Simultaneous synthesis and design of reaction–separation–recycle processes using rigorous models
Ma, Yingjie  ; Yang, Zekun  ; El-Khoruy, Aline  ; Zhang, Nan; Li, Jie; Zhang, Bingjian  ; Sun, Li

Published in:
Industrial & Engineering Chemistry Research

DOI:
10.1021/acs.iecr.1c00250

Published: 19/05/2021

Document Version
Publisher's PDF, also known as Version of record

Link to publication on the UWS Academic Portal

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the UWS Academic Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact pure@uws.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Simultaneous Synthesis and Design of Reaction–Separation–Recycle Processes Using Rigorous Models

Yingjie Ma, Zekun Yang, Aline El-Khoruy, Nan Zhang, Jie Li,* Bingjian Zhang, and Li Sun

ABSTRACT: Simultaneous synthesis and design of reaction–separation–recycle processes using rigorous models is highly desirable to improve process efficiency. However, it often leads to a large-scale highly challenging optimization problem. In this work, we propose a computationally efficient optimization framework for the challenging problem. The reactor and separator networks are modeled using the generalized disjunctive programming, which are reformulated into a highly nonconvex mixed-integer nonlinear programming (MINLP) formulation using the convex-hull method. To solve the complex MINLP model, a systematic solution approach is proposed in which an initialization strategy is first proposed to generate a feasible solution for a partially relaxed synthesis problem using the hybrid steady-state and time-relaxation optimization algorithm. A successive relaxed MINLP solution strategy is then adopted to solve the original model to local optimality. The computational results demonstrate that the proposed framework obtains better solutions with less computational effort, by ~1 order of magnitude, than the existing algorithms.

1. INTRODUCTION

Developing sustainable chemical processes is a key issue in the current world.1 Since chemical processes are usually composed of reactors, separators, and recycle streams to recover raw materials, we often need to determine the types of reactors and separators, their connection structure, and optimal operating conditions when designing a chemical process. Since there are many degrees of freedoms, a systematic and efficient methodology must be developed to obtain the optimal process.2 To reduce the combinatorial complexity, such design problem is often decomposed into reactor network synthesis and separation network synthesis, which are designed separately.3

The reactor network synthesis and design problems often involve the selection of reactor types, their connection configurations, optimal operating conditions, and equipment sizes. Several methods including the heuristic method, attainable region method and superstructure-based method have been proposed. In the heuristic method, the reactor type is first selected based on existing experience and the operating parameters and reactor sizes are then determined,4 which is difficult to be used for complex reaction systems.5 A more general method, the concept of attainable region, was first proposed by Horn6 referring to the set of points in concentration space that can be obtained through reaction and mixing. This method was further improved by Glasser et al.7 and Hildebrandt et al.8 as a geometry method. Although it is considered as a rigorous and elegant method, it is difficult to apply to solve problems with more than three dimensions.9 To conquer the limitations on the dimensions of the geometry method, the concept of attainable region was also extended to become an optimization-based method by several researchers.10−12 The superstructure-based method constructs the superstructure to represent all possible splitting and mixing between predefined reactors.13,14

Most researches on the separation network synthesis and design focus on distillation network synthesis and design, since 90% of all separations in the chemical industry are performed by distillation15 and distillation continues to be the most important separation technique in chemical processes.16 The distillation synthesis and design problem determines the optimal distillation sequence, size of each distillation column, and optimal operating conditions. There are two main methods that have been proposed, including the graphical method and the superstructure-based method. The graphical method, like the residual curve method (RCM), can be easily
used for synthesis and analysis of special distillation systems, such as azeotropic distillation,\textsuperscript{17} extractive distillation,\textsuperscript{18} and reactive distillation.\textsuperscript{19} However, they are usually limited to the number of components and cannot directly provide optimal design parameters. The superstructure-based method can be applied to distillation systems with an arbitrary number of components\textsuperscript{20,21} and generate optimal distillation sequence and operating parameters simultaneously.\textsuperscript{22} The superstructure is usually modeled using the generalized disjunctive programming (GDP) with rigorous distillation models.\textsuperscript{23}

Synthesizing reactor network and separation network individually usually cannot get the really optimal flowsheet, because of the strong interaction between them.\textsuperscript{24} To synthesize and design the reaction–separation–recycle process, Douglas\textsuperscript{3} decomposed the synthesis problem into different decision levels and used heuristics to determine the configurations and parameters in various levels individually, which, however, is hard to manage the interactions between different levels. The superstructure-based optimization method was also used to determine configuration alternatives and operating parameters of the reaction–separation–recycle process systematically by Kocis and Grossmann.\textsuperscript{25} They employed short-cut models to decrease the complexity, which may lead to inaccuracy or even infeasibility when the optimal solution was validated using rigorous models. Ye et al.\textsuperscript{26} also synthesized the reaction-separation-recycle process using short-cut models, which involved complex distillation configuration in the superstructure. Recker et al.\textsuperscript{17} first used short-cut models to select several good flowsheet variants, and then conducted optimization to the candidates using rigorous models, which may miss the really optimal flowsheet at the initially screening phase.

To obtain the really optimal design of the reaction–separation–recycle process, incorporating rigorous models at the process synthesis and design level is highly necessary. Gross and Roosen\textsuperscript{28} synthesized the hydrodealkylation process using evolutionary algorithms based on rigorous Aspen Plus simulation.\textsuperscript{29} However, their algorithm requires a long computational time and is often difficult to find an optimal solution when there is a large increase in the number of decision variables.\textsuperscript{30} In addition, the solution optimality from the evolutionary algorithm cannot be guaranteed. Henao and Maravelias proposed a surrogate-based superstructure optimization framework for process synthesis, where the surrogate models are obtained through fitting simulation data from rigorous models and then used for optimization, reducing the complexity and nonlinearity.\textsuperscript{31} However, the accuracy of the surrogate-based optimization approach is hard to be guaranteed.\textsuperscript{27} Recently, a rigorous model-based process synthesis framework was proposed by Zhang et al.,\textsuperscript{27} where the reactor network and distillation system were synthesized simultaneously using rigorous models. The superstructure for the reactor network and distillation system was modeled using GDP, which was then converted to a nonconvex MINLP problem using the big-M method. This framework using rigorous models can find a locally optimal solution and the generated design has higher accuracy than the surrogate-based optimization approach. However, a large number of discrete variables were introduced in their model, resulting in low computational efficiency. The performance of the branch and

Figure 1. A superstructure for the reaction–separation–recycle system.
bound optimizer they used is significantly dependent on the initial point, which may lead to infeasibility with some initial points. No systematic strategy is provided for generation of a good initial point. In addition, they assumed fixed purities in the recycle stream without considering the effect of varying purity in the recycle stream, which could miss a much better solution.\(^{32}\)

In this work, we develop a computationally efficient model-based optimization framework for the optimal synthesis and design of the reaction—separation—recycle process using rigorous models. The superstructure of the reaction—separation—recycle process is modeled using the GDP method\(^{23}\) and then reformulated into a MINLP optimization problem using the convex hull method.\(^{33}\) To solve the resulting highly nonconvex MINLP model, a systematic solution approach is proposed in which an initialization strategy is first used to generate a feasible solution for a partially relaxed synthesis problem of the reaction—separation—recycle process through the hybrid steady-state and time-optimization algorithm\(^{34}\) and the GRG solution algorithm.\(^{35}\) This feasible solution is then used to initialize the optimization of the reaction—separation—recycle system. A successive relaxed MINLP solution strategy is adopted to solve the original model to local optimality. The computational results demonstrate that we can not only solve all the case studies in Zhang et al.\(^{25}\) with better optimal solutions and significantly less computational effort by \(\sim\)1 order of magnitude, but also optimize a much more complex HDA process within 2 min. In addition, optimizing the purity of the recycle stream simultaneously with the whole reaction—separation—recycle process can result in a cost savings of \(\sim\)13%.

The paper is organized as follows: The next section defines the problem. In section 3, the mathematical formulation is proposed. Following that, the tailored solution approach is described. In section 5, three process synthesis and design problems are solved to show the efficiency and good convergence of the proposed optimization framework. Finally, section 6 concludes this work.

### 2. PROBLEM STATEMENT

Figure 1 illustrates a superstructure for the reaction—separation—recycle system, which is composed of a reactor network and a separation network for ternary separation. There are total \(K\) \((k = 1, 2, 3, \ldots, K)\) stages in the reactor network proposed by Lakshmann and Biegler.\(^9\) At each stage, there are \(R\) \((r = 1, 2, 3, \ldots, R)\) types of reactors that can be used, such as isothermal or adiabatic plug flow reactor (PFR), and continuous-stirred tank reactor (CSTR). At most, one reactor type can be selected at each stage. The feed stream is allowed to bypass reactors at each stage. The outlet stream at a stage \(k\) is split and fed into downstream reactors. There are \(I\) \((i = 1, 2, 3, \ldots, I)\) components in each stream. Distillation is assumed to be the only separation technology considered in the separation network, since it is widely used for separation. The separation network is represented using the SEN representation in which a separation sequence is defined with a series of separation tasks.\(^{23}\) There are \(J\) \((j = 1, 2, 3, \ldots, J)\) distillation columns \((J = I - 1)\) with \(D\) \((d = 1, 2, 3, \ldots, D)\) possible separation tasks \((D = (P - I)/6)\). Each distillation column has \(N\) \((n = 1, 2, 3, \ldots, N)\) trays.

The problem can be stated as follows: Given:

1. Raw materials, product and production requirements, e.g., production rates and product purities.
2. \(K\) stages in the reactor network, suitable reactor types, reaction kinetics, and related thermodynamic data.
3. \(I\) components and their thermodynamic properties.
4. \(J\) distillation columns, and maximum number of trays in each column.

Determine:

1. Optimal flowsheet structure and reactor types.
2. Optimal operating conditions including stream flow rates, pressure, and temperature.
3. Optimal sizes of reactors, distillation columns, and heat exchangers.

Assumptions:

1. Only CSTR and PFR are considered in the reactor network because other reactors can always be approximated by a cascade of PFRs and CSTRs.\(^{37}\)
2. Constant liquid density is also assumed in the PFR.
3. Only the simple distillation column with one feed and two outlet streams is considered as this type of distillation column is widely used in chemical plants.
4. Ideal vapor and negligible enthalpy of mixing are assumed because the mixtures are close to being a ideal mixture.

The objective is to minimize total annualized cost \((TAC)\) or maximize profit.

### 3. MODELING OF THE REACTOR NETWORK

We use stages 0 and \((K + 1)\) to denote raw materials and final products, respectively, in the reactor network of Figure 1. Therefore, there are \((K + 2)\) \((k = 0, 1, 2, \ldots, K + 1)\) stages in total in the reactor network. At stages 0 and \((K + 1)\), there is no reactor. We use a set \(K^R = \{k | 1 \leq k \leq K\}\) to denote stages involving reactors. We define variables \(\eta^R_{i,k}\) to denote the feed flow rate of a component \(i\) to stage \(k\) and \(\eta^R_{o,k}\) to denote the outlet flow rate of a component \(i\) from stage \(k\). Since a stream from a stage can be fed into all downstream stages, \(\eta^R_{i,k,k'}\) is introduced to denote the flow rate of a component \(i\) from stage \(k\) to stage \(k'\). The feed rate of a component \(i\) to stage \(k\) is given by

\[
\eta^R_{i,k} = \sum_{k'=0}^{k-1} \eta^R_{i,k',k} \quad \forall \ k > 0, i
\]

The outlet flow rate of a component \(i\) from a stage \(k\) is calculated by

\[
\eta^R_{o,k} = \sum_{k'=k+1}^{K+1} \eta^R_{i,k,k'} \quad \forall \ k < K + 1, i
\]

We define \(\eta^R_{i,k,k'}\) to denote the total flow rate of a stream from stage \(k\) to stage \(k'\). Then,

\[
\eta^R_{i,k,k'} = \sum_i \eta^R_{i,k,k'} \quad \forall \ k < K + 1, k' > k
\]

Each outlet stream at a stage \(k\) is split into different streams, which are fed into downstream stages. We define \(\eta^R_{i,k}\) to denote the split ratio from stage \(k\) to stage \(k'\),

\[
\eta^R_{i,k,k'} = \eta^R_{i,k,k'} \eta^R_{o,k'} \quad \forall \ i, k < K + 1, k' > k
\]
The enthalpy balance at the inlet point of a stage k is given below:

$$\sum_{i} \left( F_{i,k}^{in} h_{i,k}^{in} \right) = \sum_{k=0}^{k-1} \sum_{i} \left( F_{i,k}^{out} h_{i,k}^{out} \right) \quad \forall \ k > 0$$  \hspace{1cm} (5)

where $h_{i,k}^{in}$ and $h_{i,k}^{out}$ are the specific molar enthalpies of pure component i in the streams feeding to stage k and leaving stage k, respectively, since the mixing enthalpy is assumed to be negligible. The enthalpy of a component is the function of temperature as follows:

$$h_{i,k}^{in} = H_{i,k}^{in} \quad \forall \ i, k$$  \hspace{1cm} (6)

$$h_{i,k}^{out} = H_{i,k}^{out} \quad \forall \ i, k$$  \hspace{1cm} (7)

where $H_{i, k}^{in}$ is a function to calculate the enthalpy of a component i. $T_{k}^{in}$ and $T_{k}^{out}$ are inlet and outlet temperatures of stage k, respectively.

At each stage different types of reactors are provided for option. However, at most, one reactor type can be selected. If a reactor type is selected, its related constraints are activated. Otherwise, all related variables including feed flow rates, size, and cost should be zero. Such logic conditions can be generally represented by the GDP:

$$\left[ \begin{array}{c} Y_{r,k} \\ F_{i,r,k}^{out} = R_{r}^{f} \left( F_{i,r,k}^{in}, T_{r,k}^{in}, \tau_{r,k} \right) \\ T_{r,k}^{out} = R_{r}^{f} \left( F_{i,r,k}^{in}, T_{r,k}^{in}, Q_{r,k}, \tau_{r,k} \right) \\ Cap_{r,k} = C_{r}^{f} \left( \tau_{r,k} \right) \end{array} \right] \quad \forall \ r, k \in \mathbb{K}^{R}$$  \hspace{1cm} (GDP0)

where $Y_{r,k}$ is a Boolean variable denoting whether a reactor type $r$ is selected at stage $k$. $F_{i,r,k}^{in}$ and $F_{i,r,k}^{out}$ are the respective inlet and outlet flow rates of a component $i$ in a reactor $r$ at stage $k$. $T_{r,k}^{in}$ and $T_{r,k}^{out}$ are the respective inlet and outlet temperatures of a component $i$ in a reactor $r$ at stage $k$. $Q_{r,k}$, $\tau_{r,k}$, and $Cap_{r,k}$ denote the heat duty, residence time, and capital cost of reactor type $r$ at stage $k$, respectively. $R_{r}^{f}()$ represents the performance model of reactor $r$ (PFR, CSTR), which is provided in the Supporting Information. $C_{r}^{f}()$ is the capital cost correlation of reactor type $r$.

The equation is then reformulated into an MINLP problem using the convex hull method as the convex hull method usually generates tighter formulations than the big-M method. We define a binary variable $y_{r,k}$ to represent if a reactor type $r$ is selected at stage $k$. We use a dummy reactor ($r = 0$) to denote the bypassing at each stage. Therefore, there are total $(R + 1)$ ($r = 0, 1, 2, ..., R$) reactor types at each stage. At each stage, exactly one reactor type can be selected.

$$\sum_{i} y_{r,k} = 1 \quad \forall \ k \in \mathbb{K}^{R}$$  \hspace{1cm} (8)

If a reactor $r$ is not selected, its inlet and outlet flow rates, inlet and outlet temperatures, heat duty, residence time, and capital cost should be zero. Otherwise, the related variables can vary within some bounds. These conditions are described by eqs 9–15.

$$F_{i,r}^{min} \leq F_{i,r,k}^{in} \leq F_{i,r}^{max} \quad \forall \ i, r, k \in \mathbb{K}^{R}$$  \hspace{1cm} (9)

$$T_{r,k}^{min} \leq T_{r,k} \leq T_{r,k}^{max} \quad \forall \ r, k \in \mathbb{K}^{R}$$  \hspace{1cm} (10)

$$Q_{r,k}^{min} \leq Q_{r,k} \leq Q_{r,k}^{max} \quad \forall \ r, k \in \mathbb{K}^{R}$$  \hspace{1cm} (11)

$$\tau_{r,k}^{min} \leq \tau_{r,k} \leq \tau_{r,k}^{max} \quad \forall \ r, k \in \mathbb{K}^{R}$$  \hspace{1cm} (12)

$$Cap_{r,k}^{min} \leq Cap_{r,k} \leq Cap_{r,k}^{max} \quad \forall \ r, k \in \mathbb{K}^{R}$$  \hspace{1cm} (13)

If a specific reactor type is given, the performance model can be explicitly presented. In this work, only CSTR and PFR are used, whose model is provided in the Supporting Information.
4. MODELING OF DISTILLATION COLUMN NETWORK

The distillation column network consists of one or more distillation columns that are used for separation of multiple components. Modeling of such network includes modeling of the distillation sequence and the distillation column.

4.1. Modeling of the Distillation Sequence. The superstructure for the distillation sequence is represented using the state-equipment network (SEN), because of the least number of equipment required. The superstructure can be constructed based on either the order of separation tasks or the order of key components. For the ternary separation, we use the superstructure based on the order of separation tasks, while for the separation of more components, we use the superstructure based on the order of key components, because it requires a smaller number of binary variables, as indicated in the Table S13 in the Supporting Information. The superstructures for ternary and quaternary separations are provided in Figure 2.

The SEN for the distillation sequence can be modeled using the GDP given in (eq GDP1).

\[
\begin{align*}
Y_{d,j} & \\
xd_{d,j} & \geq \xi_{d} \quad \forall \ i, \ j \\
zd_{d,j} & \geq \phi_{d} \quad \forall \ i, \ j
\end{align*}
\]  

(GDP1)

where \(D_i\) is the set denoting all possible tasks that a column \(j\) can possess. \(Y_{d,j}\) is a Boolean variable to denote if a column \(j\) performs a task \(d\). \(xd_{d,j}\) and \(zd_{d,j}\) are variables denoting the purity and recovery of component \(i\) in the distillate of a column \(j\) with the separation task \(d\), respectively, which should be greater than the desired purity \(\xi_{d} \) and recovery \(\phi_{d}\) in the distillate stream, respectively. On the other hand, \(xb_{d,j}\) and \(zb_{d,j}\) are variables denoting the purity and recovery of component \(i\) in the bottom of a column \(j\) with the separation task \(d\), respectively, which should be greater than the desired purity \(\xi_{d} \) and the desired recovery \(\phi_{d}\) in the bottom stream, respectively.

We usually only have required purity for the final product stream and do not need to specify the minimum purity for the middle streams. In addition, we only need to specify the recovery of key components for each separation task. The purity and recovery are usually fixed values for each final product, so we only need to use \(\xi_{d} \) and \(\phi_{d}\) to denote purity requirement and recovery requirement for each component \(i\), respectively. With these, GDP1 can be simplified to GDP2 as follows,

\[
\begin{align*}
Y_{d,j} & \\
x_{d,j} & \geq \xi_{d} \quad \forall \ i, \ j \\
zd_{d,j} & \geq \phi_{d} \quad \forall \ i, \ j
\end{align*}
\]  

(GDP2)

where \(DD_{p}\) is the set of separation tasks that can obtain pure component \(i\) in the distillate stream of a column, while \(DB_{p}\) is the set of separation tasks in \(D\) that can obtain pure component \(i\) in the bottom stream of a column. For example, \(DD_{p,a} = \{A/B, A/BC, A/BCD\}\) and \(DB_{p,a} = \{B/C, AB/C\}\) for quaternary separation in Figure 2b. \(DD_{k}\) is the set of separation tasks in \(D\) that take component \(i\) as a light key component, while \(DB_{k}\) is the set of separation tasks in \(D\) that take component \(i\) as a heavy key component. For example, \(DD_{k,a} = \{A/B, A/BC, A/BCD\}\) and \(DB_{k,a} = \{B/C, AB/C, B/CD, AB/CD\}\) for quaternary separation in Figure 2b.

Equation GDP2 can be converted to the following MINLP model.

\[
\sum_{d \in D_j} y_{d,j} = 1 \quad \forall \ j
\]  

(26)

\[
xd_{d,j} \geq \xi_{d} \quad \forall \ i, \ j, \ d \in D_j \cap DD_{p}
\]  

(27)

\[
zd_{d,j} \geq \phi_{d} \quad \forall \ i, \ j, \ d \in D_j \cap DD_{k}
\]  

(28)

\[
xb_{d,j} \geq \xi_{d} \quad \forall \ i, \ j, \ d \in D_j \cap DB_{p}
\]  

(29)

\[
xb_{d,j} \geq \phi_{d} \quad \forall \ i, \ j, \ d \in D_j \cap DB_{k}
\]  

(30)
There are some logical relations among the separation tasks, which can be used to exclude some infeasible solutions. For convenience, \( d(i_o, i_s, i_j) \) is used to represent a separation task in column \( j \), where \( i_o, i_s \) and \( i_j \) denote different components in the mixture. The subscript \( s, k \) and \( e \) are integers satisfying \( 1 \leq s \leq k \leq e \leq I \), and \( i_s \) represent the first and last components in the mixture, respectively. The larger subscript for \( i \) means the heavier component. With these symbols, \( d(i_o, i_s, i_j) \) can be interpreted as a task separating a mixture of \( i_o, i_s, \ldots, i_j \) with \( i_s/i_{j+1} \) as key components in column \( j \). One logical relation is that if a separation task \( d(i_o, i_s, i_j) \) is selected, then, in columns \( j \) and \( j+1 \) (\( j \neq j+1 \)), there must be a separation task \( d(i_o, i_s, i_j) \) and a separation task \( d(i_o, i_s, i_j) \), respectively, where \( s \leq k_1 < k_2 < k \) and \( k_1 + 1 \leq k_2 < k_3 \) which can be represented by eqs 31 and 32.

\[
\sum_{i_o=1}^{k-1} \sum_{i_s=1}^{k} \sum_{i_j=s}^{k-1} y_{(i_o,i_s,i_j)} \geq y_{d(i_o,i_s,i_j)} \\
\forall \ j, d(i_o, i_s, i_j) \in D_j
\] (31)

\[
\sum_{i_o=1}^{k-1} \sum_{i_s=1}^{k} \sum_{i_j=s-1}^{k-1} y_{(i_o,i_s,i_j)} \geq y_{d(i_o,i_s,i_j)} \\
\forall \ j, d(i_o, i_s, i_j) \in D_j
\] (32)

The second logical relationship is that if a separation task \( d(i_o, i_s, i_j) \) is selected, then in column \( j \neq j+1 \), there must be a separation task \( d(i_o, i_s, i_j) \), where \( 1 \leq k_3 \leq e-1 \) or a separation task \( d(i_o, i_s, i_j) \), where \( e+1 \leq k_4 \leq I \). This can be represented by eq 33:

\[
\sum_{i_o=1}^{k-1} \sum_{i_s=1}^{k} \sum_{i_j=s-1}^{k-1} y_{(i_o,i_s,i_j)} + \sum_{i_o=1}^{k-1} \sum_{i_s=1}^{k} \sum_{i_j=e+1}^{k} y_{(i_o,i_s,i_j)} \geq y_{d(i_o,i_s,i_j)} \\
\forall \ j, d(i_o, i_s, i_j) \in D_j
\] (33)

An example of applying the model of the distillation sequence to a quaternary separation is shown in the Supporting Information.

4.2. Modeling of the Distillation Column. The rigorous tray-by-tray model (i.e., the MESH model) is usually employed for modeling a distillation column. Many existing efforts have developed MINLP formulations for the distillation column in which binary variables are used to denote if a tray exists or not. \(^{37,38}\) When a large number of trays are required to meet the desired product purity and recovery, the MINLP model involves a large number of binary variables. In addition, there are many nonconvex nonlinear terms in the rigorous model that mainly come from phase equilibrium equations, energy balance equations, and thermodynamic property equations. Directly solving such a large-scale nonconvex MINLP model is fairly computationally challenging. It is even more computationally intensive to incorporate such model into a more complicated reaction—separation—recycle process synthesis problem. To improve the computational efficiency, the bypass efficiency method \(^{39}\) is used to model the number of stages. Let \( N (n = 1, 2, \ldots, N) \) denotes stages in a column, \( N \subset N \) denotes the feed stages, and \( N \subset N \) denotes all conditional stages, which do not include the reboiler, condenser, and feed stages.

The bypass for a conditional stage \( n (n \in N) \) is illustrated in Figure 3.

When a vapor stream from stage \( (n-1) \) with component flow rate \( V_{n-1,i} \) and specific enthalpy \( hV_{n-1,i} \) feeds stage \( n \), a part of the vapor stream goes through stage \( n \) and reaches phase equilibrium with the liquid stream leaving this stage, while the remaining bypasses the stage. While the component flow rate and specific enthalpy of the vapor stream reaching phase equilibrium with the liquid stream at this stage \( n \) are represented by \( eV_{n,i} \) and \( eV_{n,i} \), respectively, they are calculated by \( V_{n-1,i}(1-e) \cdot hV_{n-1,i} \), and \( (1-e) \cdot V_{n-1,i} \) respectively, for the bypass vapor stream. The final component flow rate \( V_{n,i} \) and specific enthalpy \( hV_{n,i} \) of the vapor stream leaving stage \( n \) are represented by eqs 34 and 35.

\[
V_{n,i} = (1-e) \cdot V_{n-1,i} + eV_{n,i} \quad \forall n \in NM, i
\] (34)

\[
hV_{n,i} = (1-e) \cdot hV_{n-1,i} + eV_{n,i} \quad \forall n \in NM, i
\] (35)

Similarly, the component flow rate \( L_{n,i} \) and specific enthalpy \( hL_{n,i} \) in the liquid stream leaving stage \( n \) are calculated by eqs 36 and 37.

\[
L_{n,i} = (1-e) \cdot L_{n-1,i} + eL_{n,i} \quad \forall n \in NM, i
\] (36)

\[
hL_{n,i} = (1-e) \cdot hL_{n-1,i} + eV_{n,i} \quad \forall n \in NM, i
\] (37)

where \( L_{n+1,i} \) and \( hL_{n+1,i} \) are the component flow rate and specific enthalpy of the liquid stream from stage \( n + 1 \), respectively. \( L_{n,i} \) and \( hL_{n,i} \) are the component flow rate and specific enthalpy of the liquid stream at phase equilibrium on stage \( n \).

Although \( eV_{n,i} \) are defined as continuous variables between 0 and 1, they have a tendency to be binary in the globally optimal solution, because partial bypass (i.e., \( eV_{n,i} \) takes values between 0 and 1) indicates thermodynamically inefficient mixing. \(^{39}\) This is also computationally proved by other research. \(^{40,41}\) However, in the case when the model is solved to local optimality or where adding or removing one stage in a column has rather small effect on the objective function in some processes, some \( eV_{n,i} \) may take values between 0 and 1. \(^{12}\) In such cases, the complementary constraints [i.e., eq 38] are added to unconditionally guarantee \( eV_{n,i} \) to be 0 or 1 only. As these constraints are often difficult to solve to optimality, they are used only when the bypass efficiencies are fractional in this work.

\[
eV_{n,i} \leq 0 \quad \forall n \in NM
\] (38)

For better numerical performance of the bypass efficiency method, we first calculate the stream variables in phase equilibrium \( V_{n,i} \), \( hV_{n,i} \), \( L_{n,i} \), \( hL_{n,i} \) through assuming that all the inlet streams flow into the stage \( n \) and reach to phase
equilibrium, then the bypass efficiency $\varepsilon_s$ is multiplied with the variables to embody the influence of bypass. This can avoid zero flow rates when the bypass efficiencies are zero. Based on the above description, $V_{n-1,i}^\text{L}$, $hV_{n-1,i}^\text{L}$, $L_{n,i}^\text{eq}$, $hL_{n,i}^\text{eq}$ are actually obtained from the normal MESH equations, as shown in eqs 39–43.

$$V_{n-1,i} + L_{n+1,i} - V_{n-1,i}^\text{L} - L_{n,i}^\text{eq} = 0 \quad \forall \ n \in \text{NM}, \ i$$

$$f_{n,i}^\text{L} = j_{n,i}^\text{V} \quad \forall \ n \in \text{NM}, \ i$$

$$\sum_i V_{n,i}^\text{eq} = \text{VAP}_n^\text{eq} \quad \forall \ n \in \text{NM}$$

$$\sum_i L_{n,i}^\text{eq} = \text{LIQ}_n^\text{eq} \quad \forall \ n \in \text{NM}$$

$$\sum_i (V_{n-1,i}hV_{n-1,i} + L_{n+1,i}hL_{n+1,i} - V_{n,i}^\text{L}hV_{n,i}^\text{L} - L_{n,i}^\text{eq}hL_{n,i}^\text{eq}) = 0$$

$$\forall \ n \in \text{NM}$$

where $\text{LIQ}_n^\text{eq}$ and $\text{VAP}_n^\text{eq}$ are the total flow rates of liquid and vapor streams on stage $n$, respectively, which can be calculated from eqs 44 and 45 when the vapor is considered to be an ideal gas.

$$f_{n,i}^\text{L} = p_{n,i}^\text{S} \gamma_{n,i}^\text{S} S_{n,i} \quad \forall \ n \in \text{NM}, \ i$$

$$f_{n,i}^\text{V} = y_{n,i}^\text{S} P_{n}^\text{S} \quad \forall \ n \in \text{NM}, \ i$$

where $p_{n,i}^\text{S}$, $\gamma_{n,i}^\text{S}$ and $S_{n,i}$ are the saturated pressure, liquid activity coefficient, and liquid molar composition of component $i$ on stage $n$, respectively. $y_{n,i}$ is the vapor molar composition of component $i$ on stage $n$ with pressure $P_{n}$, $P_{n}^\text{S}$ and $\gamma_{n,i}^\text{S}$ can be calculated from the Antoine equation and various activity coefficient equations. $x_{n,i}$ and $y_{n,i}$ are obtained from eqs 46 and 47:

$$L_{n,i}^\text{eq} = \text{LIQ}_n^\text{eq} x_{n,i} \quad \forall \ n \in \text{NM}, \ i$$

$$V_{n,i}^\text{eq} = \text{VAP}_n^\text{eq} y_{n,i} \quad \forall \ n \in \text{NM}, \ i$$

The entire optimization problem (denoted as $P_0$) for the reaction–separation-recycle system can be summarized below.

$$\min \ obj$$

$$\text{s.t.} \quad \text{eqs } 1–25, 26–33, 34–37, \text{ and } 39–47$$

$$x^L \leq x \leq x^U, \quad y_{r,k} \in \{0, 1\}, \ y_{d,i} \in \{0, 1\}$$

(47)

Here, $obj$ is the function to be minimized, which can be the total annualized cost (TAC) or minus profit. Equations 1–25 pertain to the modeling reactor network, eqs 26–33 pertain to the modeling distillation sequence, and eqs 34–37 and 39–47 pertain to the modeling distillation column with the bypass efficiency method. $x$ denotes continuous variables with lower bound $x^L$ and upper bound $x^U$. $y_{r,k}$ represents binary variables for determining the reactor network, and $y_{d,i}$ represents binary variables for determining the distillation sequence. The problem $P_0$ is a strongly nonconvex MINLP problem.

5. SOLUTION APPROACH

The optimization problem $P_0$ is extremely difficult to solve to optimality or even find a feasible solution, because of many strongly nonconvex nonlinear terms existing in the reactor model and the MESH equations. General global optimization solvers such as BARON and ANTIGONE often fail to find a feasible solution within acceptable computational effort. To solve the problem $P_0$, an efficient solution approach with a systematic and effective initialization procedure using the hybrid steady-state and time-relaxation optimization algorithm is proposed. The entire solution approach is illustrated in Figure 4.

5.1. Initialization. Many off-the-shelf optimization solvers for MINLP problems such as DICOPT and SBB must solve a relaxed MINLP (RMINLP) problem in the beginning, which is actually a NLP problem. However, the MESH
equations for the distillation system are large-scale and strongly nonlinear, which often results in infeasibility of the RMINLP problem. As a result, problem $P_0$ is regarded as infeasible. It is important to get an optimal solution for the RMINLP problem. As the CONOPT solver, using the generalized reduced gradient (GRG) algorithm often demonstrates good convergence for solving strongly nonlinear problems based on our computational experience, it is used in this work to solve NLP problems. However, the CONOPT solver has a tendency to fail due to difficulty in finding a feasible solution to start the optimization. To enhance the convergence of the CONOPT solver, a systematic and robust initialization strategy is proposed to identify a feasible initial solution for the NLP problem to be solved. In this initialization strategy, the reaction–separation–recycle process is decomposed into the reaction system, flash units, and the distillation system without recycling of the recovered raw materials.

We first optimize the reaction system and flash units to generate the feed condition to the distillation system. The mathematical formulation for the reaction system is also a MINLP, which can be solved to local optimality using the SBB solver. However, it is found that, in some situations, SBB fails at the root node, where a relaxed MINLP problem is required to be solved, leading to infeasibility for the MINLP problem. To enhance the robustness of SBB, a global optimization solver BARON is used to identify a feasible solution for the relaxed MINLP problem. As a result, we can ensure SBB can find a feasible solution at the root node. Since BARON is used to only generate a feasible solution instead of an optimal solution, the objective function with a constant value (e.g., 0) is used instead of TAC of the reaction system, and the equations related to the economic evaluation is removed from the relaxed MINLP problem to further reduce the complexity. After solving, the economic evaluation is conducted and a feasible solution for the relaxed MINLP is generated. With this feasible solution as an initial point, the reactor network model can be solved to local optimum easily using SBB. Sequentially, the reactor network, together with flash units, can be solved using SBB if there are flash units following the reactor network, generating the feed condition to the distillation system.

Because of large-scale strongly nonlinear and nonconvex MESH equations in the distillation system, BARON cannot provide a feasible solution for the distillation system within an acceptable time. Hence, the hybrid steady-state and time-relaxation optimization algorithm is employed to optimize the distillation system to get a feasible solution with the assumption of the direct sequence used and the bypass efficiencies for all stages to be 1. This hybrid algorithm decomposes the original NLP problem for the distillation system into a small-scale optimization problem in the outer level and a large-scale simulation problem in the inner level. This hybrid algorithm has demonstrated good convergence in optimizing the distillation system. To guarantee better convergence and a higher computational efficiency, this hybrid algorithm is only used to identify a feasible solution for the distillation system instead of a locally optimal solution. Therefore, a constant objective function (e.g., 0) is used, which can reduce the computational effort for solving the difficult NLP problem. Simulation of the distillation system in the inner level is conducted with the pseudotransient continuation (PTC) distillation model developed by Ma et al. Note that, at the beginning of the hybrid algorithm, a converged simulation of the distillation system is required, which provides a starting point for the hybrid algorithm. This converged simulation solution is not difficult to obtain as the desired purity or recovery is not necessary to be achieved. For instance, it can be generated by setting the reflux ratio to 1 and reboiled fraction to 0.5 in each distillation column. The pressure is set as 1 bar in each column if it also needs to be optimized. For more details about this hybrid optimization algorithm, the reader can be referred to the report by Ma et al.

Since the hybrid algorithm is implemented in a different platform as discussed later, which may cause the tolerance mismatch problem when the solution from the hybrid algorithm is used in GAMS. To avoid such tolerance mismatch, we use the solution from the hybrid algorithm as an initial point for GAMS/CONOPT to optimize the distillation system model again in GAMS with the feed stream, distillation sequence, and bypass efficiencies fixed to those in the initial point. As we only want to find a feasible solution for the distillation system model in GAMS, a constant objective function (e.g., 0) is also used instead of TAC of the distillation system. After that, a feasible solution for the reaction–separation–recycle process without recycling of recovered raw materials is obtained in GAMS. Although this solution may not be feasible for the original reaction–separation–recycle process, it is close to a feasible solution as the purity of the recovered raw material streams are constrained to be at least 99% (molar) in the feasible solution found for the distillation system. We then use this feasible solution for the reaction–separation process without recycle as an initial point to solve a partially relaxed MINLP problem of the original problem $P_0$, where $y_{i,k}$ is relaxed to be a continuous variable with a value of 0–1, but the distillation sequence and bypass efficiencies are still fixed to the values in the feasible solution for the reaction–separation process without recycle. This partially relaxed MINLP problem is solved using GAMS/CONOPT to generate a locally optimal solution, which is a good initial point for the original problem $P_0$.

5.2. Relaxation. With the initialization, the convergence performance of the off-the-shelf optimization solvers for MINLP problems such as DICOPT and SBB can be improved at the root node. However, they still demonstrated infeasibility or converged to a worse locally optimal solution in some cases. This is because the change in the distillation sequence represented by the values of binary variables $y_{i,k}$ can result in a significant change in the optimal values of other continuous variables, such as temperature, pressure, and composition distribution in each distillation column. As a result, it has significant influence on the convergence performance of the algorithm for NLP subproblems during the branch and bound (B&B). In a B&B algorithm, the optimal solution of a new NLP subproblem may be largely different from that of the last NLP subproblem in order to meet the product purity and recovery. Therefore, the solution from the last NLP subproblem may not be good enough to be used to initialize the current NLP subproblem, leading to infeasibility of the current NLP subproblem, even if a feasible solution actually exists. If this happens to all nodes, then no feasible solution can be found for the MINLP problem. This happens when high purity and/or recovery such as 0.9995 are required. Even if a locally optimal solution is finally found but with many infeasible NLP subproblems during B&B (e.g., 71 infeasible nodes out of 150 searched nodes in the work by Zhang et al.), a better locally optimal solution may have been
excluded, because of infeasibility in solving NLP subproblems at some nodes.

To avoid the above issues during the B&B algorithm, a successive relaxed MINLP (SRMINLP) solution approach\(^8\) is adopted to solve a few NLP problems. In this approach, the integrality constraints for \(y_{d,j}\) are relaxed. In other words, \(y_{d,j}\) becomes a continuous variable with a value of 0–1. The complementary constraints with a controlled parameter \(\mu_{d,j}\) are introduced, as shown in eq. 48.

\[
y_{d,j}(1 - y_{d,j}) \leq \mu_{d,j} \quad \forall \; j, \; d \in D_j
\]  
\[
(48)
\]

where \(\mu_{d,j}\) is a parameter that is used to control the range of \(y_{d,j}\). When \(\mu_{d,j} > 0.25\), the feasible region given by eq. 48 is larger than that given by the hypercube of \(y_{d,j}\). As a result, eq. 48 is actually inactive. On the other hand, when \(\mu_{d,j} = 0\), \(y_{d,j}\) can only be 0 or 1. The relaxation problem of P0 is given as follows, which is denoted as eq SP0.

\[
\min \quad \text{obj}
\]
\[
\text{s.t.} \quad \text{eqs } 1-25, 26-33, 34-37, 39-47, 48
\]
\[
x^L \leq x \leq x^U, \quad 0 \leq y_{d,j} \leq 1
\]
\[
(\text{SP0})
\]

Note that if \(\mu_{d,j} = 0\), then the optimal solution of eq SP0 is also the optimal solution of eq P0. However, when \(\mu_{d,j} = 0\), the problem is hard to solve. To alleviate the solution difficulty, we first set \(\mu_{d,j} = 0.25\) to solve problem SP0, which allows some relaxation for the complementary constraints and, hence, a locally optimal solution can be generated more easily. Then, we asymptotically reduce the value of \(\mu_{d,j}\) to zero as the solution of problem SP0 with a larger value of \(\mu_{d,j}\) provides a good initial point for solving problem SP0 with a smaller value of \(\mu_{d,j}\).

We treat \(\mu_{r,k}\) in a similar way with eq 49.

\[
y_{r,k}(1 - y_{r,k}) \leq \mu_{r,k} \quad \forall \; r, \; k \in R^k
\]
\[
(49)
\]

The original problem P0 is then transformed to a pure NLP problem, which corresponds to problem SP1 and is solved with both \(\mu_{r,k}\) and \(\mu_{d,j}\) at 0.25 in the beginning and then asymptotically reduce their values to zero.

\[
\min \quad \text{obj}
\]
\[
\text{s.t.} \quad \text{eqs } 1-25, 26-33, 34-37, 39-47, 48, 49
\]
\[
x^L \leq x \leq x^U, \quad 0 \leq y_{d,j} \leq 1, \quad 0 \leq y_{r,k} \leq 1
\]
\[
(\text{SP1})
\]

The entire solution approach in Figure 4 can be stated as follows. We first use BARON\(^{46}\) to find a feasible solution for the reactor network model with binary variables relaxed and economic evaluation equations excluded. The economics then are evaluated with the derived reactor sizes and heat duties. With the above results as an initial point, the whole reactor network is solved to a local optimum using an existing MINLP solver, such as GAMS/SBB\(^{48}\) and then the flash units are solved together with the reactor network if there are flash units following the reactor network. The obtained effluent leaving the reactor network or flash units is used as the feed to the distillation system, which is optimized with the direct distillation sequence and all bypass efficiencies being 1 to get a feasible solution using the hybrid optimization algorithm.\(^{44}\) Note that the feasible solution for the distillation system from the hybrid algorithm should be further optimized with a fixed separation flowsheet configuration using GAMS/CONOPT solver to avoid inconsistent tolerance between different platforms. Combining the solutions from both reactor network/flash units and the distillation system, we generate a feasible solution for the reaction—separation system without recycle with the purity of the recycle stream constrained to be at least 0.99. With this solution as an initial point, we further use GAMS/CONOPT solver to find a feasible solution for a partially relaxed optimization problem of the reaction—separation—recycle process in which the binary variables in the reactor network are relaxed and the configuration for distillation sequence is fixed. We then enter the successive relaxed algorithm, SRMINLP. We use s and S to denote the iteration and the maximum number of iterations in SRMINLP, respectively. The algorithm is stated as follows.

**Step 1:** let \(s = 0\), \(S = 50\), \(\mu_{d,i}^* = 0.25\), \(tol = 1 \times 10^{-7}\)

**Step 2:** Solve problem SP1 with \(\mu_{d,i}^*\) and \(\mu_{r,k}^*\). If an optimal solution for problem SP1 is generated, then go to step 3. Otherwise, go to step 4.

**Step 3:** Calculate the current values of \(y_{d,i}(1 - y_{d,i})\) and \(y_{r,k}(1 - y_{r,k})\) denoting as \(\mu_{d,i}^*\) and \(\mu_{r,k}^*\). If \(\mu_{d,i}^*\) and \(\mu_{r,k}^*\) are less than \(tol\), exit the iteration and go to step 5. Otherwise, let \(\mu_{d,i}^* = 0\) and \(s = s + 1\), and go to step 2.

**Step 4:** If \(0 < s < S\), then \(\mu_{d,i}^*\) and \(\mu_{r,k}^*\) increase to 0.5\((\mu_{d,i}^* + \mu_{r,k}^*)\) and 0.5\((\mu_{d,i}^* + \mu_{r,k}^*)\), respectively. Let \(s = s + 1\) and go to step 2 to solve SP1 again, which is easier to solve and still leads to some progress as \(\mu_{d,i}^*\) and \(\mu_{r,k}^*\) are smaller than \(\mu_{d,i}^*\) and \(\mu_{r,k}^*\). If \(s = S\), go to step 7 and the whole optimization fails.

**Step 5:** If all bypass efficiencies are 0 or 1 only, then go to step 7. Otherwise, go to step 6.

**Step 6:** Add constraint (38) to problem SP1 and solve it to generate a locally optimal solution where all bypass efficiencies are 0 or 1 only.

**Step 7:** Return the results.

Note that the final optimal solution is a local optimum, because the proposed algorithm cannot guarantee global optimality.

### 5.3. Implementation.

The proposed solution approach is implemented in Python\(^{51}\) which drives GAMS 24.3.1 (Generalized Algebraic Modeling System)\(^{52}\) for optimization and Aspen Custom Modeler (ACM)\(^{53}\) for simulation. The slsqp optimizer\(^{54}\) built-in SciPy\(^{55}\) is used in the hybrid steady-state and time-relaxation optimization algorithm. The steady-state simulations and PTC simulations of the distillation system are conducted in ACM. While GAMS/SBB solver\(^{54}\) is used to solve MINLP problems, the GAMS/CONOPT solver\(^{53}\) is used to solve all NLP/RMINLP problems, apart from the RMINLP problem for the reactor network in the initialization, which is solved by GAMS/BARON solver.\(^{46}\)

### 6. Case Studies

Three case studies are used to illustrate the capability of the proposed solution approach. These three case studies include the benzene chlorination process, the cyclohexane oxidation process, and the hydrodealkylation (HDA) process. While the first two case studies are from Zhang et al.\(^{57}\) the third one with three distillation columns is from Douglas.\(^{58}\) All case studies are solved on a desktop with 3.19 GHz Intel Core i7–8700 CPU, 16GB RAM and 64-bit operating system.


This case study is from Example 2 of Zhang et al.\(^{27}\) It uses benzene and chlorine to produce desirable product chlorobenzene.
Suitable amount of chlorination generates the desired product chlorobenzene, while excess amount of chlorination leads to the generation of byproduct dichlorobenzene. The reactions are illustrated in Figure 5. In this case study, the capital cost evaluation formula does not include a fixed term, so we do not need to introduce the bypass stream in the reactor network. Fresh and recycled raw materials (benzene and chlorine) are mixed together and fed into the reactor network. The highly volatile hydrochloric acid and chlorine are removed from the reactor effluent easily using a flash unit. The mixture of benzene, chlorobenzene, and dichlorobenzene is heated to the bubble point and fed to the distillation system to produce pure products chlorobenzene and dichlorobenzene and recover unreacted benzene. The unreacted benzene is recycled back to the reactor system.

The productivity of chlorobenzene should be at least 50 kmol h\(^{-1}\). The recovery and purity of chlorobenzene and dichlorobenzene are at least 99%. Zhang et al.\(^{27}\) assumed that the recycled benzene stream was pure so that the recycle stream causes negligible change to the feed composition of the reactor, after it is mixed with the fresh benzene. However, this assumption could remove some optimal solution, as demonstrated later. For a fair comparison, we use the same assumption. The physical property model (ideal properties are used) is validated through running the GAMS file provided by Zhang et al.\(^{27}\) The same optimal design results are generated. The optimal design parameters from Zhang et al.\(^{27}\) then are applied to fix the corresponding variables in our model to validate our model. It is found that we generate the same design results with TAC of 0.624 M\(\text{€}\) yr\(^{-1}\), as those reported in Zhang et al.\(^{27}\) After validation, we then use the proposed algorithm to optimize our model. The optimization results are provided in Table 1. There are 8 binary variables (2 binary variables for the distillation sequence and 6 binary variables for reactor selection in the reactor network), 2424 continuous variables and 2455 constraints. We generate the optimal TAC of 0.614 M\(\text{€}\) yr\(^{-1}\) within 23 s. The optimal design is illustrated in Figure 6. As seen from Figure 6, one PFR reactor is selected for the reactor network and the direct sequence is used for the distillation system.

We compare our optimal design with the design of Zhang et al.\(^{27}\) The comparative results are provided in Table 1. From Table 1, it can be seen that the optimization time from the proposed solution approach is only 8% of that of the literature report,\(^{27}\) because of much fewer NLP subproblems solved. It can also be seen that the initialization time is \(~\)7 times of the optimization time in the proposed algorithm, while the initialization method and time are not reported in the literature.\(^{27}\) The TAC in our optimal design is 1.6% smaller than that of the existing work,\(^{27}\) which is benefitted from the slight difference in the distillation column design. The volume of our reactor is slightly larger than that of Zhang et al.\(^{27}\) (20.68 m\(^3\) vs 20.03 m\(^3\)), so we have 2% higher reaction conversion but 0.6% lower reaction selectivity. Our optimal design uses 18 stages in the first column and 11 stages in the second column, which is reduced by 2 and 3, respectively, compared to those of Zhang et al.\(^{27}\) (20 and 14). The reboiler duty in the first column of our design is also slightly lower than that of the literature report\(^{27}\) (0.49 MW vs 0.50 MW), while the duties of the second reboiler in both designs are almost the same.

We remove the assumption of pure benzene in the recycle stream and allow its purity to vary between 0 and 1. The optimal TAC is 0.619 M\(\text{€}\) yr\(^{-1}\) with the optimal purity of benzene in the recycle stream being 0.98. The new optimal TAC is slightly larger than 0.614 M\(\text{€}\) yr\(^{-1}\), which is mainly due to the effect of the impurity considered in the recycle stream. Such effect increases the recycle flow rate and leads to higher capacity requirement, which was not considered in Zhang et al.\(^{27}\) Although the benefit of optimizing the recycle purity is not evident in this case study as the current optimal purity is close to the given purity requirement, it will be found to be rather important in the next case study.

![Figure 5. Superstructure for the benzene chlorination process.](Image)

![Table 1. Comparative Results of Our Design and the Design Reported by Zhang et al.\(^{27}\) for Case Study 1](Table)

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of binary variables</td>
<td>86</td>
<td>6</td>
</tr>
<tr>
<td>reactor network</td>
<td>one PFR</td>
<td>one PFR</td>
</tr>
<tr>
<td>reaction conversion</td>
<td>69%</td>
<td>71%</td>
</tr>
<tr>
<td>reaction selectivity</td>
<td>91.3%</td>
<td>90.7%</td>
</tr>
<tr>
<td>reactor volume (m(^3))</td>
<td>20.03</td>
<td>20.68</td>
</tr>
<tr>
<td>separation sequence</td>
<td>direct</td>
<td>direct</td>
</tr>
<tr>
<td>(N_{\text{tray}}) in column 1</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>reboiler duty in column 1 (MW)</td>
<td>0.50(^a)</td>
<td>0.49</td>
</tr>
<tr>
<td>(N_{\text{tray}}) in column 2</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>reboiler duty in column 2 (MW)</td>
<td>0.64(^a)</td>
<td>0.64</td>
</tr>
<tr>
<td>TAC (M(\text{€}) yr(^{-1}))</td>
<td>0.624</td>
<td>0.614</td>
</tr>
<tr>
<td>initialization time (s)</td>
<td>(b)</td>
<td>20</td>
</tr>
<tr>
<td>optimization time (s)</td>
<td>36</td>
<td>3</td>
</tr>
</tbody>
</table>

\(^{a}\)The heat duties reported in Zhang et al.\(^{27}\) divided by a constant 3.6 as a mistake is found in their work. \(^{b}\)The initialization time is not reported in Zhang et al.\(^{27}\)

This case study is also from Example 1 of Zhang et al.²⁷ It uses cyclohexane to produce a mixture of cyclohexanol and cyclohexanone (i.e., KA oil) through oxidation. Suitable oxidation leads to KA oil, while overly oxidation will generate byproduct adipic acid. The reaction pathways are provided in Figure S2 in the Supporting Information. The superstructure for this case study is provided in Figure S3 in the Supporting Information, where only CSTR reactors are used, which is the same as that of Zhang et al.²⁷ Air and cyclohexane are fed to the reactor network with high pressure. Only liquid effluent is drawn from the reactor network, which is separated using two distillation columns to generate the desired product and byproduct after passing a pressure-relief valve to atmospheric pressure. The recovered cyclohexane from the distillation system is recycled back to the reactor network. The purity and recovery of the products and raw materials are at least 0.9995. The Wilson equation is used to calculate liquid activity coefficients. The vapor is assumed to be ideal gas as the distillation columns are operated at the atmospheric pressure. The objective is to minimize TAC including energy cost and annualized capital cost. The models and parameters for physical property and TAC calculation are given in the Supporting Information.

Similar to case study 1, we first validate the design results reported in Zhang et al.²⁷ through running the GAMS file provided. The optimal design parameters are used to validate our model. It is shown that the design results with TAC of 0.231 M€ yr⁻¹ is obtained, which is almost the same as that of the 0.229 M€ yr⁻¹ reported in Zhang et al.²⁷ After validation, we use the proposed algorithm to solve the model. The optimization problem involves 12 binary variables (2 binary variables to denote the distillation sequence and 10 binary variables to select the reactor), 4647 continuous variables, and 4646 constraints. The optimization results are provided in Table 2. From Table 2, we generate the optimal TAC of 2.120 M€ yr⁻¹ within 219 s. The optimal design is illustrated in Figure 7. The optimal flowsheet requires 3 series CSTRs. The conversion and selectivity reach 10.8% and 58%, respectively. The distillation column 1 requires 15 trays, and the column 2 requires 4 trays.

We also compare our optimal design with that reported in Zhang et al.²⁷ The comparative results are provided in Table 2. From Table 2, it can be seen that the optimization time from the proposed approach is 11 s, which is 5% of that in Zhang et al.²⁷ The initialization time of the proposed algorithm is still much longer than the optimization time. The TAC in our design is 4.9% less than that of Zhang et al.²⁷ This is because our reactor network achieves both higher conversion (10.9% vs 10.6%) and higher selectivity (61% vs 58%), which lowers the burden of the distillation system. As a result, the distillation system in our design consumes less energy in both columns and requires less number of stages at the same time. To achieve such better reaction performance, the annualized capital cost of the reactor system in our design is slightly higher than in the literature²⁷ (0.58 M€ yr⁻¹ vs 0.53 M€ yr⁻¹). However, the annual cost of the distillation system in our design is lower (1.54 M€ yr⁻¹ vs 1.69 M€ yr⁻¹). Evidently, the decrease in the annual cost of the distillation system in our design outweighs the increase in the annual cost of the reactor system.

Similar to case study 1, the assumption of pure cyclohexane in the recycle stream is also removed. We allow the purity of cyclohexane to vary between 0 and 1. The optimal TAC is reduced to 1.845 M€ yr⁻¹, which is decreased by 13.0%, compared to the optimal TAC of 2.120 M€ yr⁻¹ and by 17.2%, compared to the optimal TAC (2.229 M€ yr⁻¹) of Zhang et al.²⁷ The cyclohexane purity in the recycle stream is 0.9870, instead of 0.9995. It indicates that assuming pure cyclohexane in the recycle stream can miss a much better locally optimal solution. In other words, the purity in the recycle stream should be optimized simultaneously in order to generate better design.


The HDA process is from Douglas,³ which uses toluene and hydrogen to produce desired product benzene and byproduct diphenyl. The reactions are illustrated

Table 2. Comparative Results of Our Design and Design of Zhang et al.²⁷ for Case Study 2

<table>
<thead>
<tr>
<th>Value</th>
<th>Zhang et al.²⁷</th>
<th>ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of binary variables</td>
<td>91</td>
<td>10</td>
</tr>
<tr>
<td>reactor network</td>
<td>2 serial CSTRs</td>
<td>3 serial CSTRs</td>
</tr>
<tr>
<td>reaction conversion</td>
<td>10.6%</td>
<td>10.9%</td>
</tr>
<tr>
<td>reaction selectivity</td>
<td>58%</td>
<td>61%</td>
</tr>
<tr>
<td>separation sequence</td>
<td>direct sequence</td>
<td>direct sequence</td>
</tr>
<tr>
<td>Ntray in column 1</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>reboiler duty in column 1 (MW)</td>
<td>5.83</td>
<td>5.27</td>
</tr>
<tr>
<td>Ntray in column 2</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>reboiler duty in column 2 (MW)</td>
<td>1.25</td>
<td>1.21</td>
</tr>
<tr>
<td>TAC (M€ yr⁻¹)</td>
<td>2.229</td>
<td>2.120</td>
</tr>
<tr>
<td>initialization time (s)</td>
<td>a</td>
<td>208</td>
</tr>
<tr>
<td>optimization time (s)</td>
<td>202</td>
<td>11</td>
</tr>
</tbody>
</table>

Note: The initialization time is not reported in Zhang et al.²⁷
The reaction kinetics are obtained from the GAMS Model Library. For an isothermal PFR, the kinetic equations are provided below:

\[ 1 - X = \frac{1}{1 + 0.372kVF_{\text{in}}^{0.372}F_{\text{tot}}^{1.544}} \] (50)

\[ 1 - Y = 0.0036(1 - X)^{-1.544} \] (51)

\[ k = 6.3 \times 10^{10} \exp \left( -\frac{26167}{RT} \right) \] (52)

where \( X \) is toluene conversion, \( k \) is a reaction kinetic constant (kmol m\(^{-3}\) s\(^{-1}\)), \( V \) is the reactor volume (m\(^3\)), and \( F_{\text{in}}^{0.372} \) and \( F_{\text{tot}}^{1.544} \) are the toluene flow rate and total flow rate (kmol s\(^{-1}\)) in the feed stream, respectively. \( Y \) is the reaction selectivity of benzene. \( R \) (kJ kmol\(^{-1}\) K\(^{-1}\)) is the ideal gas constant and \( T \) is the reaction temperature (K). For an adiabatic reactor, the average temperature between the inlet and outlet temperatures are used in eq 52.

The superstructure is illustrated in Figure 8. We do not consider additional process options such as recovery of H\(_2\) from the purge stream using membrane as only distillation columns are used as the separation technologies in this work. The required benzene molar purity is 99.97% with a production rate of 124.8 kmol h\(^{-1}\). The objective is to maximize the economic potential, which is calculated by the revenue of benzene and diphenyl minus annualized capital cost and operating cost. Most data are obtained from Kocis and Grossmann, including raw material and product price, and economic evaluation correlations, which are shown in the Supporting Information. We also assume ideal vapor and liquid.
in the system, which is consistent with those from Kocis and Grossmann. The Antoine equation from the HDA GAMS file in GAMS Model Library is used to calculate the saturated pressure. The enthalpies of H$_2$ and CH$_4$ in the liquid are assumed to be the same as their vapor enthalpies, because the critical temperatures of H$_2$ and CH$_4$ are much lower than the temperature lower bound in the system (i.e., 300 K). The Aspen ideal gas heat capacity polynomial is used to calculate temperature lower bound in the system (i.e., 300 K). The enthalpies of a component are calculated using heat capacity and the DIPPR correlation is used to calculate the liquid heat capacity instead of assuming a constant in Kocis and Grossmann. The vapor and liquid enthalpies of a component are calculated using heat capacity correlations with the reference enthalpy at 300 K. Similar to Kocis and Grossmann, the overall stage efficiency is assumed to be 0.5. All equations used to calculate the saturated pressure, heat capacity, and enthalpy are provided in the Supporting Information.

There are 8142 constraints, 8643 continuous variables, and 13 binary variables (10 binaries for distillation sequence and 3 binaries for the reactor network) in the optimization problem. We generate the optimal TAC of 4.956 M$\text{S yr}^{-1}$ within 111 s. The time for initialization and optimization is 93 and 18 s, respectively. The optimal design is depicted in Figure 9. From Figure 9, we can see that one adiabatic reactor is used for reaction, and the direct sequence is used for separation in our design, which are the same as those in the work of Kocis and Grossmann. However, that work reported a higher profit of 5.887 M$\text{S yr}^{-1}$ than our design (4.957 M$\text{S yr}^{-1}$), because of membrane separation used to recover hydrogen, which is not considered in the current work. Another possible reason is that the short-cut distillation model is used in the existing work, while we use the rigorous distillation column model.

We also use DICOPT and SBB to solve these three case studies. The computational results are provided in Table 3. Note that the feasible solution generated from the proposed initialization procedure is used as an initial point for DICOPT and SBB. This is because DICOPT and SBB often fail at the root node with the default initial point. From Table 3, it can be seen that SBB and DICOPT can generate the same optimal TAC (0.614 M$\text{S yr}^{-1}$) as that from the proposed algorithm for case study 1. While DICOPT and SBB require 94 and 29 s, respectively, the proposed algorithm requires 23 s. Neither SBB nor DICOPT can find a feasible solution for case study 2; even both of them use the initial point provided from the proposed initialization procedure, while the proposed algorithm still can find a locally optimal solution. For case study 3, SBB can find a locally optimal solution of 4.957 M$\text{S yr}^{-1}$ within 117 s, which is almost the same as that (4.956 M$\text{S yr}^{-1}$) from the proposed algorithm, while DICOPT fails to provide a feasible solution.

We also use the proposed algorithm to solve these three case studies using other GAMS versions to illustrate its robustness. These GAMS versions include GAMS 24.0.1, 24.3.3, and 24.5.6. The computational results are presented in Table 3. From Table 3, it can be seen that the proposed algorithm can provide locally optimal solutions for all three of these case studies under different GAMS versions, while SBB can only converge under some of the tested GAMS versions for cases 2 and 3. DICOPT has the worst convergence performance, because it cannot solve case study 2 and 3. The model of Zhang et al. generates evidently different solutions under different GAMS versions. Most importantly, it fails to provide a feasible solution for case study 2 using the GAMS 24.3.3 and 24.5.6. From Table 3, it can also be observed that the proposed solution approach may converge to a worse locally optimal solution for case study 3 using GAMS 24.0.1, because of different initial points generated from the proposed initialization procedure.

7. CONCLUSION

In this work, we proposed a computationally efficient optimization-based framework for simultaneous synthesis and design of the reaction—separation—recycle processes using rigorous models. While the reactor network was represented using the superstructure of Lakshmanan and Biegler, the separation network was represented using SEN. The superstructure was modeled through GDP, which was reformulated to a MINLP model, using the convex hull method. The bypass efficiency method was adopted to represent the number of stages in a distillation column, which did not introduce additional integer variables. To solve the highly nonconvex MINLP model to local optimality, a systematic solution approach was proposed in which a systematic initialization procedure was proposed to generate a feasible solution for a partially relaxed problem of the reaction—separation—recycle process, which was used as an initial point for optimization of the reaction—separation—recycle process. To alleviate the difficulty in solving the MINLP problem directly, a successive relaxed MINLP solution strategy was then adopted to solve the model to local optimality. Although the complementary constraints resulted in more difficulty, it seems that adding the complementary constraints asymptotically could guarantee good convergence performance. The case studies demonstrated that we can not only solve all the case studies in the work of Zhang et al. with better optimal solutions and significantly less computational effort by ~1 order of magnitude, but also synthesize a much more complex HDA process within 2 min. In addition, we showed that optimizing the purity of the recycle stream simultaneously with the whole reaction—separation—recycle process can save ~13% for the TAC.

Table 3. Optimal TAC for Case Studies 1–3 Using Different GAMS Versions

<table>
<thead>
<tr>
<th>GAMS</th>
<th>Case Study 1</th>
<th>Case Study 2</th>
<th>Case Study 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zhang et al.</td>
<td>ours</td>
<td>SBB</td>
</tr>
<tr>
<td>24.3.1</td>
<td>0.624</td>
<td>0.614</td>
<td>0.614</td>
</tr>
<tr>
<td>24.0.1</td>
<td>0.631</td>
<td>0.614</td>
<td>0.614</td>
</tr>
<tr>
<td>24.3.3</td>
<td>0.624</td>
<td>0.614</td>
<td>0.614</td>
</tr>
<tr>
<td>24.5.6</td>
<td>0.778</td>
<td>0.614</td>
<td>0.614</td>
</tr>
</tbody>
</table>

*DICOPT fails for case studies 2 and 3.*
Although the developed framework is robust and fast for the simultaneous synthesis and design of the reaction—separation—recycle processes using rigorous models, it could generate a worse locally optimal solution from some initial points. This will be addressed in the future work. In addition, we assumed ideal vapor phase, neglected mixing enthalpy and used the ideal PFR and CSTR reactors in the reactor network. In the future, we will extend the proposed optimization-based framework to consider nonideal vapor phase, mixing enthalpy, and other nonideal reactors in the reactor network, which may increase the difficulty for the proposed algorithm to find a locally optimal solution, because of additional nonconvex terms introduced. Another future work is to consider more-complex distillation configurations such as thermally coupled distillation columns and dividing-wall columns for higher separation efficiency. Although these complex distillation configurations may result in further challenges for both superstructure model and solution algorithms, such as numerical difficulty caused by zero flow rates, the proposed solution algorithm still has great potential to solve such more-complex distillation configurations.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.iecr.1c00250.

Parameters and specifications in the case study 1 (Table S1); parameters for component physical properties in the case study 1 (Table S2); steam used in different equipment in the case study 1 (Table S3); parameters and specifications in case study 2 (Table S4); parameters for component physical properties in case study 2 (Table S5); parameters for Wilson equations (Table S6); feedstock and product specification and price (Table S7); utility prices for HDA process (Table S8); investment costs for HDA process (Table S9); parameters for Antoine equation (Table S10); parameters for ideal vapor heat capacity polynomial (Table S11); parameters for DIPPR liquid heat capacity polynomial (Table S12); number of binary variables required for separation of a mixture with different numbers of components in the two SEN representations (Table S13); reactions in benzene chlorination process (Figure S1); reaction pathway for cyclohexane oxidation (Figure S2); superstructure for the cyclohexane oxidation process (Figure S3); reactions in the HDA process (Figure S4) (PDF)

AUTHOR INFORMATION

Corresponding Author
Jie Li — Centre for Process Integration, Department of Chemical Engineering and Analytical Science, School of Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom; orcid.org/0000-0001-5196-2136; Phone: + 44 (0) 161 306 8622; Email: jie.li-2@manchester.ac.uk; Fax: +44 (0) 161 236 7439

Authors
Yingjie Ma — Centre for Process Integration, Department of Chemical Engineering and Analytical Science, School of Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom; orcid.org/0000-0003-1465-3873

Zekun Yang — Centre for Process Integration, Department of Chemical Engineering and Analytical Science, School of Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom

Aline El-Khoury — Centre for Process Integration, Department of Chemical Engineering and Analytical Science, School of Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom

Nan Zhang — Centre for Process Integration, Department of Chemical Engineering and Analytical Science, School of Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom

Bingjian Zhang — School of Materials Science and Engineering, Guangdong Engineering Technology Research Center for Petrochemical Energy Conservation, Sun Yat-sen University, Guangzhou S10275, China; orcid.org/0000-0001-9830-9193

Li Sun — School of Computing, Engineering and Physical Sciences, The University of the West of Scotland, Scotland PA1 2BE, United Kingdom

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acs.iecr.1c00250

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge financial support from China Scholarship Council—The University of Manchester Joint Scholarship (No. 201809120005).

REFERENCES


https://doi.org/10.1021/acs.iecr.1c00250


(57) General Algebraic Modeling System (GAMS), Release 24.0.1; GAMS Development Corporation; Fairfax, VA, USA, 2012.

(58) General Algebraic Modeling System (GAMS), Release 24.3.3; GAMS Development Corporation; Fairfax, VA, USA, 2014.

(59) General Algebraic Modeling System (GAMS), Release 24.5.6; GAMS Development Corporation; Fairfax, VA, USA, 2015.