillate component purity), then used as optimization variables only put extra effort on the algorithm development.
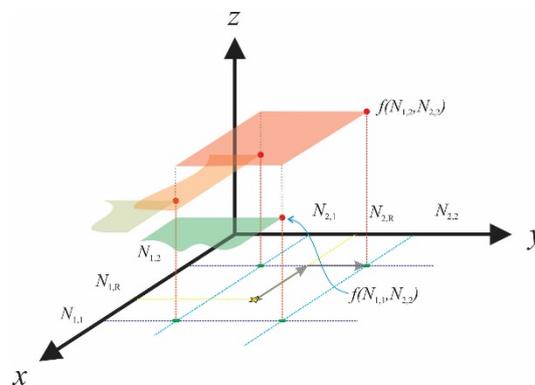

Figure 2. Typical rounding treatment of randomization approach for discrete variables

There are two ways to handle process constraints: 1) explicitly define them in the optimization algorithm and take them as free variables, and 2) implicitly define them in built-in «design specifications» and take them as dependent variables (fixed variables). Both cases were discussed in this work. Additionally, in randomized algorithms, the restriction is handled with an arbitrary value that punishes the function for the cases of the restrictions (e.g., purities) are not satisfied. Since the punishment value is arbitrary, an inadequate selection of it can make impossible the optimization algorithm to find optimal solutions.

2.2 Normalization

The bounds of the optimization variables differ from each other, and even for some variables, the used system of units can change the range between tens and thousands. In most of the works about global stochastic optimization, it not clear whether the variables are handled as their real values or not. Nevertheless, typical randomized algorithms use real values for the bounds of variables.

Considering these changes in the magnitude of variables, they also will add stiffness to the mathematical model. This work proposes a normalization of variables. The normalization will allow moving freely and fairly for the variables value ranges; also can avoid numerical effort because of the precision of the continuous variables.

A normalization between 1 and 0 can be obtained as described in Equation 1, where the UL and LL are the upper and lower limit of the variable (x), respectively.

$$\bar{x} = \frac{UL - x}{UL - LL} \qquad (1)$$

2.3 Discrete variables treatment

Taken two discrete variables, if we have a surface, as shown in Figure 3, the discrete point can move adequately on the surface in direction to the feasible point with the minimum value. However, the value of the real point cannot be calculated as an average of the four vertexes. This is because these vertexes can generate a convex surface instead.
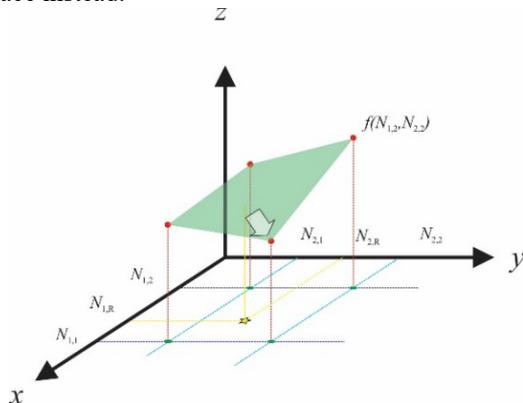

Figure 3. Hypothetical surface between feasible values

Figure 4 shows how the convex surface in Figure 3 can be generated (Case A and B), despite the "real mathematical surface" can be a little more complex as shown for example with a blue surface in the figure, which is not of interest because it contains infeasible solutions. Therefore, only we need to generate proper surfaces that emphasize the slop through the minimum value.
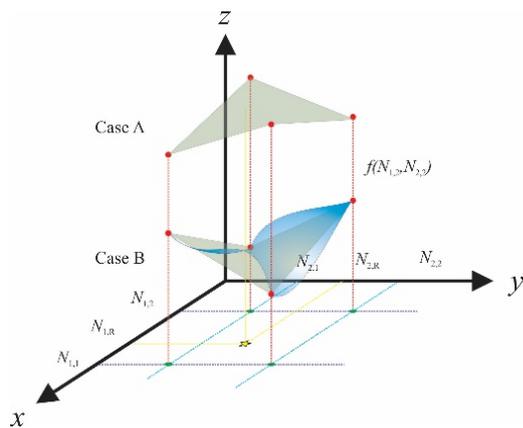

Figure 4. Hypothetical convexity of the surface of the feasible values

In order to warranty that a plain surface is generated we need to reduce the dimensional space to three vertexes. The combination of two discrete variables implies four hypertriangles. The selection of the three vertexes is based on the three nearest vertexes points (Figure 5). Then an evaluation of the three functions values it will be required for the algorithm instead of four.
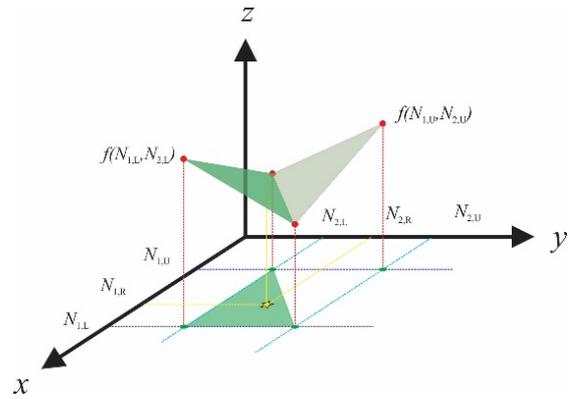

Figure 5. Reduction of the space to generate a flat surface

Figure 6 shows the evaluation of the distances from the real point to the feasible ones. The distances are calculated according to Equation 2. Then, the three vertexes can be defined by selecting the minimum distances of the plain surface and evaluating their functions.
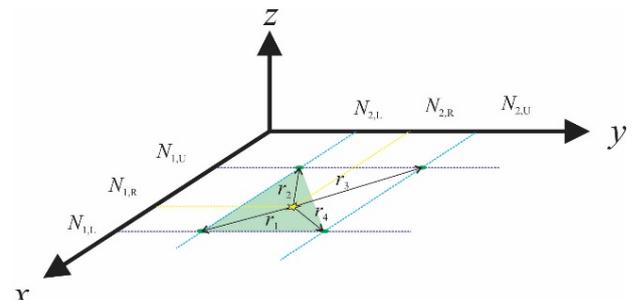

Figure 6. Selection of the vertex of the hyper-triangle trough the minimum distances

$$r_j^2 = \sum_{i=1}^{n} (N_i - N_R)^2 \qquad (2)$$

Where j represents each distance between the real node and the feasible nodes.

$$\vec{N} = \overrightarrow{AB} \times \overrightarrow{AC} \qquad (3)$$

$$\overrightarrow{AP_R} \cdot \vec{N} = 0 \qquad (4)$$

$$F_R = F_A + \frac{\left((A_x - x_R)N_x + (A_y - y_R)N_y\right)}{N_z} \qquad (5)$$

With the three vertexes evaluated it can be calculated the normal vector (**N**) that defines this surface (Equation 3), then the function value of the real point must satisfy the condition of the dot product of **AP** and **N** vectors equal to zero (Equation 4) if this point belongs to the same surface. Specifically for these two discrete variables, the function value of the real point is calculated as shown the Equation 5; where x_R and y_R are the real continuous values, A_x, A_y and F_A are the coordinates that define the minimum point and N_x, N_y, and N_z the components of the normal vector, the F_R is the function value for the real point. Figure 7 shows graphically the points mentioned before accordingly with the discrete variables.
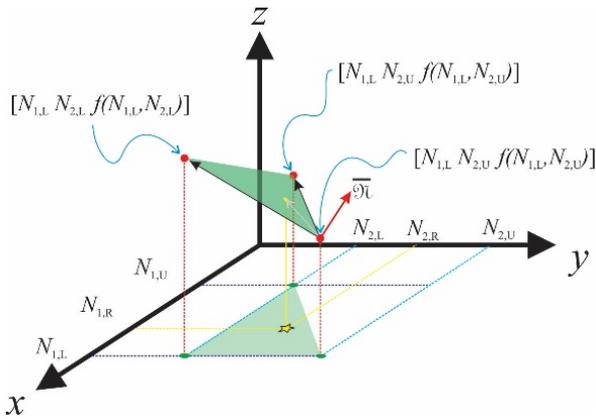
Figure 7. Feasible nodes (discrete values and function value) and the normal vector of the surface generate

2.4 Study case

For evaluating the proposed improvements to optimization algorithms, the taken case study is a conventional distillation that separates a binary mixture. This is due to the fact we know the optimal trivial solution of the separation processes; this is the maximum number of stages with the minimum reflux ratio gives the minimum heat duty consumption. Then we know the solution that must be obtained in order to compare the optimization algorithm with the traditional algorithms and with the proposed one. The optimizations are defined by Equation 6 and 7, where N_T is the total number stages, N_F is the feed stage, P is the pressure of the column and R the reflux ratio (see Figure 8). Where the purities of each component must be at least 99% mol.
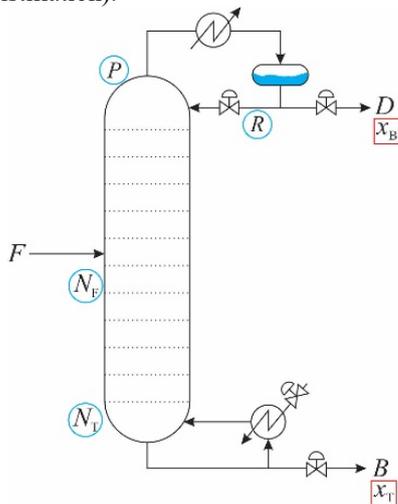
$$\min(Q_R) = f(N_T, N_F, P, R) \quad (6)$$
$$\min(TAC) = f(N_T, N_F, P, R) \quad (7)$$

Subject to

$$\vec{x} \geq 0.99 \quad (8)$$

As mentioned above, the solution for this case is already known because the total number of stages will be at the upper bound. However, this algorithm can be extended to the optimization of intensified processes (e.g., reactive distillation).

Figure 8. Feasible nodes (discrete values and function value) and the normal vector of the generated surface

The binary case is an equimolar benzene/toluene mixture, simulated in Aspen Plus®, and the optimization algorithm used is Simulated Annealing in Matlab® toolbox. The framework is through an Excel® macro. The Simulated Annealing parameters have kept as default and changed for a reduced version of the stall generation stop of 400. Noticed that the reflux ratio as a variable would disappear for the cases that simulation use design specification.

For the estimation of the total annual cost (TAC) is used the Guthrie methodology as reported by Richard Turton, *et al.* (2018) using CEPCI of 607.2 (July 2018). The TAC is calculated considering a feedback period of 5 years for the investment equipment and 8600 hours for the labor hours of the process

The optimizations were for the combinations considering the traditional framework implementation (without normalization, without design specifications and with upper rounding), and using a combination of the normalization, design specification and discrete values approach.

3 Results and Observations

Tables 1 to 4 shows most of the combinations of the use of design specification, normalization values, and discrete value approach. It can be noticed how the performance of the optimization improves with the combination of the options, and also exist and improvement of the efficacy to find the optimal solution. For these cases looks the design specification of the process has the bigger influence of finding the optimal solution, nevertheless the situation will change for the TAC cases. Any case, it is evident that using the design specification not only we reduce one design variable; also, the function value is not restrained to arbitrary punish values, and then avoiding to the optimization loses finding the reflux ratio proper to the restriction instead to searching for the other variables. Furthermore, it cannot be denied the combination of the three parameters improves the efficiency and efficacy of the optimal finding solution.

Table 1. Results for reboiler duty objective function

		Non-normalized
No design specific ation	Optimal variables values	[22.704, 11.937, 1, 1.631]
	Objective function value	1.136 MW
	Iteration number	2766
design specific ation	Optimal variables values	[29.22, 15.161,1]
	Objective function value	1.061 MW
	Iteration number	2766

Table 2. Results for reboiler duty objective function

		Normalized
No design specification	Optimal variables values	[0.2369 0.6609 1 0.9702] {23.6, 11.16, 1, 1.587}
	Objective function value	1.116 MW
	Iteration number	1865
Design specification	Optimal variables values	[0.016, 0.492, 1] {29.568 15.716, 1}
	Objective function value	1.061 MW
	Iteration number	1566

Table 3. Results for reboiler duty objective function

		Non-normalized and integer approach
No design specification	Optimal variables values	[24.907 15.002 1.024 1.758]
	Objective function value	1.121 MW
	Iteration number	1111
Design specification	Optimal variables values	[29.996, 15, 1]
	Objective function value	1.061 MW
	Iteration number	872

Table 4. Results for reboiler duty objective function

		Normalized and integer approach
No design specification	Optimal variables values	[0.172, 0.463 0.999, 0.964] {25.356, 14.005, 1.003 1.896}
	Objective function value	1.177 MW
	Iteration number	710
Design specification	Optimal variables values	[0, 0.481, 1] {30, 16.013, 1}
	Objective function value	1.061 MW
	Iteration number	546

In Tables 5 to 8 show the same combinations for the TAC evaluation as function objective. For these cases, it is more clear that using a design specification has an important role in finding the optimal solution. Nevertheless, it generates a bigger numerical effort. The implementation of the normalization helps to reduce this numerical effort, and at the same time improves the efficacy of the optimization slightly.

Table 5. Results for TAC objective function

		Non-normalized
No design specification	Optimal variables values	[17.19, 8.918, 1, 1.997]
	Objective function value	165.26 M$/year
	Iteration number	2386
Design specification	Optimal variables values	[29.418, 15.016, 1]
	Objective function value	141.79 M$/year
	Iteration number	5722

Table 6. Results for TAC objective function

		Normalized
No design specification	Optimal variables values	[0.5926, 0.6363, 1, 0.895] {13.999, 11.819, 1, 5.3395}
	Objective function value	146.51 M$/year
	Iteration number	1939
Design specification	Optimal variables values	[0.007, 0.484, 1] {29.811, 15.834, 1}
	Objective function value	141.79 M$/year
	Iteration number	2314

Table 7. Results for TAC objective function

		Non-normalized and integer approach
No design specification	Optimal variables values	[21.467, 9.974, 1, 1.776]
	Objective function value	151.47 M$/year
	Iteration number	1196
Design specification	Optimal variables values	[29.998 16 1]
	Objective function value	141.32 M$/year
	Iteration number	1128

Table 8. Results for TAC objective function

		Normalized and integer approach
No design specification	Optimal variables values	[0.151, 0.389, 0.997, 0.967] {25.923, 16.006, 1.009 1.7467}
	Objective function value	142.97 M$/year
	Iteration number	734
Design specification	Optimal variables values	[0.001, 0.481, 1] {29.973, 15.999, 1}
	Objective function value	141.32 M$/year
	Iteration number	685

It can be considered the most important part of the optimization is the design specification, due to all the options with design specification aim to reach the optimal solution. This is not a trivial coincidence because the algorithm is released of the arbitrary punishment. A good contribution of the normalization avoids many Aspen running simulations because the introduced values can be restricted dynamically then, that avoid taking an unfeasible design that will be rejected for the simulator and eventually took as an inappropriate area to optimize. Finally, the discrete values approach, reduce enormously the numerical effort by generating a softer surface reducing the stiffness of the system.

4 Conclusions

The proposed improvements for the optimization of the design processes has shown an evident improvement of the efficiency and efficacy of the optimization; this is the reduction of the number of iterations (until 80% fewer iterations) and the localization of the optimal design getting closer than the feasible discrete values. To guarantee the effectivity of this approach the three recommendations: (1) normalization (2) design specification, and (3) approach to discrete value points treatment, must be used at the same time for the sake of better performance.

References

Cabrera-Ruiz, J., E. Y. Miranda-Galindo, J. G. Segovia-Hernandez, S. Hernández, and A. Bonilla-Petriciolet; "Evaluation of Stochastic Global Optimization Methods in the Design of Complex Distillation Configurations." pp. 441–456, Stoch. Optim. - Seeing Optim. Uncertain, Ed. A Review of Recent Advances in Global Optimization InTech Ed., Croatia (2012)

Cavazzuti, M.; Optimization Methods: From Theory to Design, Springer, Berlin, Heidelberg, Germany (2013)

Floudas, C. A. and C. E. Gounaris; "A Review of Recent Advances in Global Optimization," *J. Glob. Optim*., **45**, 3 (2009)

Gitizadeh, M., A. A. Vahed, and J. Aghaei; "Multistage Distribution System Expansion Planning Considering Distributed Generation Using Hybrid Evolutionary Algorithms," *Appl. Energy,* **101**, 655–666 (2013)

Goryachih, A. S. and M. A. Rachinskaya; "Multidimensional Global Optimization Method Using Numerically Calculated Derivatives," *Procedia Comput. Sci.,* **119**, 90–96 (2017)

Grossmann, I. E. and A. W. Westerberg; "Research Challenges in Process Systems Engineering." *AIChE J*. **46**, 1700–1703 (2000)

Munawar, S. A. and R. D. Gudi; "A Nonlinear Transformation Based Hybrid Evolutionary Method for MINLP Solution," *Chem. Eng. Res. Des*., **83**, 1218–1236 (2005)

Popović, Ž. N., V. Dj. Kerleta, and D. S. Popović; "Hybrid Simulated Annealing and Mixed Integer Linear Programming Algorithm for Optimal Planning of Radial Distribution Networks with Distributed Generation," *Electr. Power Syst. Res*., **108**, 211–222 (2014)

Renders, J. and S. P. Flasse; "Hybrid Methods Using Genetic Algorithms for Global Optimization." *IEEE Trans. Syst. Man*, Cybern. Part B **26**, 243–258 (1996)

Ricardez-Sandoval, L. A., H. M. Budman, and P. L. Douglas; "Simultaneous Design and Control: A New Approach and Comparisons with Existing Methodologies." *Ind. Eng. Chem. Res*., **49**, 2822–2833 (2004)

Santaella, M., G. Rodríguez, and A. Orjuela; "Design of a Thermally Coupled Reactive Distillation Sequence for Triethyl Citrate Production," 14th Annual Meeting AIChE, Atlanta, USA (2014)

Sayyaadi, H., A. Saffari, and A. Mahmoodian; "Various Approaches in Optimization of Multi Effects Distillation Desalination Systems Using a Hybrid Meta-Heuristic Optimization Tool." *Desalination,* **254,** 138–148 (2010)

Segovia-Hernández, J. G., S. Hernández, and A. Bonilla Petriciolet; "Reactive Distillation: A Review of Optimal Design Using Deterministic and Stochastic Techniques," *Chem. Eng. Process. Process Intensif.*, **97**, 134–143 (2015)

Turton, R., J. A. Shaeiwitz, D. Bhattacharyya, and W. B. Whiting; Analysis, Synthesis, and Design of Chemical Processes, 5th ed., Prentice-Hall International Series in Engineering, Prentice Hall, Upper Saddle River, USA (2018)

Vazquez-Castillo, J. A., J. A. Venegas-Sáchez, J. G. Segovia-Hernández, H. Hernández-Escoto, S. Hernández, C. Gutiérrez-Antonio, and A. Briones-Ramírez; "Design and Optimization, Using Genetic Algorithms of Intensified Distillation Systems for a Class of Quaternary Mixtures." *Comput. Chem. Eng.,* **33**, 1841–1850 (2009)

Yuan, Z., B. Chen, G. Sin, and R. Gani; "State-of-the-Art and Progress in the Optimization-Based Simultaneous Design and Control for Chemical Processes." *AIChE J.,* **58**, 1640–59 (2012)