

A hybrid stochastic-deterministic optimization approach for integrated solvent and process design

Teng Zhou ^{a,*}, Yageng Zhou ^b, Kai Sundmacher ^{a,b}

^a *Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical
Systems, Sandtorstr. 1, D-39106 Magdeburg, Germany*

^b *Process Systems Engineering, Otto-von-Guericke University Magdeburg, Universitätsplatz 2,
D-39106 Magdeburg, Germany*

**Corresponding author: Tel.: +49 391 6110 406; Fax: +49 391 6110 353*

E-mail address: zhout@mpi-magdeburg.mpg.de

Abstract

The best solution to computer-aided solvent and process design problems can be only achieved by the simultaneous optimization of solvent molecules and process operating conditions. In this contribution, a hybrid stochastic-deterministic optimization approach is proposed for integrated solvent and process design. It is a combination of a genetic algorithm (GA) that optimizes the discrete molecular variables and a gradient-based deterministic algorithm that solves the continuous nonlinear optimization problem of the process at fixed molecular variables as proposed by the GA. The method is demonstrated on a coupled absorption-desorption process where solvent molecular structures as well as the operating conditions of the absorption and desorption columns are optimized simultaneously. While deterministic mixed-integer nonlinear programming (MINLP) algorithms rely on well-selected initial estimates, the proposed hybrid approach can reliably and steadily solve the problem under random initializations. The combination of the advantages of stochastic and deterministic algorithms makes the approach a promising alternative to conventional MINLP algorithms for solving integrated solvent and process design problems.

Keywords: Computer-aided molecular design (CAMD); Integrated solvent and process design; Genetic algorithm; Hybrid stochastic and deterministic optimization; Absorption-desorption processes

Introduction

The selection of solvents or other functional chemicals is significant for achieving processes with better economics and lower environmental impact (Pistikopoulos et al., 2010). During the past several decades, the computer-aided molecular design (CAMD) technique (see review Ng et al., 2015) has been proposed and widely applied to solvent design. For the CAMD of solvents, desirable solvent structures are identified to match user-specified target properties via either multi-level generate-and-test methods (Gani et al., 1991; Pretel et al., 1994; Harper and Gani, 2000; Karunanithi et al., 2005, 2006) or optimization-based mathematical programming methods (Odele and Macchietto, 1993; Kim and Diwekar, 2002; Sinha et al., 1999; Pistikopoulos and Stefanis, 1998; Buxton et al., 1999; Giovanoglou et al., 2003; Marcoulaki and Kokossis, 2000; van Dyk and Nieuwoudt, 2000; Roughton et al., 2012; Cheng and Wang, 2010; Folić et al., 2007, 2008; Samudra and Sahinidis, 2013; Zhou et al., 2015a, 2015b).

The traditional CAMD of solvents is based on solvent property objectives, such as solute solubility (Karunanithi et al., 2006), separation capacity and selectivity (Odele and Macchietto, 1993), reaction rate (Folić et al., 2007; Zhou et al., 2015a), etc. However, when designing solvents for a specific process, it is very often difficult to relate the performance of the process to a specific solvent property. For example, Kossack et al. (2008) used the separation selectivity as the objective to design entrainers for an extractive distillation process. However, the obtained entrainer was finally found to be undesirable in terms of process economics. Usually, solvents have multiple effects on a process they are designed for. For example, separation agents must offer high selectivity and capacity for the compound being separated, yet they must be easily recovered. Reaction solvents are required to provide high reaction rates but meanwhile, they should be easily separated from products. In order to capture all tradeoffs between different properties of designed solvents, integrated solvent and

process design where the process performance is optimized is highly desirable (Burger et al., 2015; Zhou et al., 2015a). Besides the property tradeoffs, there are also strong interdependencies between the selection of solvents and the process operating conditions. On the one hand, the specification of process conditions strongly influences the selection of a suitable solvent and on the other hand, the selected solvent reversely determines the optimal operating point of the process. Based on the above considerations, it is very clear that the best solution to computer-aided solvent and process design problems can be only achieved by the simultaneous optimization of solvent molecules and process operating conditions based on a single process-wide objective.

Integrated solvent and process design usually leads to complex mixed-integer nonlinear programming (MINLP) problems. The solution of such problems using standard MINLP algorithms is usually prohibitive due to strong nonlinearities in thermodynamic and process models as well as the large mixed discrete-continuous search space. However, the problems can be tackled by using simplified process models instead of rigorous models. Hamad and El-Halwagi (1998) simultaneously designed mass separating agents and interception networks by using linearized process models. Zhou et al. (2015a) developed a method for the integrated design of reaction solvents and processes where ideal mixtures were assumed and shortcut distillation models were used in the solvent recovery column. Burger et al. (2015) considered reduced process models in a first-step optimization to generate solutions that were used as initial guesses for solving a second MINLP problem where full process models were applied.

Besides using simplified models, complex integrated solvent and process design problems are more often solved via decomposition-based strategies. Hostrup et al. (1999) first used heuristics to eliminate solvent and process flowsheet options, the reduced design problem was then solved numerically. Buxton et al. (1999) proposed a decomposition-based algorithm to solve the integrated solvent and process design problem. The algorithm decomposes the

problem into a primal nonlinear programming (NLP) problem with fixed solvent structures to optimize process conditions and a master problem to optimize the solvent molecular structure. The primal problem starts with several preprocessing steps where candidate solvents are generated and assessed by physical property tests. Unsuitable candidates are discarded and solvent structures are reinitialized. The algorithm was later extended to dynamic processes in Giovanoglou et al. (2003). Papadopoulos and Linke (2006a, 2006b, 2009) proposed a framework that uses multi-objective optimization (MOO) to screen solvents for non-inferior options with respect to properties impacting process performances. A set of Pareto-optimal solvents were generated from the MOO. A molecular clustering approach was then employed to integrate solvent selection and process synthesis based on the obtained optimal solvent set. This method has been extended for use in the optimal design of working fluids (Papadopoulos et al., 2010) and working fluid mixtures (Papadopoulos et al., 2013) for Organic Rankine Cycles (ORC). Eden et al. (2004) and Eljack et al. (2007) suggested another decomposition strategy where molecular properties are first optimized to maximize the process performance. Molecular structures that can match the property targets are then identified from solving a separate CAMD problem. The corresponding algorithm has been developed and used to design blanket wash solvents in Chemmangattuvalappil et al. (2009). Bardow et al. (2010) and Oyarzún et al. (2011) proposed a two-stage method for integrated molecular and process design where continuous parameters characterizing molecules are optimized together with process variables in the first stage. A hypothetical optimal molecule is obtained and mapped onto an existing molecule in the second stage. The method was applied to the integrated design of solvents and a CO₂ capture process. Similarly, Pereira et al. (2011) simultaneously designed alkane solvents and a CO₂ absorption process where the length of alkane was treated as a continuous variable and optimized together with process operating conditions. A real alkane solvent was identified in a subsequent step.

In summary, it is still challenging to directly solve the integrated solvent and process design problems using standard MINLP algorithms. Most of the previous works rely on decomposition-based solution strategies. In the present contribution, a hybrid stochastic-deterministic algorithm is proposed to address the integrated design problem.

This article is organized as follows. After the integrated solvent and process design problem is generalized and mathematically formulated, the hybrid stochastic-deterministic algorithm is introduced. The effectiveness of the algorithm is then demonstrated on an absorption-desorption (AD) process where the absorption solvent and the AD process are simultaneously designed to maximize the process economic performance. Finally, results are discussed and some conclusions are given.

Problem statement

The integrated solvent and process design problem is described as: Given a batch or continuous process requiring a solvent and a process performance index, find an optimal solvent and operating conditions of the process that give a best process performance. The resulting MINLP problem is given by:

$$\min_{x,y} F = f(x, y, q)$$

$$s.t. \quad \mathbf{h}(x, y, q) = \mathbf{0}$$

$$\mathbf{g}(x, y, q) \leq \mathbf{0}$$

$$q = q(x, y)$$

$$\mathbf{c}(y) = \mathbf{0}$$

$$\mathbf{d}(\mathbf{y}) \leq \mathbf{0}$$

$$\mathbf{x} \in \mathbf{R}^m, \mathbf{y} \in \mathbf{N}^r, \mathbf{q} \in \mathbf{R}^n$$

where F is the process performance index, such as a cost function to be minimized. \mathbf{y} is a r -dimensional integer vector of solvent structural variables. It indicates the number of structural groups present in a solvent molecule. \mathbf{x} is a m -dimensional continuous vector of process variables, such as temperatures, flow rates, compositions, etc. Both \mathbf{x} and \mathbf{y} are independent design variables being optimized. \mathbf{q} is a n -dimensional continuous vector of dependent state variables that are determined by \mathbf{x} and \mathbf{y} . $\mathbf{h}(\mathbf{x}, \mathbf{y}, \mathbf{q}) = \mathbf{0}$ represent thermodynamic and process models, such as activity coefficient equations, mass and heat balances, equipment sizing, etc. $\mathbf{g}(\mathbf{x}, \mathbf{y}, \mathbf{q}) \leq \mathbf{0}$ are specifications on process operating limits. $\mathbf{c}(\mathbf{y}) = \mathbf{0}$ are structural feasibility rules and $\mathbf{d}(\mathbf{y}) \leq \mathbf{0}$ represent molecular complexity and solvent property constraints. Both of them contain molecular variables only.

Hybrid stochastic-deterministic algorithm

Due to the non-convexity of the integrated solvent and process design problem, global solutions can be only obtained by using global MINLP algorithms (Sahinidis, 1996; Misener and Floudas, 2013). However, they are at the moment too expensive to be implemented, especially for problems where rigorous process models are used. For such problems, local deterministic methods are normally employed and these methods cannot guarantee global optimality. Without good initial estimates, poor suboptimal or infeasible solutions are often obtained. Burger et al. (2015) considered reduced process models in a first-step multiobjective optimization to generate a set of Pareto-optimal solutions. These solutions were used as initial

guesses for solving a second MINLP problem where full process models are applied. Starting from different initial estimates, the problem was solved multiple times and the probability of finding high-quality solutions was thus increased.

In comparison with deterministic methods, stochastic optimization methods favor the search of global or near-global designs with random initializations (Biegler, 2014; Rangaiah, 2010). They have been proven to be very efficient for solving discrete optimization problems (Rangaiah, 2010). However, stochastic optimization methods are usually restricted to unconstrained or simple boundary-constrained problems. By contrast, deterministic algorithms can efficiently handle nonlinear constraints and are preferable methods for large-scale nonlinear programming (NLP) problems. Moscato (1989) first introduced a memetic algorithm that combines population-based stochastic algorithms with local refinement strategies. Since then, many combined stochastic and deterministic algorithms have been proposed and applied to various optimization problems including chemical process synthesis (Athier et al., 1997; Urselmann et al., 2011a; Skiborowski et al., 2015) and industry-scale distillation and reactive distillation design (Gómez et al., 2006; Urselmann et al., 2011b). Studies have shown that the optimization approach that combines advantages of stochastic and deterministic algorithms can considerably improve the optimization performance in terms of solution quality and computational cost (Lima et al., 2006; Molina et al., 2010).

The genetic algorithm (GA) is one of the most prominent evolutionary-based stochastic algorithms (Affenzeller et al., 2009). It has been proven to be very efficient in solving optimization problems with large combinatorial and/or discontinuous search spaces, such as

molecular design (Venkatasubramanian et al., 1994; van Dyk and Nieuwoudt, 2000; Xu and Diwekar, 2005; Herring and Eden, 2015). In order to apply the GA to molecular design, molecular structures must be expressed in a readable and operable form. In our previous work (Zhou et al., 2015c), an efficient molecular encoding method was proposed where molecules are represented as tree structures and encoded in the form of a dynamic list with UNIFAC groups as tree nodes. In the dynamic tree structure, each group node is a structural array with four fields (see Figure 1). The first two fields provide group identity information including group ID number and group valence. The last two contain group connectivity information with “Size” denoting the number of group nodes in the sub-tree and “*Child” a pointer indicating the location of connected child nodes. The number of child nodes of a group node depends on the valence of the group and the position of the node. The dynamic tree structure is exemplified with the 2-butanol molecule in Figure 2.

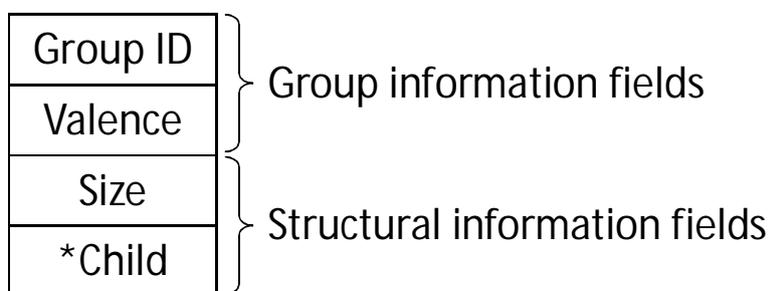


Figure 1: Structure of the group node in molecular trees

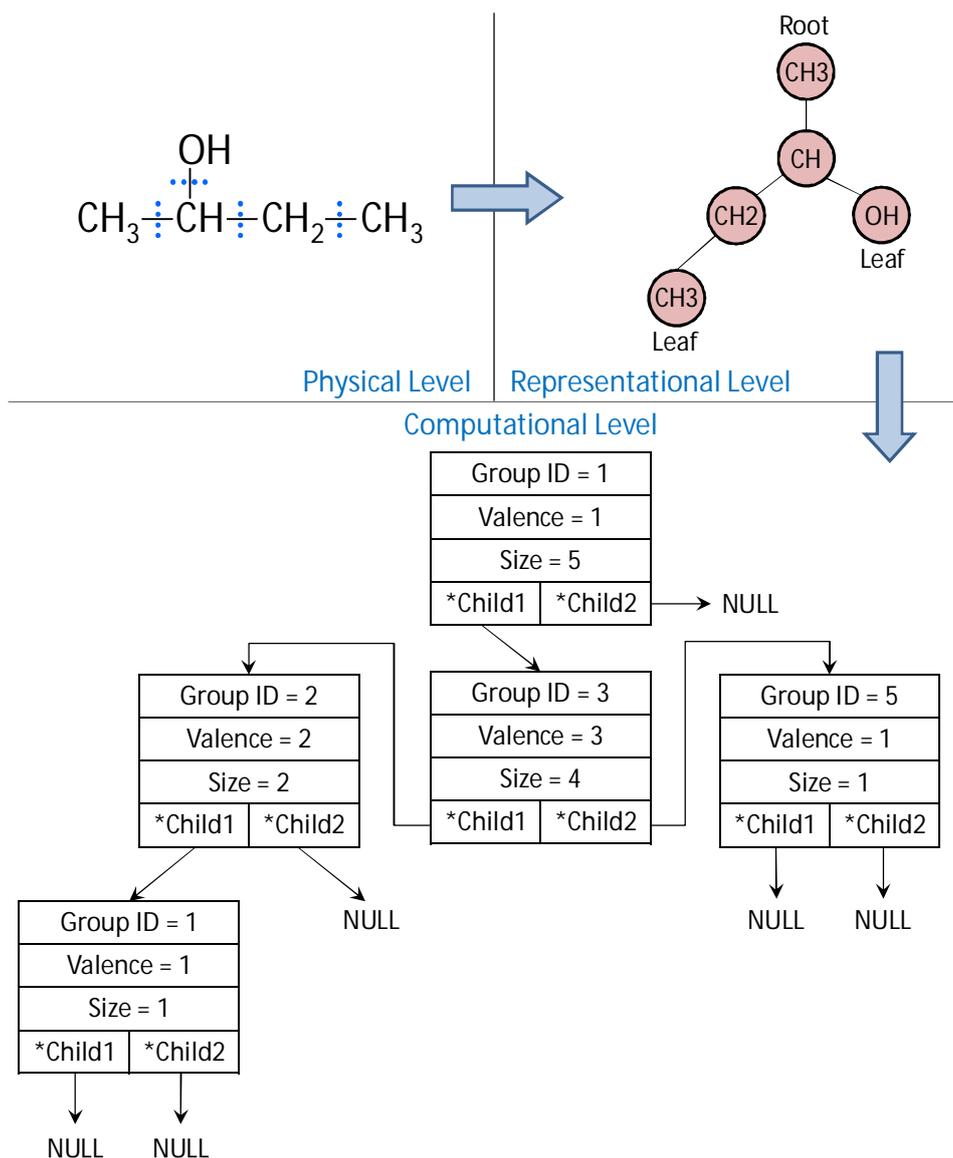


Figure 2: The dynamic tree structure of 2-butanol (Each group is assigned with an ID number.

ID = 1, 2, 3, 4, and 5 represent the CH₃, CH₂, CH, C, and OH groups, respectively)

It is worth noting that the proposed dynamic tree structure facilitates the handling of genetic operations and always guarantees structurally feasible molecules (Zhou et al., 2015c). Based on the molecular encoding method, a GA has been developed and used to design solvents for chemical reactions (for more information, please refer to Zhou et al., 2015c).

In this contribution, a hybrid optimization algorithm that combines the GA developed in Zhou et al. (2015c) with a deterministic NLP algorithm is proposed to solve complex integrated

solvent and process design problems. In the hybrid stochastic-deterministic algorithm, the optimization of discrete molecular variables is performed by the GA and all the continuous process variables are optimized by a gradient-based NLP solver at fixed solvent molecules proposed by the GA. The general structure of the algorithm is shown in Figure 3 with the major steps summarized as follows.

1. To start a program, first set GA parameters and specify a set of structural groups from which solvent molecules are generated.
2. Randomly initialize solvent molecules for the first generation using the dynamic tree structure.
3. Optimize process variables using the gradient-based CONOPT solver (Drud, 1994) at fixed molecular variables and return the optimal solution as well as process objective. This procedure is performed for all the molecules in the generation.
4. Assign a fitness value to each molecule in the generation based on the corresponding process objective value.
5. Select parent molecules from the current generation based on the *roulette wheel selection* rule (Lipowski and Lipowska, 2012).
6. Create offspring molecules for the next generation via genetic operations (Zhou et al., 2015c) performed on the selected parents. Repeat Step 3 and Step 4.
7. The computation is terminated if predefined stopping criteria are satisfied (usually a maximum number of generations is specified). The best solution in the current generation represents the final solution to the design problem which includes an optimal solvent structure and the corresponding best process operating conditions. If the criteria are not satisfied, $i = i + 1$ and return to Step 5.

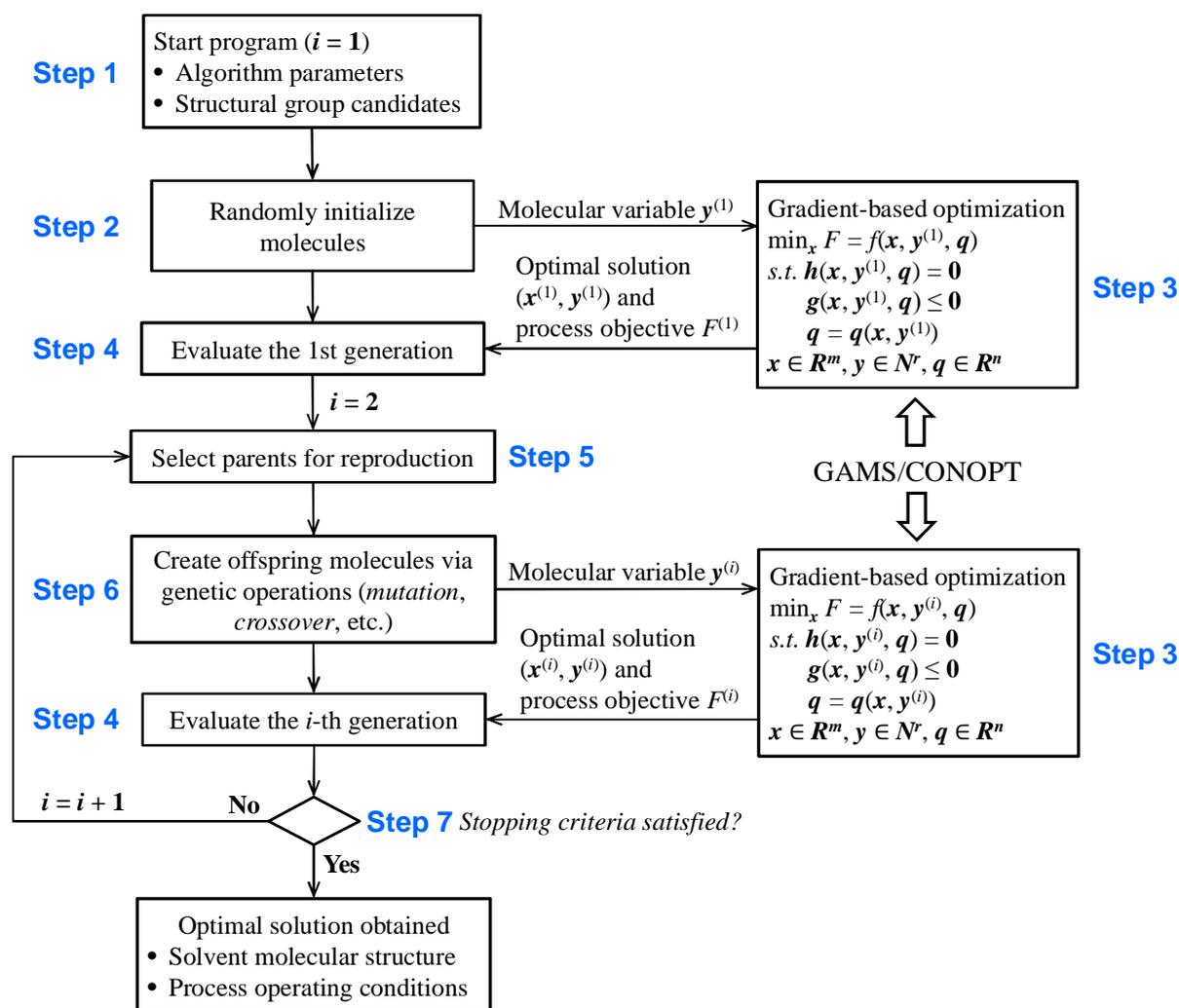


Figure 3: Schematic diagram of the hybrid stochastic-deterministic algorithm

The result of each calculation including the optimal solution and process objective value is recorded in a database, which provides two advantages. On the one hand, if a molecule that has been previously tested is created, the result is directly exported from the database without re-optimizing the process using the NLP solver. This strategy helps saving a lot of computational efforts. On the other hand, the construction of the database facilitates the generation of a list of top solutions after the entire computation terminates. These solutions can be more rigorously evaluated e.g., by experiments before a final decision is made.

Case study

The proposed algorithm is demonstrated on an absorption-desorption (AD) process which consists of an absorption column to separate acetone from air, a desorption column for recovering the absorption solvent, two internal heat exchangers, one cooler, and one heater (see Figure 4). Both absorption and desorption columns are operated at atmospheric pressure. The operating temperatures (T^{ab} and T^{de}) are considered as key process variables to be optimized. Ideal gas phase behaviour is assumed due to the low operating pressure. The UNIFAC model is used to predict the non-ideality of the components in the liquid phases. The dissolution of air in the solvent and the vaporization of the solvent are neglected. A rate-based mass transfer approach is employed to determine the packing heights of the absorption and desorption columns.

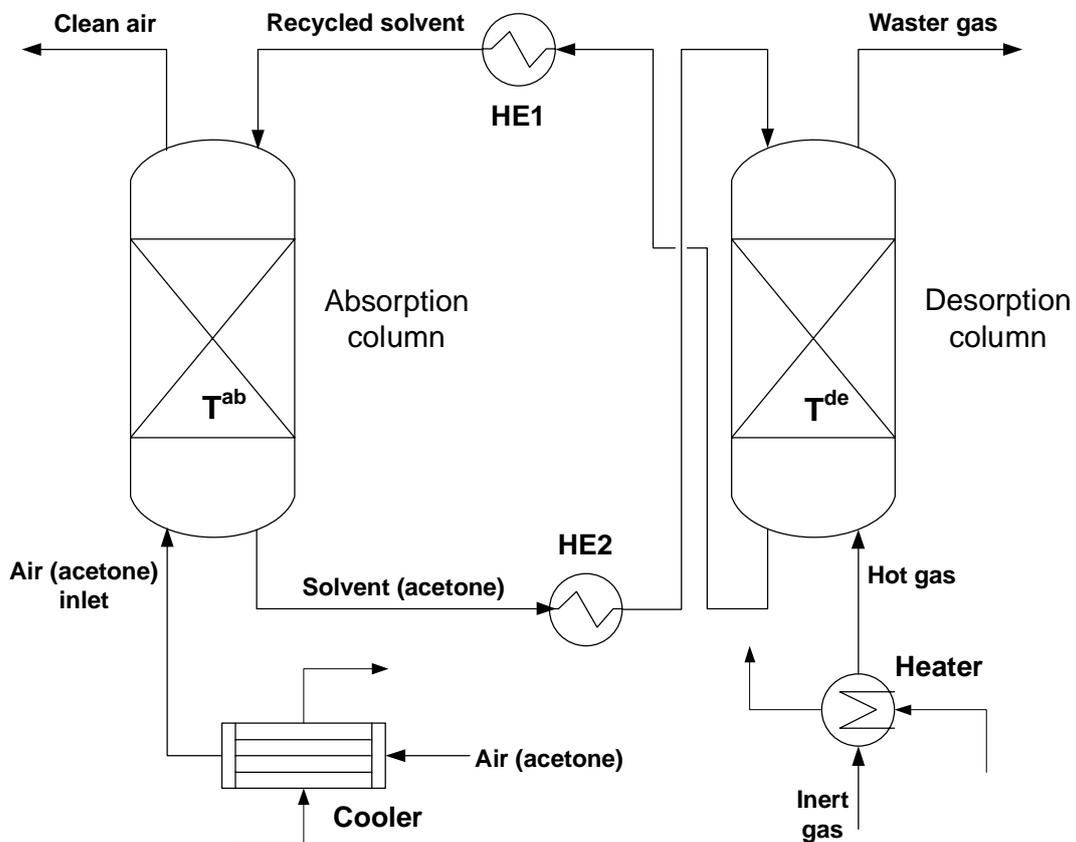


Figure 4: Flowsheet of the AD process

Based on environmental considerations, only non-halogenated acyclic hydrocarbons are considered as potential solvents. As a result, the structural groups from which solvent molecules are built are limited to: CH₃, CH₂, CH, C, OH, CH₃CO (carbonyl), CH₂CO (carbonyl), CHO (aldehyde), CH₃COO (acetate), CH₂COO (acetate), HCOO (formate), OCH₃ (ether), OCH₂ (ether), OCH (ether), and COOH. Structural feasibility constraints are omitted because the employed molecular encoding method already guarantees structurally feasible molecules. Taking into account the limitation and accuracy of first-order group contribution (GC) methods, the designed molecules should have no more than 10 groups and at most two functional groups (Giovanoglou et al., 2003). In order to rationalize the assumption of non-volatility of the solvent, the lower bound of solvent boiling point is set to 370 K. Whereas these molecular complexity and solvent property constraints are handled as inequality constraints in deterministic solvent design, in this work each molecule generated by the GA is tested against these constraints and those violating the constraints are immediately discarded. Solvent molecules are continuously created and tested until a pre-specified number (i.e., the population size of a generation) of molecules are obtained.

In order to obtain high-quality solutions in a reasonable computational time, GA parameters should be properly selected. The selection of GA parameters is task-specific. The population size and maximum number of generations are normally selected according to the size of the design space. In this case study, the maximum number of generations was set to 30 and each generation includes 11 molecule individuals. Increasing the probabilities of genetic operations facilitates the generation of new individuals during the evolution and therefore increases the possibility of finding high-quality solutions. However, this action can slow down the

convergence of the GA computation. Considering the tradeoff between the solution quality and computational cost, the probabilities of performing mutation, crossover, insertion, and deletion operations were set to 0.3, 0.3, 0.6, and 0.6, respectively. The probabilities of the insertion and deletion operations were set higher than those of mutation and crossover in order to favor the generation of molecules with different sizes. In order to ensure that the best individual in a generation does not deteriorate with the increasing of generation, the best molecule in one generation is directly passed into the next generation without any modifications. In GAs, each individual solution is evaluated and assigned with a fitness value that tells how desirable the solution is. In this case study, the fitness function is defined as the total annual cost (TAC) of the AD process.

Physical property models

Pure component properties

Physical properties of solvent and acetone are estimated by first-order GC models where the required information is the number and type of structural groups present in the molecule. There are a few GC methods (Constantinou and Gani, 1994; Marrero and Gani, 2001; Scilipoti et al., 2014) for predicting pure component physical properties. The most recently developed models reported in Scilipoti et al. (2014) are considered to have the overall highest prediction accuracy and therefore are mainly employed in this work.

Critical pressure (Scilipoti et al., 2014)

$$P_c = \frac{\sum_j a_j n_j DMW_j}{\sum_j a_j n_j DP_j + 0.34 \sum_j \frac{0^2}{\emptyset}} \quad (1)$$

Critical volume (Scilipoti et al., 2014)

$$V_c = 3.3127 \sum_j \frac{a_j}{V_j} n_j D V_j^{0.9542} \quad (2)$$

Molar volume at the normal boiling temperature (Scilipoti et al., 2014)

$$V = 0.285 V_c^{1.048} \quad (3)$$

Ratio of normal boiling temperature to critical temperature (Scilipoti et al., 2014)

$$\frac{T_b}{T_c} = 1 - \frac{\sum_j \frac{a_j}{V_j} n_j D M W_j^{0.0268}}{1 + \sum_j \frac{a_j}{V_j} n_j D T_j} + \frac{\sum_j \frac{a_j}{V_j} n_j D A_j}{0.122 + 0.1219 \sum_j \frac{a_j}{V_j} n_j D M W_j^{1.3769}} \quad (4)$$

Normal boiling temperature (Scilipoti et al., 2014)

$$T_b = V_c P_c \sum_j \frac{a_j}{V_j} n_j D \left[0.002876 \ln P_c \frac{T_b}{T_c} + 0.02603 \sum_j \frac{a_j}{V_j} n_j D \right] \quad (5)$$

Vapor pressure estimated from the Clapeyron equation

$$P^{sat}(T) = P_c \exp \left[\sum_j \frac{a_j}{V_j} n_j D \left(\frac{T_c}{T} - \frac{T_b}{T_c} \right) \ln \frac{P_c}{1.01325} \right] \quad (6)$$

Viscosity (Scilipoti et al., 2014)

$$m = \frac{\sum_j \frac{a_j}{V_j} n_j D m_j}{(P^{sat})^{0.2 + \sum_j \frac{a_j}{V_j} n_j D V_j}} \quad (7)$$

Heat capacity (Rihani and Doraiswamy, 1965)

$$C_p = \sum_j \frac{a_j}{V_j} n_j a_j + \sum_j \frac{a_j}{V_j} n_j b_j T + \sum_j \frac{a_j}{V_j} n_j c_j T^2 + \sum_j \frac{a_j}{V_j} n_j d_j T^3 \quad (8)$$

Normal melting temperature (Constantinou and Gani, 1994)

$$T_m = 102.425 \ln \left(\frac{\sum_j \frac{a_j}{e_j} n_j t m_j}{\sum_j \frac{a_j}{e_j} n_j t m_j} \right) \quad (9)$$

Mixture properties

Infinite dilution diffusion coefficient (Poling et al., 2001)

$$D_{AS}^0 = \frac{7.4 \cdot 10^{-8} (f_S MW_S)^{0.5} T}{m_S V_A^{0.6}} \quad (10)$$

The equation for calculating D_{SA}^0 can be derived by simply exchanging the indices of A and S in Eq. (10).

Diffusion coefficient of the acetone and solvent mixture at arbitrary compositions

$$D_{AS} = (D_{AS}^0 x_S + D_{SA}^0 x_A) \quad (11)$$

Liquid-phase mass transfer coefficient in packed columns (Cussler, 2009)

$$\frac{k_L d_p}{D_{AS}} = 25 \left(\frac{a_L v_0}{e n_{AS}} \right)^{0.45} \left(\frac{a_{AS}}{e D_{AS}} \right)^{0.5} \quad (12)$$

where:

$$n_{AS} = \frac{m_{AS}}{r_{AS}} \quad (13)$$

$$\ln m_{AS} = x_A \ln m_A + x_S \ln m_S \quad (14)$$

$$r_{AS} = x_A (MW_A / V_A) + x_S (MW_S / V_S) \quad (15)$$

Superficial flow velocity

$$v_0 = \frac{\sum_{i=A,S} N_i^L \cdot \sum_{i=A,S} x_i V_i}{\frac{1}{4} \rho d^2} \quad (16)$$

The overall mass transfer coefficient is calculated from:

$$\frac{1}{K_y a} = \frac{1}{k_y a} + \frac{m_A}{k_x a} \quad (17)$$

where

$$k_x = k_L / \sum_{i=A,S} \dot{a} x_i V_i \quad (18)$$

K_y values in the absorption and desorption columns are determined individually according to Eqs. (10) – (18).

Process and cost models

Mass balances in the absorption and desorption columns

$$0 = \frac{dN_A^L}{dz^{ab}} - \frac{dN_A^G}{dz^{ab}} \quad (19)$$

$$0 = \frac{dN_A^L}{dz^{de}} - \frac{dN_A^G}{dz^{de}} \quad (20)$$

Inert gas and nonvolatile solvent assumptions lead to:

$$N_S^G = N_{air}^L = 0 \quad (21)$$

$$\frac{dN_{air}^G}{dz^{ab}} = \frac{dN_S^L}{dz^{ab}} = 0 \quad (22)$$

$$\frac{dN_{air}^G}{dz^{de}} = \frac{dN_S^L}{dz^{de}} = 0 \quad (23)$$

The flow rates of acetone in the gas and liquid phases are expressed by:

$$N_A^G = \frac{N_{air}^G y_A}{1 - y_A} \quad (24)$$

$$N_A^L = \frac{N_S^L x_A}{1 - x_A} \quad (25)$$

Eqs. (21), (24), and (25) holds for the entire columns including inlet and outlet streams.

Connectivity constraints for the recycle streams

$$N_i^L(ab, out) = N_i^L(de, in) \quad (i = A, S) \quad (26)$$

$$N_i^L(ab, in) = N_i^L(de, out) \quad (i = A, S) \quad (27)$$

Process specifications

$$\begin{aligned} N_{air}^G(ab, in) &= 10 \text{ mol / s} \\ y_A(ab, in) &= 0.03 \\ y_A(ab, out) &= 0.0015 \\ N_S^L(ab, in) &= 4 \text{ mol / s} \\ x_A(ab, in) &= 0.005 \\ N_{air}^G(de, in) &= 5 \text{ mol / s} \\ y_A(de, in) &= 0 \end{aligned} \quad (28)$$

Eq. (29) is used to determine the packing heights of the absorption and desorption columns (i.e., Z^{ab} and Z^{de}) with the vapor-liquid equilibrium condition stated in Eq. (30). Finite difference methods (FDM) were used to solve the differential equations. Specifically, the entire solute-concentration range (from $y_{A,in}$ to $y_{A,out}$) was discretized into 30 regions and for each region, the differential equations were discretized and solved as algebraic equations.

$$- \frac{d}{dz} \left(\frac{N_{air}^G y_A}{1 - y_A} \right) = K_y a (y_A - y_A^{eq}) \rho d^2 \quad (29)$$

$$y_A^{eq} = m_A x_A = \frac{g_A(T) P_A^{sat}(T)}{P_{tot}} x_A \quad (30)$$

The temperature and solvent dependent activity coefficient of acetone γ_A is calculated by the

reformulated UNIFAC model (Buxton et al., 1999). In the model, the activity coefficient of component i is given by:

$$\ln(g_i) = \ln(g_i^C) + \ln(g_i^R) \quad (31)$$

Combinatorial part

$$\begin{aligned} \ln(g_i^C) = & \ln(R_i) - \ln \frac{\sum_{ii} \frac{R_i}{R_i} x_{ii} R_{ii}}{\sum_{ii} R_{ii}} - 5Q_i \ln(R_i) - 5Q_i \ln \frac{\sum_{ii} \frac{Q_i}{Q_i} x_{ii} Q_{ii}}{\sum_{ii} Q_{ii}} \\ & + 5Q_i \ln(Q_i) + 5Q_i \ln \frac{\sum_{ii} \frac{R_i}{R_i} x_{ii} R_{ii}}{\sum_{ii} R_{ii}} - 5Q_i + 1 + C_i \end{aligned} \quad (32)$$

$$C_i = \frac{\sum_{ii} \frac{R_i}{R_i} x_{ii} Q_{ii}}{\sum_{ii} R_{ii}} - 1 \frac{R_i}{\sum_{ii} R_{ii}} \quad (33)$$

Residual part

$$\ln(g_i^R) = R1_i - R2_i \quad (34)$$

$$R1_i = Q_i - \sum_j n_{i,j} q_j \ln(R3_j) + Q_i \ln \frac{\sum_{ii} \frac{Q_i}{Q_i} x_{ii} Q_{ii}}{\sum_{ii} Q_{ii}} - \sum_j \frac{R4_{i,j}}{R3_j} \quad (35)$$

$$R2_i = Q_i \ln(Q_i) - \sum_j n_{i,j} q_j \ln \frac{\sum_{mj} \frac{Q_i}{Q_i} q_{mj} n_{i,mj} y_{mj,j}}{\sum_{mj} Q_{mj}} \quad (36)$$

$$R3_j = \sum_{mj} q_{mj} \sum_{ii} n_{ii,mj} x_{ii} y_{mj,j} \quad (37)$$

$$R4_{i,j} = \sum_{mj} n_{i,mj} q_{mj} q_j \sum_{ii} n_{ii,j} x_{ii} y_{mj,j} \quad (38)$$

Van der Waals volume R_i and surface area Q_i for component i

$$R_i = \sum_j n_{i,j} r_j \quad (39)$$

$$Q_i = \sum_j n_{i,j} q_j \quad (40)$$

Interaction parameters between group mj and group j

$$y_{mj,j} = \exp(-a_{mj,j}/T) \quad (41)$$

In the UNIFAC model, $n_{i,j}$ represents the number of group j in component i . In this case study, there are only two components (acetone and the solvent) present in the liquid phase. $n_{A,j}$ is known for the acetone molecule and $n_{S,j}$ is the molecular variable of solvent to be optimized.

In order to ensure that the designed solvent is liquid in both absorption and desorption columns, the following inequality constraints are defined.

$$T_{m,S} < T^{ab}, T_{m,S} < T^{de} \quad (42)$$

$$T_{b,S} > T^{ab}, T_{b,S} > T^{de} \quad (43)$$

The TAC of the process includes annual utility cost (UC) and annual capital investment (CI).

Cooling duty:

$$Q_G^{Cooling} = N_A^G(ab, in) \int_{T^{am}}^{T^{ab}} C_{p,A} dT + N_{air}^G(ab, in) C_{p,air} (T^{ab} - T^{am}) < 0 \quad (44)$$

$$Q_L^{Cooling} = \sum_{i=A,S} \sum_e N_i^L(de, out) \int_{T^{de}}^{T^{ab}} C_{p,i} dT < 0 \quad (45)$$

$$Q_{total}^{Cooling} = Q_G^{Cooling} + Q_L^{Cooling} \quad (46)$$

Heating duty:

$$Q_G^{Heating} = N_A^G(de, in) \int_{T^{am}}^{T^{de}} C_{p,A} dT + N_{air}^G(de, in) C_{p,air} (T^{de} - T^{am}) > 0 \quad (47)$$

$$Q_L^{Heating} = \sum_{i=A,S} \sum_e N_i^L(ab, out) \int_{T^{ab}}^{T^{de}} C_{p,i} dT > 0 \quad (48)$$

$$Q_{total}^{Heating} = Q_G^{Heating} + Q_L^{Heating} \quad (49)$$

The utility cost majorly consists of refrigeration electricity cost UC^C and heating steam cost UC^H (Towler and Sinnott, 2012).

$$UC_G^C = 330 \cdot 24 \cdot 60 \cdot 60 \cdot \left| Q_G^{Cooling} \right| \cdot \frac{\beta_{cool} T^{am} - T^{ab}}{Z \cdot T^{ab}} \cdot y_{EL} \quad (50)$$

$$UC_L^C = 330 \cdot 24 \cdot 60 \cdot 60 \cdot \left| Q_L^{Cooling} \right| \cdot \frac{\beta_{cool} T^{de} - T^{ab}}{Z \cdot T^{ab}} \cdot y_{EL} \quad (51)$$

$$UC^C = UC_G^C + UC_L^C \quad (52)$$

$$UC^H = 330 \cdot 24 \cdot 60 \cdot 60 \cdot Q_{total}^{Heating} \cdot y_{HS} / DH_{vap}^{water} \quad (53)$$

Total annual utility cost

$$UC^{total} = UC^C + UC^H \quad (54)$$

Annual capital investment of the absorption and desorption columns (Loh et al., 2002)

$$CI^{ab} = (C_{shell}^{ab} + C_{packing}^{ab}) / PBT \quad (55)$$

$$C_{shell}^{ab} = y_{s0} + y_s \cdot Z^{ab} \quad (56)$$

$$C_{packing}^{ab} = \frac{\alpha}{\epsilon^4} \rho d^2 \cdot Z^{ab} \cdot y_p \quad (57)$$

$$CI^{de} = (C_{shell}^{de} + C_{packing}^{de}) / PBT \quad (58)$$

$$C_{shell}^{de} = y_{s0} + y_s \cdot Z^{de} \quad (59)$$

$$C_{packing}^{de} = \frac{\alpha}{\epsilon^4} \rho d^2 \cdot Z^{de} \cdot y_p \quad (60)$$

Total annual capital investment

$$CI^{total} = CI^{ab} + CI^{de} \quad (61)$$

Finally,

$$TAC = CI^{total} + UC^{total} \quad (62)$$

Results and discussion

The GA was encoded in C and run in Microsoft Visual C++ 2005 (Kruglinski et al., 1998). The nonlinear process optimization problem was solved by use of CONOPT (Drud, 1994) in the GAMS modelling environment (Rosenthal, 2006). In order to evaluate the performance and robustness of the hybrid algorithm, 10 optimization runs were consecutively performed with the results summarized in Table 1. The CPU time for a typical run is about 20 minutes. 9 of the 10 runs result in the same optimum (Solution 1, Table 1) with CH_3COOH as the best solvent. In the remaining one run, another solution (Solution 2, Table 1) with the solvent $\text{CH}_3\text{CH}_2\text{COOH}$ was found. In order to more clearly recognize the reliability of the method, all the solvents (the total number is 1696) in the design space were enumerated and evaluated. Solution 1 and Solution 2 have been proven to be the best and the second best solutions, respectively. This high probability of finding the best solutions from random initializations demonstrates the high reliability and robustness of the method.

Table 1: Results of 10 consecutive optimization runs

	Solution 1	Solution 2
Optimal solvent	CH_3COOH	$\text{CH}_3\text{CH}_2\text{COOH}$
T^{ab} (K)	303.54	304.17
T^{de} (K)	325.38	326.28
Z^{ab} (cm)	199.89	241.12
Z^{de} (cm)	311.93	382.62
UC^{C} (\$/year)	261.28	362.78
UC^{H} (\$/year)	2264.75	2846.80
CI^{total} (\$/year)	1728.36	1827.49
TAC (\$/year)	4254.39	5037.07

For better comparison, the problem has also been handled by deterministic MINLP algorithms, including the global solver BARON and local solvers DICOPT and SBB (Brooke et al., 1998). For each solver, the problem has been solved multiple times starting from different initial estimates. The result shows that the problem cannot be successfully settled by BARON within a reasonable amount of time. The performances of DICOPT and SBB are very similar. Both of them rely much on good initializations. Over 80% of the runs failed to find any feasible solutions and only about 5.0% of them succeeded to identify the best known solution (Solution 1, Table 1). More details on the solved MINLP problem are provided in Table 2.

Table 2: Computational details of the solved MINLP problem (CONOPT for NLP subproblems and CPLEX for MILP subproblems)

Solver	DICOPT
Number of continuous variables	1198
Number of discrete variables	14
Number of equations	1190
Number of NLPs	6
NLP CPU time / MILP CPU time (s)	231.2 / 4.0
Total CPU time (s)	235.8

As described before, the result of each calculation including the optimal solution and the corresponding process TAC value is recorded in a database. A list of top solutions can be obtained by sorting all solutions in the database according to their TAC values after an entire computation is completed. For a random optimization run, the top five solutions are obtained and presented in Table 3. Taking into account the error in the employed property predictive models, the generation of multiple promising solutions is desirable since the solutions can be further evaluated by experiments before making a final selection.

Table 3: Top five solutions obtained for the integrated solvent and process design problem

Ranking	Optimal solvent	T ^{ab} (K)	T ^{de} (K)	TAC (\$/year)
1		Solution 1, Table 1		
2		Solution 2, Table 1		
3	CH ₃ CH(CHO) ₂	295.51	319.28	5291.23
4	CHO(CH ₂) ₂ CHO	295.54	319.31	5394.91
5	CH ₃ CH(OH)CHO	282.72	307.30	5490.98

The evolution of the best solution in each generation is depicted in Figure 5. The computation experiences three turning points before it converges to the best known solution at the 15th generation. All the corresponding solutions including the solvent molecule as well as the absorption and desorption temperatures are shown in the figure.

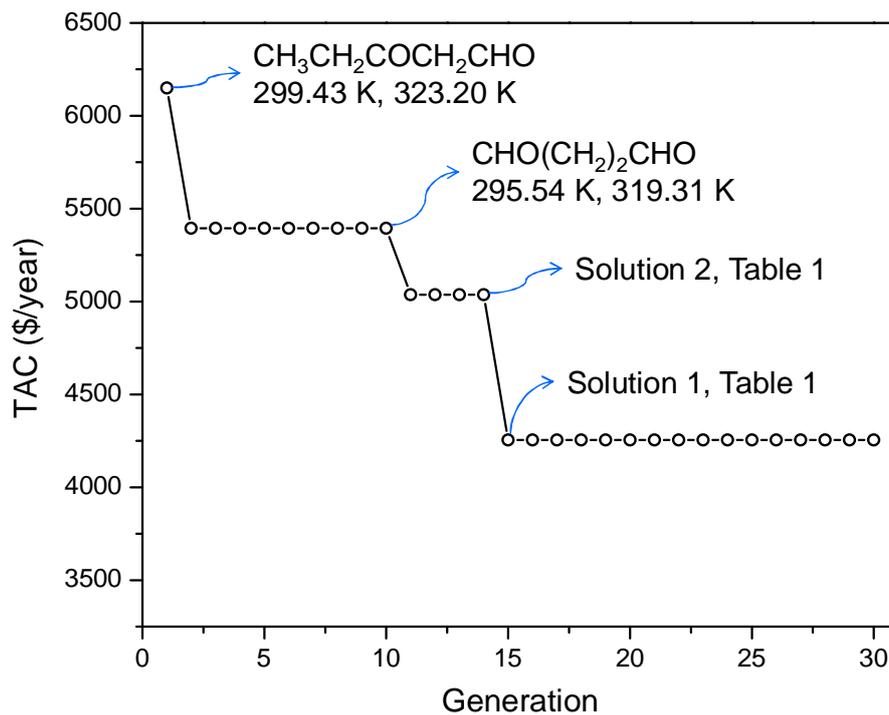


Figure 5: The evolution of the best solution in each generation

Conclusion

This paper proposes a hybrid stochastic-deterministic optimization method for integrated solvent and process design. The efficiency and robustness of the method has been demonstrated on a gas absorption process where solvent molecular structures and process operating conditions were simultaneously optimized to maximize the overall economic performance of the process. This approach is applicable beyond the integrated solvent and process design problem addressed in this work, and can be recommended for use in other large-scale MINLP optimizations.

The reliability of solvent design depends on the accuracy of the predictive models used for estimating solvent physical properties. One commonly used approach to improve the prediction accuracy is to use higher-level property models, such as higher-order GC models (Marrero and Gani, 2001). However, it should be noted that the computational complexity is significantly increased when higher-order groups are included due to the much larger number of design variables and equations. It is recommended that at the beginning of solvent and process design, complex predictive models which largely complicate the computation without significantly improving the reliability of results should be avoided. Instead of providing only one optimal solution, the proposed method facilitates the generation of multiple promising solutions, which can then be more rigorously evaluated using higher-level property models or experimental properties, if available.

Nomenclature

List of parameters

Symbol	Parameter description	Value or source
ΔMW_j	contribution of group j to MW	group-specific in g/mol
ΔP_j	contribution of group j to P_c	Scilipoti et al. (2014)
ΔV_j	contribution of group j to V_c	Scilipoti et al. (2014)
ΔT_j	contribution of group j to T_b/T_c	Scilipoti et al. (2014)
ΔA_j	association contribution of group j	Scilipoti et al. (2014)
$\Delta \mu_j, \Delta N_j$	contribution of group j to μ	Scilipoti et al. (2014)
a_j, b_j, c_j, d_j	contribution of group j to C_p	Rihani and Doraiswamy (1965)
tm_j	contribution of group j to T_m	Constantinou and Gani (1994)
ϕ	dimensionless association factor	1.0
d_p	packing size of 2-inch saddles	5.08 cm
d	column diameter	30.0 cm
k_y	local mass transfer coefficient in gas phases	1.70×10^7 mol/(cm ² ×s)
a	surface area per volume for the saddle packing	1.05 cm ² /cm ³
P_{tot}	operating pressure of column	1 atm
$a_{mj,j}$	UNIFAC interaction parameters between groups m_j and j	Gmehling et al. (1982)
r_j	UNIFAC volume parameter of group j	Gmehling et al. (1982)
R_i	Van der Waals volume of component i	N/A
q_j	UNIFAC surface area parameter of group j	Gmehling et al. (1982)
Q_i	Van der Waals surface area of component i	N/A
$R1_i$	modified residual activity coefficient of component i in the mixture	N/A
$R2_i$	modified residual activity coefficient of component i in a reference solution containing only molecules of component i	N/A
$C_{p,air}$	heat capacity of air	29.19 J/(mol×K)
T^{am}	ambient temperature	305.15 K
ζ	cooling cycle efficiency	0.80
ψ_{EL}	electricity price	0.06 \$/kwh
ψ_{HS}	heating steam price	2.57×10^{-4} \$/mol

ΔH_{vap}^{water}	enthalpy of vaporization of water at 250 °C	1715.8 kJ/kg
PBT	payback time	8 years
ψ_{SO}	parameter of linearized shell cost model	5100 \$
ψ_S	parameter of linearized shell cost model	6.56 \$/cm
ψ_P	packing price	7.42×10^{-4} \$/cm ³

List of variables

Symbol	Variable description	Unit
n_j	number of group j in the molecule	N/A
P_c	critical pressure	atm
V_c	critical volume	cm ³ /mol
V	molar volume at the normal boiling temperature	cm ³ /mol
T_b	normal boiling temperature	K
T_c	critical temperature	K
P^{sat}	saturated vapor pressure	atm
T	temperature	K
μ	viscosity	mPa×s
C_p	heat capacity	cal/(mol×K)
T_m	normal melting temperature	K
MW	molecular weight	g/mol
D_{AS}^o	infinite dilution diffusion coefficient of A in S	cm ² /s
D_{SA}^o	infinite dilution diffusion coefficient of S in A	cm ² /s
D_{AS}	diffusion coefficient of the A and S mixture	cm ² /s
x	mole fraction in the liquid phase	N/A
y	mole fraction in the gas phase	N/A
k_L	liquid-film mass transfer coefficient	cm/s
ν	kinematic viscosity	$0.01 \times \text{cm}^2/\text{s}$
ν_0	superficial flow velocity	cm/s
N	molar flow rate	mol/s
K_y	overall mass transfer coefficient	mol/(cm ² ×s)
m	vapor-liquid equilibrium constant	N/A
k_x	local mass transfer coefficient in liquid phases	mol/(cm ² ×s)
dz	differential packing height of column	cm
Z	packing height of column	cm

γ	activity coefficient	N/A
γ^C	combinatorial part for the activity coefficient	N/A
γ^R	residual part for the activity coefficient	N/A
TAC	total annual cost	\$/year
UC	annual utility cost	\$/year
CI	annual capital investment	\$/year
$Q^{Cooling}$	cooling duty	J/s
$Q^{Heating}$	heating duty	J/s
UC^C	annual cooling utility cost	\$/year
UC^H	annual heating utility cost	\$/year
C_{shell}	cost of column shell	\$
$C_{packing}$	cost of column packing	\$

Indices

ab	absorption column
de	desorption column
G	gas phase
L	liquid phase
A	acetone
S	solvent
j, mj	group index
i, ii	component index
in	inlet steams
out	outlet steams
eq	equilibrium composition

Acknowledgements

This work was conducted partly in cooperation with the Collaborative Research Center SFB/TRR 63 "Integrated Chemical Processes in Liquid Multiphase Systems (InPROMPT)". The financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged. Moreover, the authors acknowledge the support provided by the "International Max Planck Research School (IMPRS) for Advanced Methods in Process and System Engineering (Magdeburg)" for Teng Zhou.

References

- Affenzeller, M., Wagner, S., Winkler, S., Beham, A., 2009. Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications. Chapman and Hall: CRC Press.
- Athier, G., Floquet, P., Pibouleau, L., Domenech, S., 1997. Process optimization by simulated annealing and NLP procedures. Application to heat exchanger network synthesis. *Comput. Chem. Eng.* 21, S475-S480.
- Bardow, A., Steur, K., Gross, J., 2010. Continuous-molecular targeting for integrated solvent and process design. *Ind. Eng. Chem. Res.* 49, 2834-2840.
- Biegler, L.T., 2014. Recent advances in chemical process optimization. *Chemie Ingenieur Technik* 86, 943-952.
- Brooke, A., Kendrick, D., Meeraus, A., Raman, R., 1998. GAMS: the Solver Manuals. GAMS Development Corporation, Washington D.C.
- Burger, J., Papaioannou, V., Gopinath, S., Jackson, G., Galindo, A., Adjiman, C.S., 2015. A hierarchical method to integrated solvent and process design of physical CO₂ absorption using the SAFT- Mie approach. *AIChE J.* 61, 3249-3269.
- Buxton, A., Livingston, A.G., Pistikopoulos, E.N., 1999. Optimal design of solvent blends for environmental impact minimization. *AIChE J.* 45, 817-843.
- Chemangattuvalappil, N.G., Eljack, F.T., Solvason, C.C., Eden, M.R., 2009. A novel algorithm for molecular synthesis using enhanced property operators. *Comput. Chem. Eng.* 33, 636-643.
- Cheng, H.C., Wang, F.S., 2010. Computer-aided biocompatible solvent design for an integrated extractive fermentation–separation process. *Chem. Eng. J.* 162, 809-820.
- Constantinou, L., Gani, R., 1994. New group contribution method for estimating properties of pure compounds. *AIChE J.* 40, 1697-1710.
- Cussler, E.L., 2009. Diffusion: Mass Transfer in Fluid Systems. Cambridge University Press.
- Drud, A.S., 1994. CONOPT—a large-scale GRG code. *ORSA J. Comput.* 6, 207-216.
- Eden, M.R., Jorgensen, S.B., Gani, R., El-Halwagi, M.M., 2004. A novel framework for simultaneous separation process and product design. *Chem. Eng. Process.* 43, 595-608.
- Eljack, F.T., Eden, M.R., Kazantzi, V., Qin, X., El-Halwagi, M.M., 2007. Simultaneous process and molecular design—A property based approach. *AIChE J.* 53, 1232-1239.
- Folić, M., Adjiman, C.S., Pistikopoulos, E.N., 2007. Design of solvents for optimal reaction rate constants. *AIChE J.* 53, 1240-1256.
- Folić, M., Adjiman, C.S., Pistikopoulos, E.N., 2008. Computer-aided solvent design for reactions: maximizing product formation. *Ind. Eng. Chem. Res.* 47, 5190-5202.
- Gani, R., Nielsen, B., Fredenslund, A., 1991. A group contribution approach to computer-aided molecular design. *AIChE J.* 37, 1318-1332.

Giovanoglou, A., Barlatier, J., Adjiman, C.S., Pistikopoulos, E.N., Cordiner, J.L., 2003. Optimal solvent design for batch separation based on economic performance. *AIChE J.* 49, 3095-3109.

Gmehling, J., Rasmussen, P., Fredenslund, A., 1982. Vapor-liquid equilibria by UNIFAC group contribution. Revision and extension. 2. *Ind. Eng. Chem. Process Des. Dev.* 21, 118-127.

Gómez, J.M., Reneaume, J.M., Roques, M., Meyer, M., Meyer, X., 2006. A mixed integer nonlinear programming formulation for optimal design of a catalytic distillation column based on a generic nonequilibrium model. *Ind. Eng. Chem. Res.* 45, 1373-1388.

Hamad, A.A., El-Halwagi, M.M., 1998. Simultaneous synthesis of mass separating agents and interception networks. *Chem. Eng. Res. Des.* 76, 376-388.

Harper, P.M., Gani, R., 2000. A multi-step and multi-level approach for computer aided molecular design. *Comput. Chem. Eng.* 24, 677-683.

Herring, R.H., Eden, M.R., 2015. Evolutionary algorithm for de novo molecular design with multi-dimensional constraints. *Comput. Chem. Eng.* 83, 267-277.

Hostrup, M., Harper, P.M., Gani, R., 1999. Design of environmentally benign processes: integration of solvent design and separation process synthesis. *Comput. Chem. Eng.* 23, 1395-1414.

Karunanithi, A.T., Achenie, L.E.K., Gani, R., 2005. A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures. *Ind. Eng. Chem. Res.* 44, 4785-4797.

Karunanithi, A.T., Achenie, L.E.K., Gani, R., 2006. A computer-aided molecular design framework for crystallization solvent design. *Chem. Eng. Sci.* 61, 1247-1260.

Kim, K.J., Diwekar, U.M., 2002. Integrated solvent selection and recycling for continuous processes. *Ind. Eng. Chem. Res.* 41, 4479-4488.

Kossack, S., Kraemer, K., Gani, R., Marquardt, W., 2008. A systematic synthesis framework for extractive distillation processes. *Chem. Eng. Res. Des.* 86, 781-792.

Kruglinski, D.J., Wingo, S., Sheperd, G.W., 1998. *Programming Microsoft Visual C++*. Microsoft Press, Redmond.

Lima, R.M., Salcedo, R.L., Barbosa, D., 2006. SIMOP: Efficient reactive distillation optimization using stochastic optimizers. *Chem. Eng. Sci.* 61, 1718-1739.

Lipowski, A., Lipowska, D., 2012. Roulette-wheel selection via stochastic acceptance. *Physica A.* 391, 2193-2196.

Loh, H.P., Lyons, J., White, C.W., 2002. *Process Equipment Cost Estimation, Final Report*. National Energy Technology Lab., Pittsburgh.

Marcoulaki, E.C., Kokossis, A.C., 2000. On the development of novel chemicals using a systematic optimisation approach. Part II. Solvent design. *Chem. Eng. Sci.* 55, 2547-2561.

Marrero, J., Gani, R., 2001. Group-contribution based estimation of pure component properties. *Fluid Phase Equilib.* 183, 183-208.

Misener, R., Floudas, C.A., 2013. GloMIQO: Global mixed-integer quadratic optimizer. *J. Global Optim.* 57, 3-50.

Molina, D., Lozano, M., García-Martínez, C., Herrera, F., 2010. Memetic algorithms for continuous optimisation based on local search chains. *Evolut. Comput.* 18, 27-63.

Moscato, P., 1989. On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms. Caltech concurrent computation program, C3P Report, 826.

Ng, L.Y., Chong, F.K., Chemmangattuvalappil, N.G., 2015. Challenges and opportunities in computer-aided molecular design. *Comput. Chem. Eng.* 81, 115-129.

Odele, O., Macchietto, S., 1993. Computer aided molecular design: A novel method for optimal solvent selection. *Fluid Phase Equilib.* 82, 47-54.

Oyarzún, B., Bardow, A., Gross, J., 2011. Integration of process and solvent design towards a novel generation of CO₂ absorption capture systems. *Energy Procedia* 4, 282-290.

Papadopoulos, A.I., Linke, P., 2006a. Multiobjective molecular design for integrated process-solvent systems synthesis. *AIChE J.* 52, 1057-1070.

Papadopoulos, A.I., Linke, P., 2006b. Efficient integration of optimal solvent and process design using molecular clustering. *Chem. Eng. Sci.* 61, 6316-6336.

Papadopoulos, A.I., Linke, P., 2009. Integrated solvent and process selection for separation and reactive separation systems. *Chem. Eng. Process.* 48, 1047-1060.

Papadopoulos, A.I., Stijepovic, M., Linke, P., 2010. On the systematic design and selection of optimal working fluids for Organic Rankine Cycles. *Appl. Therm. Eng.* 30, 760-769.

Papadopoulos, A.I., Stijepovic, M., Linke, P., Seferlis, P., Voutetakis, S., 2013. Toward optimum working fluid mixtures for organic Rankine cycles using molecular design and sensitivity analysis. *Ind. Eng. Chem. Res.* 52, 12116-12133.

Pereira, F.E., Keskes, E., Galindo, A., Jackson, G., Adjiman, C.S., 2011. Integrated solvent and process design using a SAFT-VR thermodynamic description: High-pressure separation of carbon dioxide and methane. *Comput. Chem. Eng.* 35, 474-491.

Pistikopoulos, E.N., Georgiadis, M.C., Dua, V., Adjiman, C.S., Galindo, A., 2010. *Process Systems Engineering. Volume 6. Molecular Systems Engineering.* Wiley-VCH, Weinheim.

Pistikopoulos, E.N., Stefanis, S.K., 1998. Optimal solvent design for environmental impact minimization. *Comput. Chem. Eng.* 22, 717-733.

Poling, B.E., Prausnitz, J.M., O'Connell, J.P., 2001. *The Properties of Gases and Liquids.* McGraw-Hill, New York.

Pretel, E.J., López, P.A., Bottini, S.B., Brignole, E.A., 1994. Computer-aided molecular design of solvents for separation processes. *AIChE J.* 40, 1349-1360.

Rangaiah, G.P., 2010. *Stochastic Global Optimization: Techniques and Applications in Chemical Engineering (Vol. 2).* World Scientific.

Rihani, D.N., Doraiswamy, L.K., 1965. Estimation of heat capacity of organic compounds from group contributions. *Ind. Eng. Chem. Fund.* 4, 17-21.

Rogers, R.D., Seddon, K.R., 2003. Ionic liquids-solvents of the future?. *Science* 302, 792-793.

Rosenthal, R.E., 2006. *GAMS-A Users' Guide*. GAMS Development Corporation, Washington D.C.

Roughton, B.C., Christian, B., White, J., Camarda, K.V., Gani, R., 2012. Simultaneous design of ionic liquid entrainers and energy efficient azeotropic separation processes. *Comput. Chem. Eng.* 42, 248-262.

Sahinidis, N.V., 1996. BARON: A general purpose global optimization software package. *J. Global Optim.* 8, 201-205.

Samudra, A.P., Sahinidis, N.V., 2013. Optimization-based framework for computer-aided molecular design. *AIChE J.* 59, 3686-3701.

Scilipoti, J., Cismondi, M., Brignole, E.A., 2014. Prediction of physical properties for molecular design of solvents. *Fluid Phase Equilib.* 362, 74-80.

Sinha, M., Achenie, L.E.K., Ostrovsky, G.M., 1999. Environmentally benign solvent design by global optimization. *Comput. Chem. Eng.* 23, 1381-1394.

Skiborowski, M., Rautenberg, M., Marquardt, W., 2015. A hybrid evolutionary–deterministic optimization approach for conceptual design. *Ind. Eng. Chem. Res.* 54, 10054-10072.

Towler, G., Sinnott, R.K., 2012. *Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design*. Elsevier.

Urselmann, M., Barkmann, S., Sand, G., Engell, S., 2011a. A memetic algorithm for global optimization in chemical process synthesis problems. *IEEE T. Evolut. Comput.* 15, 659-683.

Urselmann, M., Barkmann, S., Sand, G., Engell, S., 2011b. Optimization-based design of reactive distillation columns using a memetic algorithm. *Comput. Chem. Eng.* 35, 787-805.

van Dyk, B., Nieuwoudt, I., 2000. Design of solvents for extractive distillation. *Ind. Eng. Chem. Res.* 39, 1423-1429.

Venkatasubramanian, V., Chan, K., Caruthers, J.M., 1994. Computer-aided molecular design using genetic algorithms. *Comput. Chem. Eng.* 18, 833-844.

Xu, W., Diwekar, U.M., 2005. Improved genetic algorithms for deterministic optimization and optimization under uncertainty. Part II. Solvent selection under uncertainty. *Ind. Eng. Chem. Res.* 44, 7138-7146.

Zhou, T., Lyu, Z., Qi, Z., Sundmacher, K., 2015b. Robust design of optimal solvents for chemical reactions—A combined experimental and computational strategy. *Chem. Eng. Sci.* 137, 613-625.

Zhou, T., McBride, K., Zhang, X., Qi, Z., Sundmacher, K., 2015a. Integrated solvent and process design exemplified for a Diels-Alder reaction. *AIChE J.* 61, 147-158.

Zhou, T., Wang, J., McBride, K., Sundmacher, K., 2015c. Optimal design of solvents for heterogeneous liquid phase reactions. *AIChE J.* Under review.