

# A feasibility-based algorithm for Computer Aided Molecular and Process Design of solvent-based separation systems

Smitha Gopinath<sup>a</sup>, Amparo Galindo<sup>a</sup>, George Jackson<sup>a</sup> and Claire S. Adjiman<sup>a\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom  
c.adjiman@imperial.ac.uk*

## Abstract

Computer-aided molecular and product design (CAMPD) can in principle be used to find simultaneously the optimal conditions in separation processes and the structure of the optimal solvents. In many cases, however, the solution of CAMPD problems is challenging. In this paper, we propose a solution approach for the CAMPD of solvent-based separation systems in which implicit constraints on phase behaviour in process models are used to test the feasibility of the process and solvent domains. The tests not only eliminate infeasible molecules from the search space but also infeasible combinations of solvent molecules and process conditions. The tests also provide bounds for the optimization of the process model (primal problem) for each solvent, facilitating numerical solution. This is demonstrated on a prototypical natural gas purification process.

**Keywords:** CAMPD, MINLP, solvent, CO<sub>2</sub> capture

## 1. Introduction

The computer-aided molecular and process design (CAMPD) problem is defined as the simultaneous optimization of a process and the molecules used in the process. CAMPD problems are large mixed-integer nonlinear problems (MINLP) that are challenging to solve for a number of reasons: the need to solve a large number of nonlinear equality constraints in process and property prediction models; the discontinuous nature of the feasible space, where optimization starting from an infeasible point may lead to numerical failure; convergence to poor local solutions due to the non-convexity of the problem.

Several approaches have been proposed to overcome these challenges. One approach is to reformulate it into a nonlinear continuous optimization problem (Pereira et al., 2011; Lampe et al., 2015). Hierarchical optimization has been used to generate initial guesses for the solution of the full CAMPD problem (Burger et al., 2015). Screening-based methods, and that reduce the molecular search space, based on a generate-and-test strategy, that employ physical property targets have been used in several studies (Hostrup et al., 1999; Karunanithi et al., 2005; Eljack et al., 2008). The determination of targets by using property “clusters” has been investigated (Eljack et al., 2008).

Buxton et al. (1999) proposed a screening approach embedded in an MINLP optimizer and applied it to the design of a solvent for a separation unit with fixed operating conditions. In their approach, tests based on property targets, the ability to initialize model subproblems and the feasibility of mass transfer were used prior to solving the primal problem. We adopt the idea of embedding tests in an optimizer and propose novel tests that do not require the user to specify arbitrary property

targets for the solvent or mixtures in the process. The tests are based on thermodynamic insights that exploit implicit constraints on phase behaviour. Examples of such constraints include the presence of two phases at vapour-liquid equilibrium or the stability of the liquid solvent at relevant process conditions. Implicit constraints are often discontinuous and their violation can lead to numerical failure. The tests developed in our study lead to a reduction in the domain of (unknown) process *and* molecular variables, thereby providing bounds for the optimization of the process model that lie within the feasible region. While the tests developed are general and may be adapted to different separation systems, they are introduced with a focus on the design of an absorption system.

## 2. Proposed tests

The CAMPD problem is given by

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{n}} \quad & f(\mathbf{u}, \mathbf{n}) && \text{(P)} \\ \text{s.t.} \quad & \mathbf{g}_p(\mathbf{u}, \mathbf{n}) = 0 \\ & \mathbf{g}(\mathbf{u}, \mathbf{n}) \leq 0 \\ & \mathbf{C}\mathbf{n} \leq \mathbf{e} \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned}$$

where  $\mathbf{u} \in \mathbf{X} \subset \mathbb{R}^c$  is a vector of continuous variables and  $\mathbf{n} \in \mathbf{N} \subset (\mathbb{N} \cup \{0\})^q$  is a vector of integer variables that represent the number of groups of each type in the molecule.  $f: \mathbf{X} \times \mathbf{N} \rightarrow \mathbb{R}$  is the process objective.  $\mathbf{g}_p$  is a set of process and property constraints,  $\mathbf{g}$  represents process constraints and  $\mathbf{C}\mathbf{n} \leq \mathbf{e}$  represents molecular feasibility constraints and bounds on the vector  $\mathbf{n}$ .  $\mathbf{u}^L$  and  $\mathbf{u}^U$  are lower and upper bounds on the continuous variables, respectively.

### 2.1. Test 0

An implied constraint on an absorption model is that the feed remains in the vapour phase. Test 0 identifies conditions of temperature and pressure under which the feed is in the desired phase. First, an iterative approach is applied to find the cricondentherm  $T_F^{L0}$ , the highest two-phase coexistence temperature for the feed of composition  $\mathbf{y}_F$ . The dew pressure is found iteratively at every temperature  $T$  ( $T$  is increased by a user-specified step size  $h$  at the end of each iteration) until no two-phase coexistence pressure is found, to yield  $T_F^{L0}$ .

Once  $T_F^{L0}$  has been computed, a lower bound on the pressure domain is determined. The pressure is assumed to be regulated using an isenthalpic valve. A reduction in pressure, for gases with a positive Joule-Thompson coefficient such as  $\text{CO}_2$  and methane, results in a reduction in temperature. Thus, one may find the lowest pressure  $P^{L0}$  that is achievable by isenthalpic expansion of the gas, without the temperature dropping below  $T_F^{L0}$ .

### 2.2. Test 1

An implied constraint on an absorption process is that the solvent is a liquid at relevant process conditions. Test 1 is used to identify whether the solvent (defined by vector  $\mathbf{n}^k$ ) is a liquid in the range of user-specified storage and handling temperatures  $[T_{s1}, T_{s2}]$  and pressures, and if it is safe to handle and feasible to transport.  $T_{melt}$ ,  $T_{boil}$ ,  $T_{flash}$  and  $v$  are the melting point, boiling point, flash point and viscosity of the proposed solvent, respectively.  $v^b$  is a user specified upper bound on viscosity. Similar to the property-based screening of Harper et al. (1999), test 1 consists of the following property tests:

$$T_{s2} - T_{boil} \leq 0, T_{melt} - T_{s1} \leq 0, T_{s2} - T_{flash} \leq 0, v(T_{s1}) - v^b \leq 0 \quad \text{(P1)}$$

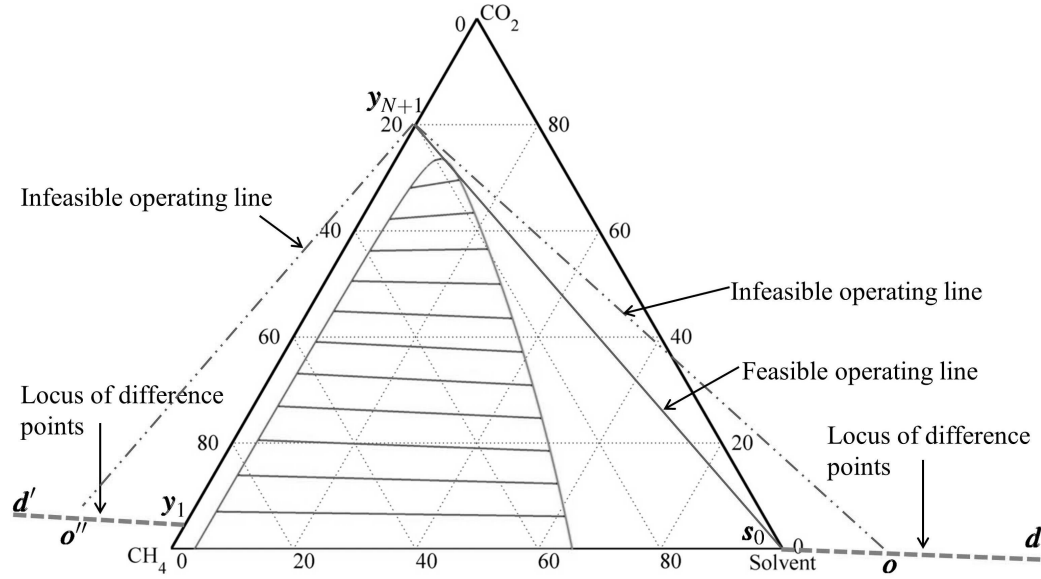


Figure 1: A phase diagram for CO<sub>2</sub>-methane-solvent (propyl-methyl ether) at  $T_N = 304.4$  K,  $P_{abs} = 9.897$  MPa

### 2.3. Test 2

An implicit constraint in an absorber model is that the solvent and feed can mix to form two phases. Test 2 is designed to identify if the solvent can form a two-phase mixture with the feed and to determine the range of operating conditions where it may do so. This test is based on the presence of a feasible operating line for separation. The operating line is calculated using the concept of difference point,  $\Delta = V_{i+1} - L_i$ , given by Hunter and Nash (1934) and Henley et al. (2011) for the design of extraction systems. Here, vapour and liquid streams leaving any stage  $i$  are represented by flowrates  $V_i$  and  $L_i$  and compositions  $y_i$  and  $x_i$ , respectively. The composition at the difference point is represented by  $\mathbf{o}$  in Figure 1.

The gas entering the absorber has a flow rate  $V_{N+1}$  and composition  $y_{N+1}$ . The solvent entering the absorber (which is assumed to be pure solvent in this feasibility test) has a flowrate  $L_0$  and composition  $s_0$ . As  $\Delta = V_{N+1} - L_N$ , an operating line must pass through the feed  $y_{N+1}$ ,  $x_N$  (a point on the two-phase boundary) and the difference point  $\mathbf{o}$ . From  $\Delta = V_1 - L_0$ , the locus of difference points lies on the line  $\overrightarrow{y_1 s_0}$ , excluding the segment  $\overrightarrow{y_1 s_0}$  (as both  $V_1$  and  $L_0$  are non-negative).

This test may be applied to systems for which the solubility of the solute in the solvent increases with an increase in pressure. At a pressure where the mixture exhibits a critical point, such that the solute and solvent are completely miscible, a further increase in pressure reduces the maximum concentration of the solute on the two-phase boundary. When the highest concentration of solute on the liquid phase boundary is lower than the concentration of the solute in the feed, a difference point that lies on ray  $\overrightarrow{y_1 d'}$  is infeasible. This is because an operating line  $\overrightarrow{y_{N+1} o''}$  cannot intersect the two-phase region, if  $o''$  is a point on  $\overrightarrow{y_1 d'}$ . Hence, the difference point can lie only on the ray  $\overrightarrow{s_0 d}$ . By visualization, if there exists an operating line with any  $\mathbf{o}$  on the ray  $\overrightarrow{s_0 d}$  that intersects the two-phase region, then the operating line with  $\mathbf{o} = s_0$  also cuts the two-phase region. If the line segment  $\overrightarrow{y_{N+1} s_0}$  does not intersect the two-phase region, then no other operating line is feasible. Figure 1 represents the operating line and difference points at the maximum value of feasible

pressure  $P^{U1}$ , where  $\overline{\mathbf{y}_{N+1}\mathbf{s}_0}$  is the only feasible operating line.

This is formalised in problem (P2), where the existence of a feasible operating line is checked by testing if the difference line  $\overline{\mathbf{y}_{N+1}\mathbf{s}_0}$ , intersects any point  $\mathbf{x}_N$  that is in equilibrium with  $\mathbf{y}_N$ . An upper bound  $P^{U1}$  is identified on the pressure of the separation unit below which the condition of feasibility is met.

$$\begin{aligned}
 P^{U1} = & \max_{P_{abs}, T_N} P_{abs} & (P2) \\
 \text{s.t.} & \frac{\mathbf{y}_{N+1}(1) - \mathbf{x}_N(1)}{\mathbf{y}_{N+1}(2) - \mathbf{x}_N(2)} = \frac{\mathbf{y}_{N+1}(1) - \mathbf{s}_0(1)}{\mathbf{y}_{N+1}(2) - \mathbf{s}_0(2)} \\
 & \mu_i(\mathbf{y}_N, T_N, P_{abs}) = \mu_i(\mathbf{x}_N, T_N, P_{abs}) \quad \forall i \in NC \\
 & \sum_{i=1}^{NC} x_i = 1 \\
 & \sum_{i=1}^{NC} y_i = 1 \\
 & \|\mathbf{y}_N - \mathbf{x}_N\|_2 > 0 \\
 & P^{L0} \leq P_{abs} \leq P^{U0} \\
 & T^{L0} \leq T_N \leq T^{U0} \\
 & T_{melt} + 10 \leq T_N \leq \min(T_{flash} - 10, T_{boil} - 10)
 \end{aligned}$$

### 3. Proposed algorithm

The tests are embedded within an outer approximation framework (Duran and Grossmann, 1986; Fletcher and Leyffer, 1994). To account for linearizations that are not strict underestimators of the nonconvex feasible space, global convexity tests (Kravanja and Grossmann, 1994) have been implemented. The main algorithmic steps are as follows:

- I. Set  $T_F$ ,  $\mathbf{y}_F$ ,  $P^{U0}$ ,  $T^{U0}$ ,  $\mathbf{n}^1$ , initial solvent structure, and  $\varepsilon$ , convergence tolerance.
- II. Evaluate Test 0 to find  $T_F^{L0}$  and  $P^{L0}$ .
- III. Set  $f^{UBD} = \infty$  and  $f^{LBD} = 0$ . Set iteration counter  $k = 1$ ,  $F^0 = \emptyset$ ,  $A^0 = \emptyset$ ,  $IC^0 = \emptyset$  and  $\mathbf{y}^0 = \emptyset$
- IV. While ( $f^{UBD} - f^{LBD} > \varepsilon$ )
  1. Evaluate Test 1. If test passed, go to step IV 2, else go to step IV 5.
  2. Evaluate Test 2. If problem (P2) is infeasible, go to step IV 5, else go to step IV 3.
  3. Set  $P_{abs}^U = P^{U1}$ . Evaluate the primal problem (P3) (Table 1).
  4. If problem (P3) is feasible, set  $f^k$  equal to the objective and if  $f^k < f^{UBD}$ , set  $f^{UBD} = f^k$  and  $\mathbf{n}^* = \mathbf{n}^k$ . If (P3) is feasible set  $F^k = F^{k-1} \cup \{k\}$  and go to step IV 6, else go to step IV 5.
  5. Add an integer cut to (P4) (Table 1) to eliminate  $\mathbf{n}^k$ . Add  $IC^k = IC^{k-1} \cup \{k\}$ . Compute  $a^k = \sum_t (b^{t-1} n_t^k)$ , where  $b$  is set such that  $b$  is the maximum value that any  $n_i$  can assume.
  6. Solve (P4) (Table 1). If feasible, the objective is  $\eta^k$ , set  $f^{LBD} = \eta^k$  and go to step IV 7. Else, stop.
  7. Set  $k = k + 1$  and go to step IV 1.

The set  $A^k$  of the master problem (P4) contains linearizations of the active constraints in problems (P3) and (P1), as well as constraints violated in (P1). Hence an integer cut is not required to avoid cycling of solvents that fail test 1.

#### 3.1. Algorithm Implementation

The overall algorithm is implemented in C++. The problems (P0), (P1), (P2), (P3) and (P4) are each launched, as necessary, by C++. Problems (P0), (P1), (P2) and (P3) have been implemented in gPROMS (Process Systems Enterprise, 1997-2015). The default local NLP solver in gPROMS (based on sequential quadratic programming) is used to solve problems (P2) and (P3). Problem (P3) is initialized by dynamically moving the process model from a solvent structure  $\mathbf{n}^0$  to  $\mathbf{n}^k$ . The master problem is solved by Gurobi (Gurobi Optimization, Inc., 2015).

Table 1: Primal and Master Problems

Primal Problem (P3)	Master Problem (P4)
$\min_{\mathbf{u}} f(\mathbf{u}, \mathbf{n}^k)$ $\text{s.t. } \mathbf{g}_p(\mathbf{u}, \mathbf{n}^k) = 0$ $\mathbf{g}(\mathbf{u}, \mathbf{n}^k) \leq 0$ $\mathbf{u}^{L^k} \leq \mathbf{u} \leq \mathbf{u}^{U^k}$	$\min_{\mathbf{u}, \mathbf{n}, \eta} \eta$ $\text{s.t. } f^i + \nabla_{\mathbf{n}} f^i [\mathbf{n} - \mathbf{n}^i]^T + \nabla_{\mathbf{u}} f^i [\mathbf{u} - \mathbf{u}^i]^T < \eta \quad \forall i \in F^k$ $g^j + \nabla_{\mathbf{n}} g^j [\mathbf{n} - \mathbf{n}^j]^T + \nabla_{\mathbf{u}} g^j [\mathbf{u} - \mathbf{u}^j]^T \leq 0 \quad \forall j \in A^k$ $L(1 - y^r) + \varepsilon \leq a - a^r \leq U y^r - \varepsilon \quad \forall r \in IC^k$ $a = \sum_t (b^{t-1} n_t)$ $\mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U$ $C\mathbf{n} \leq \mathbf{d}$ $y^r \in \{0, 1\} \quad \forall r \in IC^k$

 Table 2: Performance of tests 1 and 2, with  $y_{F\text{CO}_2} = 0.8$ ,  $T_F = 320$  K,  $P^{U0} = 100$  MPa,  $T^{U0} = 340$  K

	Test 1	Test 2
Number of molecules tested	1109	449
Number of molecules eliminated by test	660	0
Number of molecules with updated bounds	N/A	392
Average updated upper bound on pressure	N/A	30.039 MPa

## 4. Case study and Results

The separation of carbon dioxide from methane is studied here. A process with a 10 stage absorption column operating at a variable pressure  $P_{abs}$  is considered. The solvent entering the absorber has a flow rate of  $F_{solvent}$ . The spent solvent is regenerated in an adiabatic flash drum at 1 bar. The property prediction model employed is the Statistical Associating Fluid Theory - $\gamma$  Mie equation of state (Papaioannou et al., 2014). More details on the process and molecular models may be found in Burger et al. (2015). The objective function is the net present value (*NPV*) of the process computed over a 10-year period. The design variables are  $P_{abs}$ ,  $F_{solvent}$  and  $\mathbf{n}$ , the structure of the optimal solvent. The treated gas exiting the absorber has to be of at least 97 (mole) % methane purity and the temperature in the flash drum must be at least ten degrees above the melting point of the solvent. The area and height of the absorber are constrained to be less than 300 m<sup>2</sup> and 50 m, respectively.

The effectiveness of each of the tests is first investigated. In test 0, for a feed of composition  $y_{F\text{CO}_2} = 0.8$ ,  $P_F = 7.961$  MPa,  $T_F = 320$  K and user-specified pressure bounds of  $0.1 \text{ MPa} \leq P_{abs} \leq 7.5 \text{ MPa}$ , the updated bounds on pressure and temperature obtained are  $P^{L0} = 4.961$  MPa and  $T_F^{L0} = 288$  K, respectively. This results in a 52 % reduction of the process domain. The reduction in the domain obtained from tests 1 and 2 when applied on a search space comprising of the groups CH<sub>3</sub>, CH<sub>2</sub>, eO (an oxygen group between a CH<sub>3</sub> and CH<sub>2</sub> group), cO (an oxygen group between two CH<sub>2</sub> groups), COO, =CH, COOH and CH<sub>2</sub>OH is shown in Table 1. A molecule has an updated pressure bound if test 2 identifies a  $P^{U1}$ , such that  $P^{U1} < P^{U0}$ . 60 % of the solvent design space is eliminated by test 1. The remaining molecules are evaluated in test 2 which updates the upper bound on absorber pressure for 87 % of these molecules.

The results of applying the feasibility-based algorithm on the case study are shown in Table 2. The average number of iterations of the algorithm (over ten runs with different values of  $\mathbf{n}^1$ ) is 16, whereas the average number of evaluations of the primal is 14. Importantly, the problem converges to the solution in each of the ten runs. The use of the algorithm has enhanced convergence as compared to previous work (Burger et al., 2015), indicating that the tests help to mitigate the numerical difficulties that arise when solving such highly nonlinear problems.

Table 3: Results of the algorithm

$$y_{F_{CO_2}} = 0.2, P_F = 7.961 \text{ MPa}, T_F = 301.4 \text{ K}, 0.1 \text{ MPa} \leq P_{abs} \leq 7.5 \text{ MPa}, 0 < F_{solvent} \leq 50 \text{ kmols}^{-1}$$

Solvent	$P_{abs}/\text{MPa}$	$F_{solvent}/\text{kmols}^{-1}$	$NPV/\text{Billion USD}$
penta(oxyethylene)dimethyl ether	3.83	0.84	1.72

## 5. Conclusions

A feasibility-based algorithm for CAMPD has been proposed with novel tests that reduce both the process and molecular domain. As demonstrated in a case study, the discrete space is reduced by test 1. Implicit discontinuities in process models are addressed by tests 0 and 2. Numerical difficulties are avoided by optimizing within a feasible search space and process model initialization, making the solution of challenging CAMPD problems feasible.

**Acknowledgements** The authors gratefully acknowledge financial support from the Engineering and Physical Sciences Research Council (EPSRC) grants EP/E016340, EP/J014958/1 and EP/J003840/1.

## References

- J. Burger, V. Papaioannou, S. Gopinath, G. Jackson, A. Galindo, C. S. Adjiman, 2015. A hierarchical method to integrated solvent and process design of physical  $CO_2$  absorption using the SAFT- $\gamma$  Mie approach. *AIChE Journal* 61, 3249–3269.
- A. Buxton, A. G. Livingston, E. N. Pistikopoulos, 1999. Optimal design of solvent blends for environmental impact minimization. *AIChE Journal* 45, 817–843.
- M. A. Duran, I. E. Grossmann, 1986. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming* 36, 307–339.
- F. T. Eljack, C. C. Solvason, N. Chemmangattuvalappil, M. R. Eden, 2008. A property based approach for simultaneous process and molecular design. *Chinese Journal of Chemical Engineering* 16, 424–434.
- R. Fletcher, S. Leyffer, 1994. Solving mixed integer nonlinear programs by outer approximation. *Mathematical Programming* 66, 327–349.
- Gurobi Optimization, Inc., 2015. Gurobi optimizer reference manual 6.5.
- P. M. Harper, R. Gani, P. Kolar, T. Ishikawa, 1999. Computer-aided molecular design with combined molecular modeling and group contribution. *Fluid Phase Equilibria* 158, 337–347.
- E. J. Henley, J. D. Seader, D. K. Roper, 2011. *Separation process principles*. Wiley.
- M. Hostrup, P. M. Harper, R. Gani, 1999. Design of environmentally benign processes: integration of solvent design and separation process synthesis. *Computers & Chemical Engineering* 23, 1395–1414.
- T. G. Hunter, A. W. Nash, 1934. The application of physico-chemical principles to the design of liquid-liquid contact equipment. part II. application of phase-rule graphical methods. *Journal of the Society of Chemical Industry*, 95–102.
- A. T. Karunanithi, L. E. K. Achenie, R. Gani, 2005. A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures. *Industrial & Engineering Chemistry Research* 44, 4785–4797.
- Z. Kravanja, I. E. Grossmann, 1994. New developments and capabilities in PROSYNan automated topology and parameter process synthesizer. *Computers & Chemical Engineering* 18, 1097–1114.
- M. Lampe, M. Stavrou, J. Schilling, E. Sauer, J. Gross, A. Bardow, 2015. Computer-aided molecular design in the continuous-molecular targeting framework using group-contribution PC-SAFT. *Computers & Chemical Engineering* 81, 278 – 287.
- V. Papaioannou, T. Lafitte, C. Avendaño, C. S. Adjiman, G. Jackson, E. A. Müller, A. Galindo, 2014. Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments. *The Journal of Chemical Physics* 140 (5).
- F. E. Pereira, E. Keskes, A. Galindo, G. Jackson, C. S. Adjiman, 2011. Integrated solvent and process design using a SAFT-VR thermodynamic description: High-pressure separation of carbon dioxide and methane. *Computers & Chemical Engineering* 35, 474–491.
- Process Systems Enterprise, 1997-2015. gPROMS Model Builder 4.1.0.