

Computer Aided Molecular Design: Fundamentals, Methods and Applications

Athanasios I. Papadopoulos¹, Ioannis Tsivintzelis², Patrick Linke³, Panos Seferlis⁴

¹Chemical Process and Energy Resources Institute, Centre for Research and Technology Hellas, 57001, Thessaloniki, Greece

²Department of Chemical Engineering, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece

³Department of Chemical Engineering, Texas A&M University at Qatar, Doha, PO Box 23874, Education City, Qatar

⁴Department of Mechanical Engineering, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece

Pre-publication, corrected proof by permission of Elsevier

Please cite as:

Papadopoulos Athanasios I., Tsivintzelis Ioannis, Linke Patrick and Seferlis Panos. (2018) Computer-Aided Molecular Design: Fundamentals, Methods, and Applications. In: Reedijk, J. (Ed.) Elsevier Reference Module in Chemistry, Molecular Sciences and Chemical Engineering. Waltham, MA: Elsevier. 26-Jul-18 doi: 10.1016/B978-0-12-409547-2.14342-2.

¹Corresponding author: Athanasios I. Papadopoulos, spapadopoulos@cperi.certh.gr

a0010

Computer-Aided Molecular Design: Fundamentals, Methods, and Applications

Athanasios I Papadopoulos, Centre for Research and Technology Hellas, Thessaloniki, Greece

Ioannis Tsivintzelis, Aristotle University of Thessaloniki, Thessaloniki, Greece

Patrick Linke, Texas A&M University at Qatar, Doha, Qatar

Panos Seferlis, Aristotle University of Thessaloniki, Thessaloniki, Greece

© 2018 Elsevier Inc. All rights reserved.

Introduction	4
Molecular Representation and Property Prediction	5
GC Approach	5
Other Molecular Representations	6
Property Prediction Models	7
<i>Pure component properties</i>	7
<i>Activity coefficient models</i>	7
<i>Cubic equations of state</i>	7
<i>EoS/G^E models</i>	9
<i>Advanced statistical thermodynamic models—The statistical associating fluid theory approach</i>	10
<i>Computational chemistry models—The conductor-like screening model approach</i>	12
Main Classes of CAMD Problems	13
Classification Overview	13
Single-Objective Computer-Aided Molecular Design	14
Computer-Aided Molecular and Process Design	15
MultiObjective Computer-Aided Molecular Design	15
Computer-Aided Mixture Design	16
Computer-Aided Molecular Design Under Uncertainty	16
Computer-Aided Molecular, Process, and Control Design	16
Computer-Aided Molecular Design for Reactive Systems	18
Computer-Aided Molecular Design Using Quantum Mechanical Models	18
Summary of Classes	18
General Solution Approaches	19
Challenges in the Solution of Molecular Design Problems	19
Classification of Solution Approaches	19
Generate and Test Approaches	20
Deterministic Optimization Algorithms	21
<i>Overview of main algorithms</i>	21
<i>Deterministic optimization approaches in molecular design</i>	23
Stochastic Optimization Algorithms	25
<i>Overview of main algorithmic concepts</i>	25
<i>Genetic algorithms (GA)</i>	25
<i>Simulated annealing</i>	26
<i>Tabu search</i>	27
<i>Ant colony optimization</i>	27
Property Clustering Approaches	27
Merits and Shortcomings of Solution Approaches	28
Methodologies for Solution of CAMPD Problems	30
Generate and Test Approaches	30
Deterministic Optimization Approaches	30
Stochastic Optimization Approaches	33
Continuous Molecular Representations	34
Property Clustering and/or Reverse Formulations	36
Methodologies for Various Problem Classes	36
Multiobjective Optimization	36
Mixture and Blend Design	38
Molecular Design Under Uncertainty	43
Molecular, Process, and Control Design	46
Exploitation of Advanced Computing Infrastructures	47
Methodologies for Reactive Systems or QM-Based Approaches	47
Molecular Design	51

<i>CAMD in reactive systems</i>	51
<i>QM models in reactive systems</i>	52
<i>QM models in nonreactive systems</i>	53
Molecular and Process Design	54
Applications	55
Solvents for Industrial Separations, Reactive Separations, and Promotion of Reactions	56
Catalysts, Adsorbents, and Ionic Liquids	56
Materials for CO ₂ Capture	63
Heat Exchange Fluids	64
Polymers	64
Bio-Based Fuels and Chemicals	64
Formulated Products and Other Chemicals	66
Future Outlook and Further Reading	66
References	69
Further Reading	76

Nomenclature

dt0015	A	Helmholtz free energy
dt0020	A_k	Matrix of number of occurrences of molecular groups in molecule k
dt0025	ACO	Ant colony optimization
dt0030	ADI	Absorption-desorption index
dt0035	AHP	Analytic hierarchy process
dt0040	a_{mn}	Group-group interaction parameter used in calculation of γ_i^c
dt0045	AUP^M	Augmented property index
dt0050	AP	Augmented penalty
dt0055	ASOG	Analytical Solution of Groups
dt0060	BB	Branch and bound
dt0065	BM	Big-M formulation
dt0070	b_{mn}	Group-group interaction parameter for higher order approximation used in calculation of γ_i^c
dt0075	C	Vector of concentration of components/molecules
dt0080	CADD	Computer-aided drug design
dt0085	CAM^bD	Computer-aided mixture design
dt0090	CAMD	Computer-aided molecular design
dt0095	CAMP_{CD}	Computer-aided molecular, process and control design
dt0100	CAMP_D	Computer-aided molecular and process design
dt0105	CAM^RD	Computer-aided molecular design for reactions
dt0110	CAM^uD	Computer-aided molecular design under uncertainty
dt0115	CAOS	Computer-aided organic synthesis
dt0120	C_j	Cluster for property j
dt0125	c_k	Concentration of k^{th} component/molecule
dt0130	c_{mn}	Group-group interaction parameter for higher order approximation used in calculation of γ_i^c
dt0135	CoMT	Continuous molecular targeting
dt0140	COSMO	Conductor-like Screening Model
dt0145	D	Vector of design variables
dt0150	DEA	Data envelopment analysis
dt0155	DFO	Derivative-free optimization
dt0160	DFT	Density functional theory
dt0165	d_i	i^{th} design variable
dt0170	D_j	Contributions of second-order groups
dt0175	E	Vector of disturbances
dt0180	EACO	Efficient ACO
dt0185	EBS	Environmentally benign solvent
dt0190	E_l	Contributions of third-order-groups
dt0195	EoS	Equation of state
dt0200	ER	Equality relaxation
dt0205	F	Vector of objective functions
dt0210	F_j	j^{th} objective function
dt0215	GA	Genetic algorithms
dt0220	GBD	Generalized Benders decomposition
dt0225	GC	Group contributions

dt0230	GDP	Generalized disjunctive programming
dt0235	g_l	Molecular group l
dt0240	GXL	Gas-expanded liquids
dt0245	h	Vector of equality constraints
dt0250	HR	Hull reformulation
dt0255	HSS	Hammersley sequence sampling
dt0260	HSTA	Hammersley stochastic annealing
dt0265	IA	Interval analysis
dt0270	IPS	In-process solvent
dt0275	LCVM	Linear combination of Huron's and Michelsen's mixing rules
dt0280	LHS	Latin hypercube sampling
dt0285	m	Vector of molecular groups
dt0290	m	Takes the values of -1 for acyclic, 0 for monocyclic and 1 for bicyclic compounds
dt0295	MHV1	First order Huron-Vidal mixing rule
dt0300	MHV2	Second order Huron-Vidal mixing rule
dt0305	MIDO	Mixed-integer dynamic optimization
dt0310	MIHDE	Mixed-integer hybrid differential evolution
dt0315	MILP	Mixed-integer linear programming
dt0320	MINLP	Mixed-integer nonlinear programming
dt0325	MIP	Mixed-integer programming
dt0330	MISQP	Trust region sequential quadratic programming algorithm
dt0335	M_j	Number of occurrences of second-order groups of type j
dt0340	M_k	Matrix indicating the group composition of molecule k
dt0345	MOEGA	Multi-objective efficient GA
dt0350	MOO	Multi-objective optimization
dt0355	N_d	Total number of design variables
dt0360	N_g	Total number of molecular groups
dt0365	N_i	Number of occurrences of first-order groups of type i
dt0370	n_l^k	Number of occurrences of molecular group l in molecule k
dt0375	NLP	Nonlinear programming
dt0380	N_M	Total number of molecules
dt0385	N_{of}	Total number of objective functions
dt0390	N_p	Total number of process design variables
dt0395	N_{rv}	Total number of manipulated variable
dt0400	N_y	Total number of controlled variable
dt0405	OA	Outer approximation
dt0410	O_l	Number of occurrences of third-order groups of type l
dt0415	ORC	Organic Rankine cycle
dt0420	P	Vector of process design variables
dt0425	P_c	Critical pressure
dt0430	P_{jg}	Contribution of group g in property j
dt0435	p_q	q^{th} process design variable
dt0440	P_r^{sat}	Reduced vapor pressure
dt0445	PR	Peng-Robinson
dt0450	PSA	Pressure swing adsorption
dt0455	Q	Van der Waals group surface
dt0460	q	Vector of inequality constraints
dt0465	QM	Quantum mechanics
dt0470	QSAR	Quantitative structure-activity relationship
dt0475	QSPR	Quantitative structure-property relationship
dt0480	R	Universal gas constant
dt0485	R_v	Van der Waals group volume
dt0490	RK	Redlich-Kwong
dt0495	RNFA	Reaction network flux analysis
dt0500	SA	Simulated annealing
dt0505	SAFT	Statistical Associating Fluid Theory
dt0510	SOO	Single-objective optimization
dt0515	SPM	Stochastic property matching
dt0520	SPO	Stochastic property optimization
dt0525	SQP	Sequential quadratic programming
dt0530	SRK	Soave-Redlich-Kwong

dt0535	SRLP	Successive regression and linear programming
dt0540	T_c	Critical temperature
dt0545	T_r	Reduced temperature
dt0550	TS	Tabu search
dt0555	UNIFAC	Universal Quasi-Chemical Functional-group Activity Coefficients
dt0560	vdW	Van der Waals
dt0565	vdW1f	Van der Waals one fluid mixing rule
dt0570	vPPD	Virtual product-process design laboratory
dt0575	VSA	Vacuum swing adsorption
dt0580	U	Vector of uncertain parameters
dt0585	W	Vector of manipulated variables
dt0590	w	Binary variable indicating use of second-order groups
dt0595	w_{l2}	t_2^{th} manipulated variable
dt0600	X	Vector of state variables
dt0605	X_i	Contributions of first-order groups
dt0610	Y	Vector of controlled variables
dt0615	y_{t3}	t_3^{th} controlled variable
dt0620	z	Binary variable indicating use of third-order groups

Greek Symbols

dt0630	γ_i	Activity coefficient of component i
dt0635	γ_i^c	Combinatorial part of activity coefficient
dt0640	γ_i^r	Residual part of activity coefficient
dt0645	ϵ_{t1}	t_1^{th} disturbance parameters
dt0650	N_ϵ	Total number of disturbance parameters
dt0655	v_l^k	Valence of group l in k^{th} molecule
dt0660	$\psi_j^M(P_j)$	Molecular property operator of the j^{th} property
dt0665	$\psi_j^{ref}(P_j)$	Molecular property operator for a reference value
dt0670	ω	Acentric factor
dt0675	Ω_{ji}^M	Normalized value of operators

s0010 Introduction

p0575 The selection of chemicals with properties which are desired in a particular application is a challenging task. A vast number of compounds may be available that possess suitable characteristics, hence they would all need to be assessed prior to the selection of the one with the desired behavior. Conventionally, such screening is performed experimentally, with significant challenges pertaining to the costs and effort required to perform experimental iterations. Computer-aided technologies offer a promising route to exploit experimental know-how and guide the search for novel and efficient chemicals as they enable the investigation of an enormous range of options. Such technologies allow the fast, cost-effective, and automated evaluation of a vast number of characteristics that may lead to highly efficient chemicals. They do not entirely replace experiments. Instead, experimentally obtained results are exploited to formulate mathematical representations of phenomena that can be used in computer simulation or optimization and further experiments can then be performed to verify the results of the optimization.

p0580 The exploitation of computer-aided technologies for the identification of chemicals has two major requirements:

- o0010 (a) The exhaustive generation and simultaneous evaluation of a very wide range (millions) of chemicals in order to identify the ones exhibiting optimum performance by any set of criteria.
- o0015 (b) The simultaneous consideration of models of sufficiently high fidelity, so that the obtained chemicals are both optimum and rigorously validated prior to practical implementation at industrially or commercially relevant conditions.

p0595 Conventionally, these requirements are addressed through an approach that can generally be described as follows:

p0600 "For a given molecule and associated model data, calculate the desired properties."

p0605 This is a forward problem formulation which gives rise to significant shortcomings. The exhaustive evaluation of all potentially known molecular structures is computationally very expensive, especially when it is combined with molecular models of high fidelity. Chemicals are selected from an arbitrarily compiled database which contains few options. This is extremely limiting in view of the vast number of chemicals that can be considered as candidates for a specific application. The selected candidates are feasible, but only offer incremental performance improvement over previously tested solutions. This limits the possibility of discovering novel molecular structures which may result in significant performance improvement compared to conventional options.

p0610 These challenges may be addressed efficiently by an alternative approach which results from a reverse formulation of the previous statement as follows:

p0615 “For given properties representing desired performance targets and associated model data, identify the molecular structure(s) that best match these targets.”

p0620 This statement represents the main idea behind computer-aided molecular design (CAMD). Unlike the forward formulation, CAMD does not require the *a priori* specification of a chemical as a candidate for a particular application. It is sufficient to specify the desired property target, while the chemical that best matches this target will result as the solution of the molecular design problem. For example, consider that a solvent must be identified for the separation of a mixture through liquid–liquid extraction. In the CAMD formulation, this goal may be represented through a property such as solvent–solute solubility and the result will be a solvent that will exhibit the maximum possible solubility toward the solute. In CAMD, it is not necessary to have a database of chemicals. Instead, molecular fragments are used as building blocks for the synthesis of molecular structures, hence the term “design.” This is possible due to the advent of group-contribution (GC) methods,¹ where properties may be calculated through models that account for the molecular structure as a combination of the groups comprising the molecule. These are the main characteristics of a CAMD approach as defined originally by Gani and Brignole.² Over the years, CAMD has evolved toward different directions including the development of^{3, 4}:

- u0010 ● Predictive models that can capture the nonideal behavior of molecules and mixtures, often provided in the form of various different quantitative structure–property relations (QSPR) such as GC.
- u0015 ● Modeling approaches and simulation or optimization algorithms which can efficiently address numerical issues resulting from nonidealities.
- u0020 ● Optimization approaches and algorithms which can handle nonlinear models of high fidelity and the combinatorial complexity that results from the countless molecular and process option candidates for desired applications.

p0640 In this article, we provide an overview of the fundamental features of CAMD together with methods to address computational complexity and applications in various sectors. See “Molecular Representation and Property Prediction” section presents QSPR and predictive models, whereas see “Main Classes of CAMD Problems” section presents a classification of major CAMD problems. See “General Solution Approaches” section presents algorithms used for the solution of CAMD problem classes, whereas see “Methodologies for Solution of CAMPD Problems,” “Methodologies for Various Problem Classes,” and “Methodologies For Reactive Systems Or QM-Based Approaches” sections discuss advanced methodologies which combine different solution algorithms. See “Applications” section presents a review of CAMD applications in various domains, and see “Future Outlook and Further Reading” section provides a future outlook of CAMD technology together with sources for further reading.

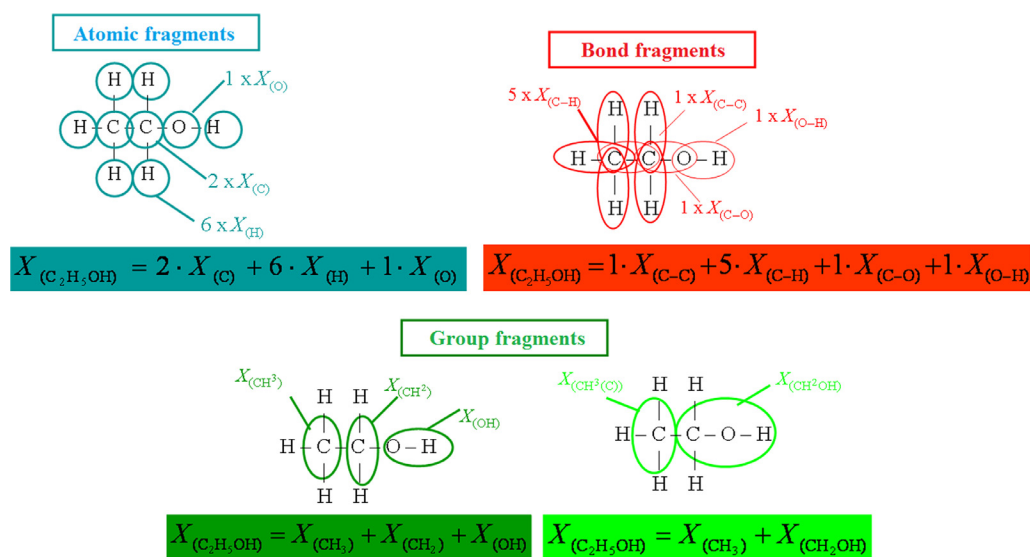
s0015 Molecular Representation and Property Prediction

s0020 GC Approach

p0645 For the prediction of molecular properties, all models require data which are generally available and easy to find when it comes to the evaluation of few or conventional molecules. However, data are very difficult to obtain for the evaluation of wide molecular sets, while the problem is amplified in cases where it is desired to assess novel molecular structures and to use high fidelity property prediction models. Conventional property modeling approaches are unable to address these challenges. For every chemical that needs to be simulated, experimental data must be fitted in a model which is then used to predict its properties at different conditions. This is clearly inefficient because large experimental effort is required for every chemical.

p0650 Instead of fitting models for entire molecules, it is more efficient to develop models for the molecular fragments (atoms, bonds, or groups) that comprise molecules. This is because the properties of a molecule are usually established based on contributions from its fragments.⁵ The intermolecular forces that determine the properties of interest depend mostly on the bonds between the atoms of the molecules. Such contributions are largely determined by the nature of the atoms involved (atom contributions), the bonds between pairs of atoms (bond contributions or equivalently group interaction contributions), or the bonds within and among small groups of atoms (GCs).⁵ It is therefore possible to use experimental data for entire molecules and fit models which are based on the effects of their fragments in the corresponding properties. It is generally assumed that fragments can be treated independently of their arrangements or their neighbors. In case of insufficient accuracy, corrections may be incorporated for specific (higher-order) multigroup, conformational or resonance effects. This is the underlying concept of the GC methods. For example, instead of fitting data into a model that directly predicts the critical temperature of ethanol, it is possible to develop a model which exploits the functional groups comprising ethanol, namely $-\text{CH}_3$, $>\text{CH}_2$, and $-\text{OH}$, to predict this property. Fig. 1 shows how a property X may be predicted using a GC approach. It is worth noting that GC models employing functional groups have prevailed over atom and bond contributions in CAMD implementations.

p0655 The idea is very convenient because the contributions of functional groups may be derived by considering experimental group interactions for a very wide set of molecular structures. As a result, the contributions of groups to the calculation of the desired property remain the same regardless of the structure in which they are utilized. For example, the calculation of the critical temperature for a completely different structure than ethanol, such as an alkyl-ester or an alkyl-amine, can be based on the same contributions used for ethanol, supplemented by contributions for ester and amine groups. This makes it possible to predict the properties for a vast number of chemicals, even of ones which were not in the set initially used to derive the models. Historically, the idea of GC first appeared in Langmuir,⁷ according to Redlich et al.⁸ The latter was a main precursor for the work of Fredenslund et al.¹ which addressed the GC-based calculation of mixture properties using Universal Quasi-Chemical Functional-group Activity Coefficients (UNIFAC). These GCs were used by Gani and Brignole² in their original CAMD formulation.



p0010 **Fig. 1** The concept of GCs based on atomic fragments, bond fragments, first- and higher-order functional groups. Reproduced from Kolska, Z.; Zabransky, M.; Randova, A., Group Contribution Methods for Estimation of Selected Physico-Chemical Properties of Organic Compounds. In *Thermodynamics-Fundamentals and Its Application in Science*. Morales-Rodriguez, R. Ed.; InTech: London, 2012, Chapter 6, pp. 135–162. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

p0660 Formally, the addition of functional groups for the prediction of a property is shown in Eq. (1)⁹:

$$f(x) = \sum_i N_i X_i + w \sum_j M_j D_j + z \sum_l O_l E_l \quad (1)$$

where N_i , M_j , and O_l represent the number of occurrences of first-, second-, and third-order groups of types i , j , l in a molecule, respectively, and X_i , D_j , E_l are the contributions of the corresponding groups. Constants w and z are assigned zero and unity values depending on the use of second- or third-order groups, in addition to first-order ones. First-order groups are basic molecular fragments which support the representation of a very ample variety of organic compounds. Second- and third-order groups introduce additional structural information about more extensive molecular fragments in order to improve the description provided by first-order groups. Second- and third-order groups further enable a better description of multifunctional or multicyclic compounds and support differentiation between isomers. It is also worth noting that function $f(x)$ may either be a linear model or a nonlinear model of the corresponding property. For example, in normal boiling point $f(x)$ is a logarithmic model of the right-hand side of Eq. (1), whereas the critical volume is a linear model.⁹

s0025 Other Molecular Representations

p0665 GC molecular representations through functional groups do not account for the internal molecular architecture, hence the molecule is represented as a loose composition of groups which includes the type of groups that comprise it and the number of occurrences of each group. Other QSPR have been developed which offer a more rigorous representation of the molecular structure (e.g., considering atoms or bonds). Main types of such QSPR include the topological indices and the signature molecular descriptors.

p0670 Topological indices characterize a compound with a single number, based on the interconnectivity and the types of atoms in the molecule.¹⁰ They are based on the concept of the molecular graph, that is, the graph representation of a molecule where atoms and bonds correspond to vertices and edges, respectively. Molecular graphs can be represented by different types of matrices such as the vertex adjacency matrix, the edge adjacency matrix, and the incidence matrix, to name but a few. Topological indices which are based on molecular connectivity are also called connectivity indices. The first connectivity indices were proposed by Randic,¹¹ who used them to describe the degree of branching in alkanes and to model enthalpy of fusion and vapor pressure. Austin et al.⁴ provide an overview and an interesting discussion of topological and connectivity indices.

p0675 Signature descriptors are based on molecular graphs where vertices denote atoms and edges denote bonds. A molecule is therefore represented by a set of canonical subgraphs, each rooted on a different vertex with a predefined level of branching, which is called height. Atomic signatures are defined as canonical subgraphs consisting of all atoms at a distance from the root which is equal to the height. The set of all unique atomic signatures and the occurrence with which they appear in the molecular graph are then defined as the molecular signature.¹² Austin et al.⁴ provide a detailed overview of signature descriptors together with GC implementations.

s0030 Property Prediction Models

s0035 Pure component properties

p0680 The original, UNIFAC-based GC concept of Fredenslund et al.¹ has been extended by Constantinou and Gani,¹³ who proposed the use of UNIFAC groups for the prediction of several pure component properties, with revisions and significant extensions to other properties presented by Marrero and Gani⁹ and Hukkerikar et al.¹⁴. The GC method of Joback and Reid¹⁵ is also used frequently in CAMD. A wide collection of GC models for pure component properties or activity coefficient models can be found in Poling et al.⁵ and Kolska et al.⁶ Hukkerikar et al.¹⁶ present a collection of GC models for the prediction of environmental and health properties. Van Krevelen and Te Nijenhuis¹⁷ provide a wide collection of GC models for the prediction of polymer properties, while Coutinho et al.¹⁸ provide a collection of GC models for ionic liquid properties. Austin et al.⁴ provide a review of GC-based models using connectivity indices and signature descriptors.

s0040 Activity coefficient models

p0685 Activity coefficient models have been used widely in the context of GC formulations for the prediction of phase equilibria. The idea is to use existing equilibrium data for predicting phase equilibria of systems for which no experimental data are available.⁵ The activity coefficient generally consists of two main parts, namely the one that accounts for differences in molecular size and shape and the other one that considers molecular interactions. According to Gmehling et al.,¹⁹ the first GC method for the prediction of vapor–liquid equilibria through an activity coefficient model was the ASOG (Analytical Solution of Groups) method.^{20, 21} ASOG contained some arbitrary assumptions in the activity coefficient estimation,⁵ which were efficiently addressed through UNIFAC proposed by Fredenslund et al.¹ In UNIFAC, the activity coefficients γ consist of a combinatorial part c which accounts for differences in molecular size and shape in a mixture, and a residual part r which accounts for energy interactions and functional group sizes and interaction surface areas.

$$\ln\gamma_i = \ln\gamma_i^c + \ln\gamma_i^r \quad (2)$$

p0690 The calculation of the activity coefficient requires only three parameters, namely the van der Waals (vdW) group volume R_v , vdW group surface area Q , and the group–group interaction parameter a_{mn} . The first two parameters participate in the calculation of the combinatorial part which requires only pure component information. The group–group interaction parameter participates in the residual part and needs to be evaluated from experimental phase-equilibrium data.⁵

p0695 The original UNIFAC¹ has been very successful, because it can predict reliably mixture properties over wide temperature and concentration ranges, even azeotropic points. However, the predictions of activity coefficients at infinite dilution and of excess enthalpies lack reliability, while results for asymmetric systems also require improvement.¹⁹ Such shortcomings have been addressed through the modified UNIFAC (Do)^{22, 23} and the modified UNIFAC (Ly).^{24, 25} The introduced modifications focus on a better representation of γ_i^c . The latter is updated with a volume term which is a nonlinear function of the vdW volume r (calculated through R_v) and a group–group interaction parameter with a second-order polynomial temperature dependence which requires data in UNIFAC (Do) for parameters b_{mn} and c_{mn} , in addition to a_{mn} .¹⁹ Further to addressing previous shortcomings, these modifications also allow UNIFAC (Do) to provide reliable predictions even for materials exhibiting very nonideal phase equilibrium behavior, such as polymers and ionic liquids.^{26, 27}

s0045 Cubic equations of state

p0700 Equations of state (EoS) are used for calculations of thermodynamic properties of mixtures in industry and in academia, especially for systems at high pressures. The term “high pressure” refers to pressures high enough to significantly alter the thermodynamic properties of both liquid and vapor phases, typically over 10 bar. In most cases, such high-pressure mixtures contain at least one component at supercritical conditions.

p0705 The most popular class of EoS is the so-called cubic. Such models originate from the vdW equation of state (EoS), which was suggested in 1873 as the first model that could describe properties of both liquid and vapor phases.²⁸ The vdW EoS has been the base for hundreds of EoS in the past years.^{29, 30} For example, considering the very popular Peng–Robinson EoS there are more than 220 modifications and uncountable studies related with parameter estimation and extension to mixtures.³⁰

p0710 Cubic EoS became very popular in oil and gas industry, since they are simple in terms of mathematical formulation and present satisfactory correlations and predictions of thermodynamic properties for mixtures of nonpolar and nonhydrogen bonding fluids.^{31, 32} Cubic EoS are reviewed in some excellent research papers and books chapters.^{29–36} In this section, the essentials of such models are presented to emphasize in routes that render those models predictive.

p0715 Besides the vdW EoS,²⁸ which now presents mostly historical value, the most well-known cubic EoS are the Redlich–Kwong³⁷ (RK), the Soave–Redlich–Kwong³⁸ (SRK), and the Peng–Robinson³⁹ (PR) EoS.^{31, 32} Such EoS, along with typical expressions often used for estimating their parameters, are presented in Table 1.

p0720 The vdW EoS can be considered as a modification of the ideal gas EoS by introducing two significant improvements, that is, the hard-core volume of molecules, which is related to the parameter b and the attractive intermolecular interactions related to the parameter a . Thus, the first term of the right-hand side of the corresponding equation of Table 1 refers to the repulsive, while the second term to the attractive intermolecular interactions.³² Parameters a and b are fluid-specific parameters and can be estimated if the critical temperature, T_c , and critical pressure, P_c , are known. However, the vdW equation provides only a quantitative description of vapor–liquid equilibrium and volumetric properties of real fluids and mixtures, even for simple ones, such as light hydrocarbons.^{31, 32} Such inefficiency of the model resulted in the development of many improved versions of cubic EoS during the last century.⁴⁰

10010 **Table 1** The most well-known cubic equations of state

EoS	Equation	Energy and covolume parameters
Van der Waals ²⁸ (vdW)	$P = \frac{RT}{V-b} - \frac{a}{V^2}$	$a = \frac{27}{64} \frac{(RT_c)^2}{P_c}$ $b = \frac{1}{8} \frac{RT_c}{P_c}$
Redlich–Kwong ³⁷ (RK)	$P = \frac{RT}{V-b} - \frac{a}{V(V+b)\sqrt{T}}$	$a = 0.42748 \frac{(RT_c)^{2.5}}{P_c}$ $b = 0.08664 \frac{RT_c}{P_c}$
Soave–Redlich–Kwong ³⁸ (SRK)	$P = \frac{RT}{V-b} - \frac{a(T)}{V(V+b)}$	$a_c = 0.42748 \frac{(RT_c)^2}{P_c}$ $b = 0.08664 \frac{RT_c}{P_c}$ $a(T) = a_c [1 + m(1 - \sqrt{T_r})]^2$ $m = 0.48 + 1.57\omega - 0.176\omega^2$
Peng–Robinson ³⁹ (PR)	$P = \frac{RT}{V-b} - \frac{a}{V(V+b)+b(V-b)}$	$a_c = 0.45724 \frac{(RT_c)^2}{P_c}$ $b = 0.07780 \frac{RT_c}{P_c}$ $a(T) = a_c [1 + m(1 - \sqrt{T_r})]^2$ $m = 0.37464 + 1.54226\omega - 0.269926\omega^2$

p0725 In this direction, Redlich and Kwong³⁷ added a temperature dependence in the parameter a , which was revised as $a/T^{0.5}$, and developed the well-known RK EoS. Later, in 1949, Soave³⁸ suggested a significant improvement to the RK EoS by introducing a new temperature dependence to the attraction parameter, which now is a function of the acentric factor ω , that resulted in the so-called SRK EoS. The acentric factor represents a measure of the nonsphericity of the molecule and can be estimated through the following relation, if data for vapor pressures are available³¹:

$$\omega = -\log P_r^{sat}|_{T_r=0.7} - 1.00 \quad (3)$$

p0730 In the latter equation, P_r^{sat} and T_r are the reduced vapor pressure and temperature, respectively.

p0735 As shown in Table 1 for the most well-known cubic EoS, such models usually require the knowledge of two fluid-specific parameters (a and b) or three fluid-specific parameters (a , b , and ω). Having such pure fluid parameters, calculations for mixtures are feasible, since the relative mixture parameters can be calculated through appropriate mixing and combining rules. The most common mixing rules are the so-called van der Waals one fluid (vdW1f) mixing rules combined with classical combining rules, that is, the geometric mean rule for the cross-energy and the arithmetic mean rule for the cross covolume parameter:

$$a = \sum_i^n \sum_j^n x_i x_j a_{ij} \quad (4)$$

$$b = \sum_i^n \sum_j^n x_i x_j b_{ij} \quad (5)$$

$$a_{ij} = \sqrt{a_i a_j} (1 - k_{ij}) \quad (6)$$

$$b_{ij} = \frac{b_i + b_j}{2} (1 - l_{ij}) \quad (7)$$

where k_{ij} and l_{ij} are binary parameters. However, the binary interaction parameter, k_{ij} , is by far the most important one in the majority of applications.³¹

p0740 Consequently, it is clear that there is a need of fluid-specific parameters in order to apply such models. Parameters a and b are usually calculated from the experimental values of critical pressure and critical temperature, as shown in Table 1, while the acentric factor can be calculated from experimental vapor pressure data. In this way, the critical point is accurately described; however, rather unsatisfactory vapor pressures are predicted away from the critical point, while significant deviations are also observed in pressure-volume-temperature (PVT) predictions.³¹ In order to obtain a more accurate description of vapor pressure and density data, such pure fluid parameters are adjusted to experimental vapor pressures and liquid densities. In this way, a satisfactory description is obtained; however, the critical point is overpredicted.³¹ Having the pure fluid parameters, mixture properties are calculated using appropriate mixing and combining rules, such as the vdW1f rules. The binary parameters k_{ij} and l_{ij} are usually adjusted to experimental binary mixture data.

p0745 In cases of lack of any experimental data, in order to apply such models in a purely predictive way, the values of pure fluid parameters should be predicted through an appropriate method. However, as shown in Table 1, the estimation of pure fluid parameters for the common cubic EoS reduces to the estimation of the critical pressure and temperature as well as of the acentric factor. For this purpose, appropriate GC methods that are described in “Property Prediction Models” section could be applied, with the most popular being the GC methods of Amprose,^{41, 42} Joback,¹⁵ and Constantinou and Gani^{13, 43} (for an extensive review of such methods the reader is referred to the collection of Poling et al.⁵ and critical review of Kontogeorgis and Tassios⁴⁴).

p0750 Another, rather not popular, approach is the calculation of parameters a and b from first principles.³¹ For example, the vdW equation can be derived from statistical thermodynamics, by assuming a hard-core potential for repulsive interactions and a Lennard-Jones potential for attractions beyond the hard-core potential.³² In this way, the parameter α is associated with the Lennard-Jones attractive interaction parameter ε and parameter b with the hard-core volume of molecules.³² However, such approach is rarely used. In contrast, the connection of the covolume parameter b with the hard-core volume of molecules allows the development of empirical correlations with the vdW volume, which makes feasible the prediction of such molecular parameter.^{31, 45, 46} However, since the covolume parameter b is a measure of the volume of a molecule that other molecules cannot penetrate, it is expected to be higher than the molecular (or the vdW) volume.³¹ Furthermore, trends of the attraction parameters with the vdW volume can be estimated. However, such correlations are different for different homologous series, in contrast with the covolume parameter b for which universal correlations with the vdW volume can be obtained.⁴⁶

p0755 Predictions for mixture properties are obtained in two ways, that is, using a zero value for the binary parameters k_{ij} and l_{ij} and using appropriate methods for estimating them (especially the binary interaction parameter k_{ij} , which is by far the most important in most applications³¹).

p0760 In this direction, it was suggested that the binary interaction parameters could be estimated if intermolecular forces are properly accounted for. Consequently, since the parameter k_{ij} reflects a correction of the geometric mean rule for the cross-energy parameter, it can be calculated through a Hudson and McCoubrey type equation, which provides the cross-interaction energy as a function of the ionization potentials of both compounds.⁴⁷ Using this approach, Coutinho et al.⁴⁸ suggested relevant equations for both the binary parameters k_{ij} and l_{ij} . Later, Coutinho et al.⁴⁹ suggested the following relation for the binary interaction parameter:

$$k_{ij} = 1 - \left(\frac{2\sqrt{b_i b_j}}{b_i + b_j} \right)^{(n/3)-2} \quad (8)$$

p0765 The parameter n in this equation depends on the symmetry of the mixture. The authors suggest using $n = 6$ for very asymmetric mixtures.

p0770 Furthermore, many correlations to estimate the k_{ij} for specific mixtures can be found in the literature (e.g., mixtures containing light hydrocarbons,⁵⁰ as well as mixtures containing hydrocarbons and methane,⁵¹ nitrogen,⁵² and carbon dioxide^{53, 54}). Such correlations are based mostly on phase equilibrium experimental data, purely empirical and often unsuitable for extrapolation.

p0775 Finally, the binary interaction parameter can be calculated through a GC approach. In this direction, Jaubert and coworkers⁵⁵⁻⁵⁷ following an approach of Peneloux and coworkers^{58, 59} developed a GC method for estimating temperature-dependent binary interaction parameters ($k_{ij}(T)$) for the PR EoS. According to their approach, the k_{ij} of a (sub)binary system is a function of temperature and of the critical temperature, critical pressure, and acentric factor of each compound. The resulting model/approach was called PPR78 (predictive 1978, PR EoS).⁵⁵ Later, the same approach was used to develop a GC method that allows the estimation of temperature-dependent binary interaction parameters for the SRK EoS, while the resulting model/approach was called PR2SRK.⁶⁰

s0050 EoS/ G^E models

p0780 Cubic EoS coupled with the vdW1f mixing rules result in successful description of mixtures that contain nonpolar fluids. However, the vdW1f mixing rules are not capable of representing highly nonideal mixtures, such as systems with polar or hydrogen bonding fluids. The inaccuracy of cubic EoS in describing properties of polar and associating fluids, which are better correlated using an activity coefficient model, as well as the inability of the latter models for high pressure calculations, resulted in an elegant approach to combine advantages of both types of models, that is, to match the excess Gibbs energy function of the mixture, as predicted by the EoS, with the one predicted by an activity coefficient model at a reference pressure. Such approach yielded a new class of models, the so-called EoS/ G^E models, which are mainly mixing rules for the energy (and covolume) parameters of cubic EoS. The EoS/ G^E models have been reviewed in many interesting research articles and books.^{31, 34, 61, 62} Consequently, here only the basic information will be provided, focusing on purely predictive models.

p0785 The idea of G^E mixing rules was firstly introduced by Huron and Vidal,⁶³ who calculated the mixture parameter a of a cubic EoS by equating the excess Gibbs energy of the cubic EoS with the relative term of the nonrandom two-liquid (NRTL) activity coefficient model at the limit of infinite pressure. Since then, several approaches have been suggested on this idea, depending on the pressure chosen for solving the EoS.^{31, 61, 62} However, using the infinite pressure approach, the existing parameter tables from activity coefficient models, which are obtained from low pressure data, cannot be used. Furthermore, Mollerup⁶⁴ questioned the validity of using infinite pressure and instead proposed the low pressure limit to equate the G^E expressions from the two models. Such approach was also used by Michelsen and coworkers, who developed the so-called modified Huron–Vidal first-order and second-order mixing rules (MHV1 and MHV2, respectively).⁶⁵⁻⁶⁷ Later, Wong and Sandler⁶⁸ matched the Helmholtz function from a cubic EoS and from an activity coefficient model.

p0790 In order to perform calculations for mixtures using such mixing rules, one has to know the parameters of the used activity coefficient model (with NRTL, UNIQUAC, or Wilson models being the most popular). Thus, such approach cannot be used in cases of lack of any experimental data for adjusting activity coefficient model parameters. However, such difficulty can be overcome using a GC activity coefficient model (with UNIFAC being the most popular of this kind of models). Consequently, assuming that all group–group interaction parameters are known, a fully predictive model (no further parameters need to be adjusted) is obtained.

p0795 Among the first attempts to develop a predictive model was the work of Vidal and coworkers,⁶⁹ who used the Huron–Vidal mixing rule in SRK and PR using the residual term of UNIFAC. However, using the infinite pressure approach, the existing parameter tables from activity coefficient models, which are obtained from low pressure data, cannot be used and new interaction parameters should be adjusted based on experimental data. The Huron–Vidal mixing rule coupled with the SRK EoS was also used in a GC approach by Soave and coworkers.^{70–72}

p0800 Many researchers used the zero pressure approach for combining a cubic EoS with UNIFAC. Two methods attracted particular interest, that is, the MHV2 method by Michelsen and coworkers^{66, 67, 73} and the PSRK method by Gmehling and coworkers,^{19, 74–77} who extended the original UNIFAC table to groups representing light gases. The latter group also developed a volume-translated Peng–Robinson EoS.^{19, 78, 79}

p0805 Furthermore, Orbey et al.⁸⁰ developed an EoS approach based on the Wong–Sandler mixing rule⁶⁸ and the UNIFAC GC method. Without needing to redefine the UNIFAC parameters, satisfactory predictions were obtained for systems for which UNIFAC provides an accurate low pressure description.

p0810 In order to address the limitation of EoS/G^E models to size asymmetric systems, Tassios and coworkers⁸¹ suggested a new mixing rule based on the linear combination of Vidal’s and of Michelsen’s mixing rules (LCVM). Such mixing rule was coupled with a modified volume-translated PR EoS and the UNIFAC activity coefficient GC model. It was proved that the model provides successful predictions of the vapor–liquid equilibrium of polar and nonpolar fluid mixtures at low and high pressures, as well as of systems with dissimilar component size, such as those containing light gases and high molecular weight alkanes.⁸¹ According to Kontogeorgis and Goutsikos,⁶² LCVM is the last one of the models offering a significant advance in the topic of EoS/G^E mixing rules.

p0815 However, recently, Voutsas et al.⁸² suggested the so-called Universal Mixing Rule, combined with a modified volume-translated PR EoS with UNIFAC GC model. Such approach yields satisfactory phase equilibrium predictions for both symmetric and asymmetric systems, including systems containing polymers and systems with highly nonideal behavior, such as mixtures with organic acids.^{83, 84}

p0820 According to de Hemptinne et al.,³³ the availability of EoS/G^E models considerably widens the applications range of EoS, since if models are well parameterized, virtually all complex systems can be accurately described. Nevertheless, EoS/G^E models carry also limitations of the constituent models, which are, in many cases, important. A thorough review of the capabilities and limitations of EoS/G^E models is provided in the excellent review article of Kontogeorgis and Goutsikos.⁶²

s0055 **Advanced statistical thermodynamic models—The statistical associating fluid theory approach**

p0825 Over the last years, advanced thermodynamic EoS models were developed based on statistical thermodynamics.³¹ These models are more complex than traditional approaches (e.g., cubic EoS and activity coefficient models), but result in significantly more accurate predictions for systems that exhibit nonideal behavior, such as polymer solutions and hydrogen bonding fluid mixtures. One successful family of such models is based on Wertheim’s first-order thermodynamic perturbation theory^{85–88} and includes all SAFT (Statistical Associating Fluid Theory) type models that are very often used in molecular design applications.

p0830 The first SAFT approaches appeared in the literature in the end of 1980s. Based on the work of Wertheim,^{85–88} Gubbins and coworkers developed the SAFT EoS for spherical and chain molecules with one or more hydrogen bonding sites.^{89, 90} At the same time, a similar model was also developed by Huang and Radosz.^{91, 92} Both models are usually referred to as “SAFT”; however, in many cases, the acronym SAFT-HS is attributed to the former one, while the acronym SAFT-HR is attributed to the latter model.⁹³

p0835 In SAFT, the Helmholtz free energy is written as the sum of contributions due to hard-sphere repulsive interactions (A^{hs}), due to chain formation through bonding of a number of hard-spheres (A^{chain}) and due to association (A^{assoc}), while a dispersion term (A^{disp}) is added as a perturbation to the reference fluid⁹⁴:

$$\frac{A}{RT} = \frac{A^{ideal}}{RT} + \frac{A^{res}}{RT} \quad (9)$$

$$\frac{A^{res}}{RT} = \frac{A^{ref}}{RT} + \frac{A^{disp}}{RT} = \frac{A^{hs}}{RT} + \frac{A^{chain}}{RT} + \frac{A^{assoc}}{RT} + \frac{A^{disp}}{RT} \quad (10)$$

p0840 However, additional terms can be added in Eqs. (9) and (10) to account for polar and ionic interactions. Based on Wertheim’s work, contributions for the chain and the association term are obtained, which are essentially unchanged in the various versions of SAFT.³¹

p0845 Numerous EoS models that are based on the SAFT approach appeared in literature, since the development of the original SAFT EoS. Such SAFT-type models are reviewed in some review articles and book chapters.^{31, 93–97} Some of the popular SAFT-type models are presented in Table 2.

s0060 **Parameter estimation**

p0850 SAFT-type models present a more realistic, but still approximate picture, of molecules and molecular interactions, compared to traditional cubic EoS. A price to pay for incorporating such important molecular details and strong specific intermolecular

t0015 **Table 2** Equation of state models that are based on SAFT

Acronym	References
SAFT-HS	Chapman et al., ⁹⁰ Jackson et al. ⁸⁹
SAFT-HR	Huang and Radosz ^{91, 92}
Simplified SAFT	Fu and Sandler ⁹⁸
LJ-SAFT	Kraska and Gubbins ^{99, 100}
Soft-SAFT	Blas and Vega ¹⁰¹
SAFT-VR	Gil-Vilegas et al. ¹⁰² and Galindo et al. ¹⁰³
SAFT-VR Mie	Lafitte et al. ¹⁰⁴
PC-SAFT	Gross and Sadowski ^{105, 106}
Simplified PC-SAFT	von Solms et al. ¹⁰⁷
tPC-SAFT	Karakatsani et al. ¹⁰⁸
SAFT + cubic	Polishuk ¹⁰⁹

interactions is the need for a larger number of fluid parameters. Most SAFT-type models use three pure fluid parameters to characterize nonhydrogen bonding fluids, while they need two additional parameters for every hydrogen bonding interaction that occurs in the system. For example, in SAFT-HR,^{91, 92} each fluid is characterized by three parameters for nonassociating and five parameters for associating components. These parameters are the number of segments per molecule, m , the segment volume, v^{00} , and the dispersion energy parameter, u^0/k , for segment–segment interactions. Additional parameters for association are the energy of association, ε^{hb}/k , and the volume of association, κ . Such pure fluid parameters are usually estimated by adjusting model predictions to experimental data for vapor pressures and liquid densities. However, such approach may result in more than one parameter sets that accurately describe the aforementioned pure fluid properties, but not all of them are able to describe other properties, such as derivative properties and the phase behavior of mixtures. For this reason, besides the use of vapor pressures and liquid densities, the use of other experimental data such as derivative properties,^{104, 110} heats of vaporization¹¹¹ and spectroscopic or calorimetric data (for estimating the hydrogen bonding parameters)¹¹² has been suggested.

p0855 Having the pure fluid parameters, calculations for mixtures are feasible. Mixture parameters are obtained using appropriate mixing rules that are needed in the dispersion term of SAFT variants. Typically, the vdW1f mixing rules are used. Moreover, the Lorentz–Berthelot combining rules are typically used for the segment energy and diameter (or volume) parameters.

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} (1 - k_{ij}) \quad (11)$$

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \quad (12)$$

p0860 Similarly to the cubic EoS, a binary interaction parameter, k_{ij} , is often used in the combining rule as a correction for the cross-energy parameter.

p0865 In mixtures of hydrogen bonding fluids, in which cross-association occurs, additional parameters for the two association parameters are needed. Many researchers tried to adopt such parameters from spectroscopic or calorimetric data,¹¹² which, however, are rare. In most cases, combining rules, often with limited physical base, are used.

p0870 In cases of lack of any experimental data, in order to apply such models in a purely predictive way, the values of pure fluid and mixture parameters should be predicted through an appropriate method. For pure fluid parameters, such problem can be, at least partially overcome, in three possible ways, that is, using correlations of the pure fluid parameters with the molecular weight or the vdW volume, adopting some of the parameters from quantum mechanics and using GC methods.

s0065 Correlations of pure fluid parameters

p0875 The development of correlations of pure fluid parameters with molar mass was shown early from the first studies that appeared in the literature. For example Huang and Radosz^{91, 92} mention (for SAFT-HR) that for chain molecules, the segment volume and segment energy are nearly constant upon increasing the molar mass, while the segment number is a linear function of molar mass. According to the authors, such observation facilitates the prediction of molecular parameters for many compounds, such as polymers, where experimental data are rare and parameters have to be estimated based on molar mass and chemical structure only. Later, many researchers developed such correlations in order to enhance the predictive ability of the model.¹¹³ However, there are significant disadvantages of this approach. Such correlations are not universal, but have to be developed for every family of compounds, for example, such correlations are different for linear and branched hydrocarbons,⁹¹ even though that such compounds present very similar chemical structure. Moreover, such correlations are not easily obtained for the hydrogen bonding parameters, because the adjustment of five pure fluid parameters is not an easy task, since the dispersion and hydrogen bonding energy parameters are intercorrelated, without all mathematical solutions to have a physical base. However, in some cases, the three nonhydrogen bonding parameters are obtained by keeping constant the hydrogen bonding energy and volume for all compounds of the same family (e.g., alcohols and primary amines).

s0070 Adopting parameters from quantum mechanics

p0880 The successful estimation of EoS pure fluid parameters through quantum mechanical calculations is a very interesting, but rather demanding, suggestion. Such approach could result in purely predictive models; however, despite the effort, this goal has not yet been accomplished. Nevertheless, interesting approaches were suggested in literature to estimate some, but usually not all, of the needed parameters. In this direction, Wolbach and Sandler¹¹⁴ used the Hartree–Fock theory and the density functional theory (DFT) to calculate the enthalpy and entropy of dimerization for water, methanol, and carboxylic acids. Furthermore, molar volumes from molecular orbital calculations were related to the segment size and chain length parameters of SAFT. The results were used in the SAFT EoS to model the phase behavior of the aforementioned pure compounds¹¹⁴ and their mixtures with nonassociating components.¹¹⁵ In a subsequent publication, Wolbach and Sandler,¹¹⁶ using molecular orbital calculations, estimated the cross-association parameters in mixtures of associating compounds. They also developed a mixing rule, which allows the estimation of the cross-association SAFT parameters from the values of the self-association parameters of the species in the mixture. Using such mixing rule, they were able to satisfactorily correlate vapor–liquid equilibrium data for mixtures of associating compounds.

p0885 Singh et al.¹¹⁷ performed quantum chemical calculations to determine molecular parameters, such as the dipole moment, the quadrupole moment, the dipole polarizability, and the dipole–dipole dispersion coefficient. Such molecular parameters were used in the polar PC-SAFT framework to calculate properties of pure fluids and mixtures.¹¹⁸ Pure compound parameters, which were not adopted from quantum mechanical calculations, were adjusted to experimental data according to the usual procedure, that is, by adjusting model predictions to experimental data. Such quantum chemical calculations were also used in order to evaluate and improve the treatment of multipolar interactions in a PC-SAFT-based EoS.¹¹⁹ In another publication of the same group, molecular descriptors for sizes, shapes, charge distributions, and dispersion interactions were computed for several compounds using quantum chemical ab initio and DFT calculations and were correlated with the remaining three parameters of a polar PC-SAFT EoS.¹²⁰ The three pure compound parameters, which were not determined directly by quantum chemistry, were adjusted to experimental vapor pressure, saturated liquid density, and critical property data. Finally, such EoS parameters, obtained from quantum mechanics, were used to estimate the surface tensions of pure fluid vapor–liquid systems.¹²¹

p0890 Finally, Lucia et al.¹²² presented a new algorithm for estimating association parameters within molecular-based EoS (and using the simplified SAFT-based equation as an example) that allows the incorporation of quantum chemical calculations.

s0075 GC methods

p0895 The GC approach is by far the most popular approach in predicting EoS parameters. In this way, EoS models become fully predictive and their application range is expanded to systems with little or none of experimental data. Within the SAFT framework, two categories of group-contribution methods were mainly developed, that is, those that refer to homonuclear and those that refer to heteronuclear models. The first category includes most SAFT-type models, in which all of the segments making up the molecular chain are identical, while in the second category the segments in a given molecule are arbitrary different.⁹³ Heteronuclear molecular models account for varying sizes and interactions of segments that constitute the molecules. Such segments can be assigned to functional groups and, therefore, a GC method can be incorporated into the SAFT approach. Consequently, heteronuclear SAFT-type models are the base for promising predictive approaches and, for this reason, attracted particular interest during the past years.¹²³ A review of the homonuclear and heteronuclear approach is presented elsewhere.⁹³

p0900 Various GC approaches have been developed through the SAFT framework and were reviewed in recent journal articles and book chapters.^{31, 93, 123–125} Some of the most popular approaches are presented in Table 3.

s0080 Computational chemistry models—The conductor-like screening model approach

p0905 All predictive models that were presented in the previous sections, such as activity coefficient approaches, group-contribution EoS, and EoS/ G^E models, require the use of experimental data to obtain the needed molecular or group-specific parameters. Furthermore, as it was mentioned in the previous section, despite the effort to predict EoS parameters from quantum mechanics, such

t0020 **Table 3** GC approaches developed for SAFT-type models

SAFT variant	References
SAFT	Lora et al. ¹²⁶ Tamouza et al. ¹²⁷ Tamouza et al. ¹²⁸
PC-SAFT	NguyenHuynh ¹²⁹ Vijande et al. ¹³⁰
sPC-SAFT	Tihic et al. ¹³¹ Tihic et al. ¹³²
SAFT-VR	Tamouza et al. ¹²⁷ Tamouza et al. ¹²⁸
GC-SAFT-VR	Peng et al. ¹³³
SAFT- γ	Lymperiadis et al. ¹³⁴ Lymperiadis et al. ¹³⁵
SAFT- γ Mie	Papaioannou et al. ¹²⁵
hs-PC-SAFT	Paduszyński and Domanska ¹²³

approaches have not yet resulted in a purely predictive model, that is, they only reduce the number of adjustable parameters, but, still, some of them are obtained using experimental data.

p0910 So far, the only models that allow the prediction of thermodynamic properties without needing any experimental data for the description of real fluids and mixtures are those that are based on the CONductor-like Screening MOdel (COSMO) approach.¹³⁶ The most widely used advanced model of this family is the COSMO-RS (COSMO-RS stands for CONductor-like Screening MOdel—for Real Solvents).^{137–139} The ultimate advantage of such model is that it allows predictions of thermodynamic properties using only data from quantum chemical calculations.

p0915 The COSMO-RS model was developed by Klamt and coworkers,^{137–139} who transformed the COSMO solvation approach in a thermodynamic model for real fluids through a novel thermodynamic treatment of the electrostatic interactions as intermolecular interactions and by approximating the screening charges of a dielectric medium of permeability ϵ with the scaled screening charges of a conductor.^{136–139} After the pioneering work by Klamt, various modifications of the COSMO-RS model appeared in the literature.^{140,141}

p0920 The ultimate advantage of such models is that they allow prediction of thermodynamic properties using only data from quantum chemical calculations. Thus, they do not require experimental data, but only use as an input the so-called sigma profiles (σ -profiles) of the interacting molecules. The σ -profile is the probability distribution of a molecular surface segment having a specific charge density. Knowing only the molecular structure, such data are generated through quantum mechanical calculations. Such advantage rendered COSMO-RS into a powerful predictive tool for phase equilibria and other thermodynamic properties. However, despite the aforementioned significant advantage of the model, COSMO-RS presents certain limitations, which are based on the fact that it is an activity coefficient model of mixtures. Thus, it can describe the liquid phase, but cannot account for the supercritical state, for high-pressure and high-temperature vapor–liquid equilibrium and for the volume changes in mixing. In order to overcome such limitations, effort has been made in order to develop EoS models based on the COSMO approach.^{142,143}

p0925 Generating the σ -profiles by quantum mechanical calculations is the most time-consuming and computationally expensive aspect of using COSMO-based methods. However, today, extensive σ -profile databases are available, which include thousands of known molecules.^{144,145} Nevertheless, in order to avoid the time-consuming step of performing quantum mechanical calculations in particular computer-aided design problems, various GC approaches appeared in the literature. In this direction, Mu et al.^{146,147} developed a GC method for predicting the surface charge density distribution (σ -profile) for the COSMO-RS(OI)¹⁴⁸ and the COSMO-SAC¹⁴⁰ models, which are modifications of the original COSMO-RS model developed by Klamt and coworkers.^{137–139} Furthermore, Austin et al.¹⁵¹ developed a GC method to estimate σ -profiles, based on the Virginia Tech's Sigma Profile Database. Such σ -profiles were obtained using the parameters of the COSMO-SAC model.¹⁴⁵

s0085 Main Classes of CAMD Problems

s0090 Classification Overview

p0930 Molecular design requires a method to enumerate the combinations which can be generated from a given set of molecular fragments (e.g., atoms or functional groups) and a GC model to predict their properties. In the most general case, molecular design is approached through the combination of GC models with optimization algorithms, exploiting the reverse problem formulation discussed in the introduction. The optimum molecular structure can be identified by formulating an optimization problem that emulates an automated molecular synthesis process (i.e., the iterative transformation and evolution of an initial structure). An optimization algorithm guides the synthesis toward optimum structures and employs performance indicators either as objective functions or as constraints representing molecular or application-related performance features.¹⁵⁰ The classification of molecular design problems can be based on whether it is desired to:

- u0025 - Consider single or multiple performance criteria.
- u0030 - Determine only optimum molecular structures or integrate process structural and operating decisions.
- u0035 - Design single or multiple components in the form of mixtures or blends.
- u0040 - Account for uncertainty in the employed models or internal and external process or application conditions.
- u0045 - Consider inert or reactive systems.
- u0050 - Include models which directly account for matter behavior at the atomic or molecular level.

p0965 Considering the earlier points together with computer-aided molecular (and process) design approaches proposed in the scientific literature to date, molecular design problems are classified as follows:

- o0020 1. Single-objective optimization computer-aided molecular design, referred to as SOO CAMD
- o0025 2. Computer-aided molecular and process design, referred to as CAMPD
- o0030 3. Multiobjective optimization computer-aided molecular design, referred to as MOO CAM(P)D
- o0035 4. Computer-aided mixture design, referred to as CAM^b(P)D
- o0040 5. Computer-aided molecular design under uncertainty, referred to as CAM^u(P)D
- o0045 6. Computer-aided molecular, process, and control design, referred to as CAMPCD
- o0050 7. Computer-aided molecular design for reactive systems, referred to as CAM^R(P)D
- o0055 8. Computer-aided molecular design using quantum mechanical models, referred to as QM-CAM(P)D

p1010 The order of the proposed classes either considers the chronological order by which corresponding works appear in published literature or facilitates the elaboration of reasons that lead to the adoption of the corresponding approaches. Note that the different problem classes may be used in various combinations. For example, MOO, SOO, or mixture/blend design formulations may be combined with any of the other classes. Furthermore, the brackets indicate that process design has also been considered in relevant publications or could be considered in the proposed class. The terms CAM^bD and CAMPD are adopted from Gani¹⁵¹ and Pereira et al.,¹⁵² respectively. The CAM^bD problem is also reported as product design¹⁵³ as well as CAMxD.¹⁵⁴ Blend design is defined as a subclass of mixture design (see "Summary of Classes" section). CAM^R(P)D refers to the use of CAMD for identification of molecules which promote chemical reactions or for the identification of molecules produced through reactions which are also identified as part of the design procedure. QM-CAM(P)D, adopted from Struebing et al.,¹⁵⁵ refers to the incorporation of quantum mechanical models in CAMD for the prediction of properties pertaining to both physical and chemical molecular characteristics. The following sections elaborate on the earlier-mentioned classes.

s0095 Single-Objective Computer-Aided Molecular Design

p1015 The first step to the mathematical formulation of a CAMD problem is the definition of a molecular structure as follows: let a vector of groups $\mathbf{m} = [g_l]_{l=1}^{N_g}$, with N_g representing the total number of available groups and a composition matrix $A_k = \text{diag}(n_l^k)_{l=1}^{N_g}$, which indicates the number of occurrences n_l^k of each group l for a k th molecule.¹⁵⁶ A molecular vector \mathbf{M}_k representing a k th molecule is defined as follows:

$$\mathbf{M}_k = A_k \cdot \mathbf{m} \quad (13)$$

where the composition matrix A_k contains information on the number of occurrences of each group.

p1020 The optimization problem can now be formulated as follows: Let a vector of N_{of} objective functions $\mathbf{F} = [F_j]_{j=1}^{N_{of}} \in \mathbb{R}^{N_{of}}$. Also let a vector of N_d design variables $\mathbf{D} = [d_i]_{i=1}^{N_d} \in \mathbb{R}^{N_d}$ with $\mathbf{M} \in \mathbf{D}$ and let $\mathbf{h}(\mathbf{X}, \mathbf{D})$, $\mathbf{q}(\mathbf{X}, \mathbf{D})$ be vectors of equality and inequality constraints, with \mathbf{X} being a vector of state variables. Further, let a total number of N_M molecules with $k \in [1, N_M]$ and a vector $\mathbf{C} = [c_k]_{k=1}^{N_M}$ representing the concentration of the components (as mass or mole fractions) in the general case of multiple components or of a mixture. For the case of designing a single molecule with a single objective, it holds that $N_M = 1$ and $N_{of} = 1$. The problem takes the following form:

$$\underset{\mathbf{D}}{\text{optimize}} F_1(\mathbf{X}, \mathbf{D}) \quad (14)$$

p1025 Subject to

$$\mathbf{h}(\mathbf{X}, \mathbf{D}) = 0 \quad (15)$$

$$\mathbf{q}(\mathbf{X}, \mathbf{D}) \leq 0 \quad (16)$$

$$\mathbf{X}^L \leq \mathbf{X} \leq \mathbf{X}^U \quad (17)$$

$$\mathbf{D}^L \leq \mathbf{D} \leq \mathbf{D}^U \quad (18)$$

p1030 As an example, assuming a set of five functional groups, that is, $N_g = 5$, vector \mathbf{D} would contain the following design variables: $d_1 = k = 1$, $d_2 = c_1 = 1$, $d_3 = n_1^1$, $d_4 = n_2^1$, $d_5 = n_3^1$, $d_6 = n_4^1$, $d_7 = n_5^1$. In this case, c_k is constant and equal to 1, because the equality vector contains the constraint $\sum_{k=1}^{N_C} c_k - 1 = 0$. Apart from d_1 and d_2 which are constant, the general case entries in \mathbf{D} for the molecular structure are $d_i = n_l^1 \quad \forall l \in [1, N_g], i \in [3, N_g + 2]$. The only decision pertains to the structure of each component, that is, the total number of design variables is N_g . Vector \mathbf{X} could represent molecular properties, while a desired molecular property could also be part of vector \mathbf{F} . The equality and inequality constraints may represent the employed molecular models that ensure feasibility or complexity of the proposed structures, and design or performance constraints, to name but a few. The indices L and U represent upper and lower bounds utilized for all the variables. Note that the earlier-mentioned formulation remains generic in order to introduce the main concepts of CAMD. Depending on the optimization approach, additional design variables may be needed, such as binary variables to model the presence or absence of groups or bonds etc.^{157,158}

p1035 According to Sioukrou,¹⁵⁹ constraints pertaining to molecular structures are based on two main classes:

- u0055 - Molecular feasibility ensuring that two adjacent groups are linked by one bond and the resulting molecule has zero valency.
- u0060 - Molecular complexity imposing upper and lower limits on the number of same type groups and on the total number of groups in a molecule.

p1050 Molecular feasibility is based on the valence of groups v_l^k , that is, the number of free bonds in each group. The most widely used constraint is proposed by Odele and Macchietto¹⁶⁰:

$$\sum_{l=1}^{N_g} (v_l^k - 2)n_l^k = 2m \quad (19)$$

where m takes the values of -1 for acyclic, 0 for monocyclic, and 1 for bicyclic compounds. This constraint refers to GC-based molecular representations employing functional groups. Different constraints are needed in the case of other descriptors, such as topological indices and signature descriptors. A detailed discussion regarding various types of connectivity constraints is reported in Austin et al.⁴

s0100 Computer-Aided Molecular and Process Design

p1055 Molecules are generally designed to serve a particular application and are often used as intermediate materials in processes. For example, they are used as solvents to facilitate the separation of components in the chemical industry, or as heat exchange fluids in refrigeration and other thermodynamic cycles, to name but a few. In this respect, they are inherent components of the processes where they are utilized, in the same way that different types of equipment are part of such processes. The chemical and physical characteristics of such materials affect directly the operating, economic, and sustainability properties of the corresponding processes, due to their thermodynamic, kinetic, and other links. Changes in such characteristics due to the use of different materials also necessitate appropriate modifications in the process equipment so that the overall process performance remains optimum. Since multiple properties of the molecule(s) to be designed influence one or more process performance objectives such as cost or environmental impact, it is difficult to develop adequate objectives for molecular design to identify those molecules that maximize process performance. In order not to develop molecules that lead to suboptimal process performance, it is advisable to integrate molecular with process design approaches.

p1060 Integrated molecular and process design can be approached in a CAMPD problem formulation, as follows:

p1065 “For given molecular and process properties representing desired performance targets, identify the molecular and process structural and operating characteristics that best match these targets.”

p1070 The CAMPD problem formulation changes the nature of the constraints in Eqs. (14)–(18). The equality and inequality constraints also include the process models, whereas the objective function may include economic-, operating-, and sustainability-related terms. Assuming a vector of $\mathbf{P} = [p_q]_{q=1}^{N_p}$ process-related design variables, a total number N_p of additional design variables will be added, hence the total number of design variables will be $N_g + N_p$.

s0105 MultiObjective Computer-Aided Molecular Design

p1075 The development of CAM(P)D approaches using the SOO formulation of (14)–(18) introduces important challenges and assumptions. Whether in CAMD or CAMPD, there might be several different criteria that need to be considered simultaneously. For example, in CAMPD, optimum sustainability targets may need to be achieved in addition to high economic performance. Formulations (14)–(18) indicate that one of the desired criteria should assume the role of the objective function that is optimized. The remaining criteria are restrained to the auxiliary role of constraints, providing performance limits that either are desired to be achieved or must not be violated. Papadopoulos and Linke¹⁶¹ showed that the selection of different criteria as F_1 in Eq. (14) (and the solution of the corresponding CAMD problem) results in different optimum molecules. Furthermore, the use of the remaining criteria as constraints requires the selection of upper and/or lower bounds representing performance limits. Such a selection is unwarranted as there is often no prior knowledge of these limits. In this context, design drives are poorly represented or misrepresented and the results may be biased toward presumably optimal options.

p1080 The earlier-mentioned challenges may be addressed through a multiobjective formulation of the CAM(P)D problem as follows:

$$\underset{\mathbf{D}}{\text{optimize}} F_1(\mathbf{X}, \mathbf{D}), \dots, F_{N_{of}}(\mathbf{X}, \mathbf{D}) \quad (20)$$

p1085 Subject to

$$\text{constraints (15 – 18)} \quad (21)$$

p1090 In the case of a multiobjective problem, the term “optimize” implies the possibility of simultaneous minimization, maximization, or both, of all objective functions. To identify a set of optimum solutions, formulations (20) and (21) need to account for an additional condition. A feasible point \mathbf{D}^{opt} is called a Pareto-optimum or nondominated solution *iff* there exists no other point \mathbf{D}^* satisfying the following condition¹⁶²:

$$\mathbf{F}(\mathbf{D}^*) \leq \mathbf{F}(\mathbf{D}) \wedge \exists j \in \{1, \dots, N_{of}\} : F_j(\mathbf{D}^*) < F_j(\mathbf{D}) \quad (22)$$

p1095 In the context of Eq. (22), the generated designs are evaluated in terms of optimality based on comparison of the objective function values representing one solution in the nondominated set with the objective function values of the other solutions contained in the set. The resulting Pareto or nondominated front indicates trade-offs among the objective functions. It shows how much performance needs to be sacrificed in one objective function in order to gain performance in another. Additional implementation details can be found in Papadopoulos and Linke¹⁶¹ and Papadopoulos et al.¹⁶² Multiobjective molecular design is a key enabler to decompose integrated molecular and process design problems as it allows introduction of a reduced (Pareto-optimal) set of molecules into computationally demanding optimal process synthesis and design efforts.¹⁶³

s0110 Computer-Aided Mixture Design

p1100 Mixtures are often preferred in industry when there is no single compound which can achieve the desired behavior in all the necessary criteria. For example, in reactive separation systems used for CO₂ capture, it is necessary to use a solvent which exhibits fast kinetics while simultaneously dissolving as much CO₂ as possible.¹⁶⁴ Most solvents considered to date either exhibit fast kinetics at the expense of low solubility toward CO₂ or vice versa. With the use of a mixture instead of a single solvent, this challenge may be addressed efficiently by incorporating two or more solvents which exhibit desired behavior in both kinetics and CO₂ solubility.¹⁶⁵

p1105 In the CAM^b(P)D problem, it is necessary to determine the optimum:

- o0060 (a) number of components k in the mixture,
- o0065 (b) structure of all components represented by n_i^k ,
- o0070 (c) concentration c_k of each component in the mixture,
- o0075 (d) characteristics p_i of the process in which the mixture is utilized.

p1130 In MOO CAM^b(P)D, formulations (20)–(21) remain the same. The total number of design variables equals to $N_M(N_g + 1) + 1 + N_p$. For $N_{of} = 1$ and $N_M = 1$, the problem formulation is reduced to a single-objective problem for the design of one component (see “Single-Objective Computer-Aided Molecular Design” section).

s0115 Computer-Aided Molecular Design Under Uncertainty

p1135 In any type of CAM(P)D formulation, the proposed molecular and/or process structures may deviate from their actual performance, when implemented in practice. This is because the employed models are approximations of the real molecular behavior which they are expected to emulate. The design procedure is therefore subject to uncertainty introduced in predictions at:

- u0065 ● The molecular level, where GC or other models are used to emulate molecular behavior.
- u0070 ● The molecular-to-process level, where thermodynamic mixture, kinetic or other models are used to deliver molecular chemistry predictions at various conditions, often relevant to process requirements.
- u0075 ● The process level, where unit operation models of different fidelity may be used to determine desired molecular and process characteristics.

p1155 Uncertainty practically means that the use of two or more different models for the prediction of the same property may result in different values. If CAM(P)D is implemented independently in two or more iterations, using a different model for the same property in every iteration, then it will result in different selected molecules.^{164–166} This is clearly undesired and should be avoided. In addition to model-related uncertainty, molecular design results will also be affected due to external or internal variability in the process system or application where the molecule is used.

p1160 With the aim to address these challenges, CAM^u(P)D involves a reformulation of (14)–(18) as follows¹⁶⁷:

$$\underset{\mathbf{D}}{\text{optimize}} P_1[\mathbf{F}(\mathbf{X}, \mathbf{D}, \mathbf{U})] \quad (23)$$

p1165 Subject to

$$P_2[\mathbf{h}(\mathbf{X}, \mathbf{D}, \mathbf{U}) = 0] \quad (24)$$

$$P_3[\mathbf{q}(\mathbf{X}, \mathbf{D}, \mathbf{U}) \leq 0] \quad (25)$$

$$\mathbf{X}^L \leq \mathbf{X} \leq \mathbf{X}^U \quad (26)$$

$$\mathbf{D}^L \leq \mathbf{D} \leq \mathbf{D}^U \quad (27)$$

$$\mathbf{U}^L \leq \mathbf{U} \leq \mathbf{U}^U \quad (28)$$

where \mathbf{U} is a vector of uncertain parameters. Vectors \mathbf{F} , \mathbf{h} , and \mathbf{q} now depend on the values of the uncertain parameters which should be different than the values predicted by the employed models, that is, the nominal values, in order to quantify the probability of obtaining different predictions. The different potential realizations of uncertain parameters can be quantified mathematically using probability density distributions.¹⁶⁸ This probability is formally represented through functions P_1 , P_2 , and P_3 which are cumulative distribution functionals such as the expected value, variance, etc.

s0120 Computer-Aided Molecular, Process, and Control Design

p1170 Using a CAM^uPD formulation to emulate variability in process operation would result in optimum molecular and process features enabling flexible operation. This means that the selected molecule and associated process features would be able to handle alternate operating conditions under steady-state operation (e.g., to avoid the need for different solvent or process equipment when it is desired to operate at different temperatures and flowrates). However, they would not be able to support efficient controllability, that is, tracking of system changes and recovery from upsets or disturbances that persistently influence its operation, under dynamic

conditions. Papadopoulos and Seferlis¹⁶⁹ and Papadopoulos et al.¹⁶² show that materials exhibiting optimum economic performance under steady-state conditions are likely to underperform at the influence of disturbances. This is shown in Fig. 2.

p1175 Each point in Fig. 2 represents one solvent and process scheme specifically designed to enable optimum economic performance under nominal conditions (i.e., steady-state). It is clear that solvents and processes represented by a point such as S2, which exhibits optimum nominal economic performance (i.e., lowest cost), have very poor controllability (i.e., very high controllability index value). This translates to a solvent and process scheme that:

o0080 (a) Operates very far from the economically optimum set-points at the presence of disturbances.

o0085 (b) Needs to consume a very large amount of resources (e.g., increase significantly the use of costly heating utilities) to bring the system back to each set-point.

p1190 On the other hand, point S5 exhibits the best controllability, but it represents an expensive solvent–process system to use at nominal conditions. Solvent–processes represented by S3 and S4 exhibit the best compromise. Without this analysis, solvent–process S2 would have been chosen with detrimental effects on process operation at the presence of disturbances and possibly requiring an expensive control system to reinstate the process back to its desired operation.

p1195 The identification of molecular and process characteristics of optimum performance in view of disturbances can be formulated as a CAMPCD problem as follows:

p1200 “For given molecular and process properties representing desired performance targets, operating disturbance scenarios representing conditions other than the nominal settings and associated molecular and process model data, identify the molecular and process structural and operating characteristics as well as the control structure(s) that best match these targets.”

p1205 In mathematical terms, the CAMPCD problem is formulated as follows:

$$\text{optimize}_{D,W} F(X, D, W, Y, E) \quad (29)$$

p1210 Subject to

$$h(X, D, W, Y, E) = 0 \quad (30)$$

$$q(X, D, W, Y, E) \leq 0 \quad (31)$$

$$X^L \leq X \leq X^U \quad (32)$$

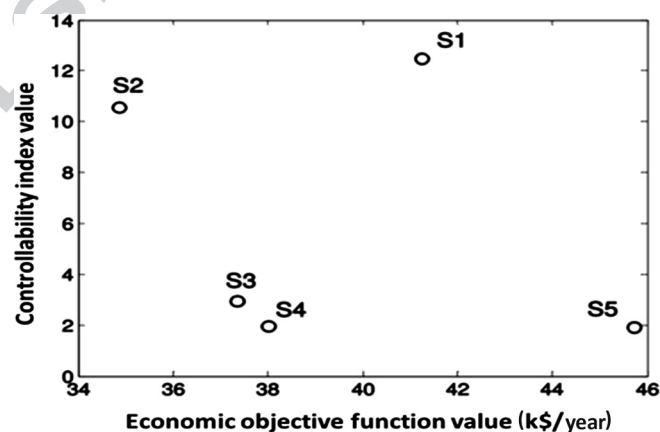
$$D^L \leq D \leq D^U \quad (33)$$

$$Y^L \leq Y \leq Y^U \quad (34)$$

$$W^L \leq W \leq W^U \quad (35)$$

$$E^L \leq E \leq E^U \quad (36)$$

where $E = [e_{i1}]_{i=1}^{N_e}$ is a vector of disturbances, $W = [w_{i2}]_{i=1}^{N_w}$ is a vector of manipulated variables, and $Y = [y_{i3}]_{i=1}^{N_y}$ is a vector of controlled variables. Notice that in Eq. (29), it is now required to determine optimum values for the manipulated variables too. Furthermore, both the manipulated and controlled variables affect the objective function; the former may be used to show the resources that need to be used in order to bring the system back to its set-point, while the latter shows the distance from the set-point. Both have a direct impact on the process economics.



r0015 **Fig. 2** Pareto front of economic performance of different solvent–process schemes under steady-state conditions (*horizontal axis*) versus controllability performance at the presence of disturbances.¹⁶⁹ Reproduced from Papadopoulos, A.I.; Seferlis, P., *Comput.-Aided Chem. Eng.* **2009**, *26*, 177–181 by permission of Elsevier.

s0125 Computer-Aided Molecular Design for Reactive Systems

p1215 This class pertains to the use of CAMD for the design of molecules which either promote chemical reactions or participate in chemical reactions as raw materials or products. With respect to reaction promoters, the kinetic characteristics of chemical reactions can be affected by the presence of materials such as solvents or catalysts which may change the rate constant and other properties. Adjiman et al.¹⁷⁰ provide an example of how a solvent may affect a ring-closing metathesis reaction by changing the rate constant, the solubility, and the deactivation of the catalyst. A solvent like acetone increases the rate of reaction but facilitates the catalyst deactivation hence hindering complete conversion. Dichloromethane enables complete conversion as it avoids catalyst deactivation, but the reaction rate is low. Cyclohexane enables high reaction rate and complete conversion at the expense of low catalyst solubility. In this context, CAM^R(P)D includes approaches to identify molecular structures which aim to strike an optimum balance among such requirements.

p1220 With respect to CAMD of products, it is possible to design reaction products which exhibit favorable properties for a potential application (e.g., biofuels with better ignition properties) but also enable production from reaction pathways that exhibit favorable characteristics compared to existing production routes (e.g., their technological implementation is more economic). CAMD may be used to design raw materials which facilitate the production of certain products through reaction synthesis pathways that are favorable compared to existing ones.

s0130 Computer-Aided Molecular Design Using Quantum Mechanical Models

p1225 QM models are used before, during, or after CAMD for the prediction of desired properties in reactive or nonreactive systems. In the case of reactive systems, QM models are used in order to predict properties related to promotion of reactions or to develop simpler models for properties such as the reaction rate constant, as they are able to describe the formation and breaking of bonds as well as the transfer of electrons,¹⁷¹ to name but a few. At the same time, properties related to phase transitions in nonreactive systems can also be predicted. QM-CAM^R(P)D and QM-CAM(P)D are therefore emerging as approaches which use atomic- or molecular-scale models to predict properties, such as activation energies, electronic charge distributions, dipole and higher moments, vibrational frequencies,¹⁷² etc., as well as free energy difference between two species in solution and reactive trajectories to determine reaction rate constants.¹⁷¹ Such properties are generally beyond the reach of standard thermodynamic models and EoS.³ QM refers to the model used to predict molecular properties, hence it can be adapted to all types of the CAMD problem classes described in the previous sections.

s0135 Summary of Classes

p1230 This section summarizes the main molecular design problem classes in Table 4 and discusses few subclasses which result from them. The latter include the following:

p1235 *Optimum molecular selection*: This class may result from formulations (20)–(21) when $N_M > 1$ and k is allowed to vary during optimization within $[1, N_M]$, but n_i^k obtains constant, prespecified values for every molecule k . This is the case of *optimum molecular selection* from a set of structurally prespecified candidates. In every optimization iteration one molecular structure is selected as a discrete option and evaluated in terms of F . It should also hold that for every evaluated molecule, $c_k = 1$. This formulation could be used to screen a large database of molecules using an optimization approach.

p1240 *Feasibility problem*: In the reported objective function formulations, it is assumed that F will vary as a function of D . An additional subclass reported in Austin et al.⁴ is based on the assumption of a constant F . This is called a *feasibility problem*, implying that a large set comprising numerous different molecular structures can be generated by enumerating parameters in D , and it can then be reduced to fewer options that satisfy the constraints.

p1245 *Continuous molecular representation*: Another subclass includes the case where the discrete n_i^k is replaced by a continuous representation of a molecule. This representation was proposed by Macchietto et al.¹⁷³ and Naser and Fournier¹⁷⁴ and originally maintained the use of UNIFAC groups, in mathematical formulations that allowed their representation through continuous variables. The representation was later approached through size- and attraction-related characteristics pertaining to the use of SAFT-type EoS¹⁷⁵ in the continuous molecular targeting (CoMT)-CAMPD formulation or through pure component properties used in cubic EoS.¹⁷⁶

p1250 *Proximity-to-property-targets*: This subclass uses as F the distance between the properties of the designed molecule and desired property targets.⁴ A similar approach is used by Bardow et al.¹⁷⁵ to match the identified set of continuous parameters with a feasible or real molecular structure.

p1255 *Blend design*: In formulations (20)–(21), if c_k is varied but n_i^k remains constantly at prespecified values, the structure of the mixture components is predetermined. This may be combined with a case where the number of components k is constant or variable. In this case, the problem is called *blend design* and the aim is to find the optimum concentration and/or number of components for a mixture containing candidates with prespecified structures. We therefore propose that the determination of the structure of the components is the necessary condition to differentiate between this class and CAM^b(P)D problems.

p1260 *Partial mixture design*: A different case may include variations of c_k for prespecified values of n_i^k for only one of the mixture components. In this case, the structure of the other component will result from optimization together with the optimum mixture concentration. This case qualifies as CAM^b(P)D too.

t0025 **Table 4** Problem classes and main characteristics for single-objective optimization (S00) cases (when $N_{of} = 1$) and multiobjective optimization (MOO) cases (when $N_{of} > 1$)

k	c_k	n_i^k	p_q	Problem class
1	1	Const.	Const.	Process simulation for prespecified single molecule
1	1	Const.	Var.	S00 or MOO process design for prespecified single molecule
1	1	Var.	Const.	S00 or MOO CAMD
1	1	Var.	Var.	S00 or MOO CAMPD
>1	Const.	Const.	Const.	Process simulation for prespecified mixture
>1	Const.	Const.	Var.	S00 or MOO process design for prespecified mixture
>1	Const.	Var.	Const.	S00 or MOO CAM ^b D with fixed mixture concentration
>1	Const.	Var.	Var.	S00 or MOO CAM ^b PD with fixed mixture concentration
>1	Var.	Const.	Const.	S00 or MOO process simulation and blend design with prespecified mixture composition (structures)
>1	Var.	Const.	Var.	S00 or MOO process and blend design with prespecified mixture composition (structures)
>1	Var.	Var.	Const.	S00 or MOO CAM ^b D
>1	Var.	Var.	Var.	S00 or MOO CAM ^b PD

Var. and Const. are abbreviations for "varied" and "constant."

p1265 The classes of Table 4 are based on the number of objective functions N_{of} , the number of components k , the concentration of components c_k , the composition (structure) of components determined by the occurrences of groups n_i^k in the molecules, the state of process-related design variables p_q , and the resulting number of design variables. Note that considering the previous criteria, the proposed classification may be repeated for CAM^u(P)D, CAMPCD, CAM^R(P)D, QM-CAM(P)D, and QM-CAM^R(P)D.

s0140 General Solution Approaches

s0145 Challenges in the Solution of Molecular Design Problems

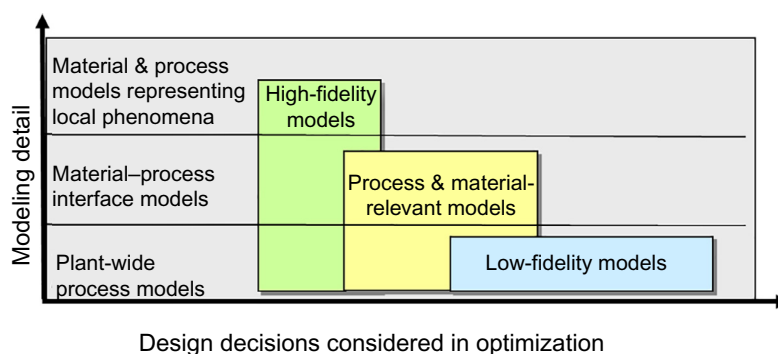
p1270 In any class of molecular design problems, the need to obtain optimum solutions that are also sufficiently reliable to ideally enable direct practical implementation imposes two requirements: (a) the consideration of numerous design decision options and (b) the incorporation of high fidelity models during optimization. These requirements are illustrated conceptually in Fig. 3 in terms of general model types and corresponding decision ranges.

p1275 Requirement (a) implies large combinatorial complexity which increases as more options are considered in different problem classes. Requirement (b) pertains to the ability to solve nonconvex and large-scale systems of equations, which is challenged as more detail is added in the models in order to obtain more realistic and accurate results. With respect to requirement (b), there are three general types of models for both the molecular and process levels as shown in Fig. 3. High fidelity models provide a very detailed representation of phenomena. However, it becomes less practical to consider more molecular and/or process combinations because of limitations due to computational effort. With low-fidelity material and process models, it is possible to consider a very large number of decision options as they enable very fast calculations (e.g., pure component GC models). These models are used in combination with optimization methods and provide useful insights into economics and design solutions that exhibit potential for further development. However, due to simplifying assumptions in phenomena representation, the obtained results require further treatment with more refined models prior to implementing design solutions in practice.

p1280 Process-relevant material and equipment/plant models may provide a reasonable compromise between modeling fidelity, range of decision options, and computational intensity. EoS, activity coefficient, and/or (semi-)empirical kinetic models are typical representatives of models that can be used for process-level evaluation, as they enable the prediction of key properties for process design. At the process side, equilibrium- or even rate-based models can be used to account for the operating characteristics of all equipment, whereas superstructure approaches¹⁷⁸ can be used with such models to account for all possible stream connections used in complete flow sheets that can determine the economic behavior of different technologies. However, the utilization of such models in CAMD problems is not without significant computational challenges.

s0150 Classification of Solution Approaches

p1285 In the published literature, there are currently few general classifications of the approaches used to solve molecular design problems. For example, Austin et al.⁴ propose the generate and test, decomposition, mathematical, and heuristic approaches. Decomposition indicates the solution of the problem in several steps, but the mathematical and heuristic classes indicate a dependence on the type of the employed solution algorithm used. As noted by the authors, many mathematical optimization methods include decomposition, hence there are overlaps. Papadopoulos et al.¹⁵⁰ propose a classification into simultaneous and integrated approaches, addressing only the CAMPD problem. In this case, the simultaneous class implies the solution of the problem within algorithmic steps executed together in one overall execution and may include problem decomposition methods to gradually reduce the optimization problem size. The integrated class implies a clearly decomposed approach which is executed in computationally



0020 **Fig. 3** Modeling detail versus range of design decisions. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. (CC BY 4.0). Adapted from Linke, P.; Papadopoulos, A.I.; Seferlis, P., *Energies* **2015**, *8*(6), 4755–4801.

independent steps, and there is a systematic way of extracting and transferring information between steps. Finally, Gopinath¹⁷⁹ proposes the direct approaches which solve the problem without any size reduction and the screening approaches which implement a decomposition scheme with size reduction.

p1290 These different classifications can be reconciled under a classification which includes direct and decomposition approaches.

u0080 - *Direct approaches* address the solution of the complete optimization problem without any intermediate design space reduction.

u0085 - *Decomposition approaches* include the formulation and successive solution of reduced-size subproblems, with adjacent transfer of information, aiming to facilitate the identification of an optimum solution. The subproblems may be computed simultaneously within the same software implementation or in separate steps.

p1305 Note that the earlier-mentioned classification focuses mainly on optimization-based CAMD. Generate and test approaches (see “*Generate and Test Approaches*” section) could be considered as part of both classes, depending on the way they are implemented. If the optimum molecule is selected by complete enumeration and evaluation of all the feasible structures, then it could be considered as part of the direct approaches. If the original set of all enumerated structures is reduced based on gradual addition and satisfaction of different constraints, then it is part of the decomposition approaches. It should be noted that the term “direct approach” does not refer to the numerical solution scheme of MINLP, that is, to the way that the algorithm handles the problem and objective function, as defined in Austin et al.¹⁴⁹

p1310 Direct approaches have the advantage that all design parameters of the problem are considered simultaneously, hence no design realization is excluded from the optimization search. Such approaches often deal with intensity in computations through appropriate formulation of the optimization problem structure⁴ or through models of lower fidelity. However, numerical difficulties may appear which render the problem intractable, while the reduction in fidelity reduces the confidence in the obtained results.

p1315 Decomposition approaches generally reduce the computational complexity of the problem by gradually removing candidate solutions which fail to satisfy increasingly stricter constraints in different subproblem formulations. At the same time, the remaining candidate solutions are introduced and evaluated in other subproblems of different orientation. Decomposition approaches have been used widely in CAMPD problems. For example, one approach is to identify good candidates in the molecular design space and then transfer them to process optimization to evaluate their performance. Decomposition approaches focus essentially on the development of methods that facilitate the identification, extraction, and exploitation of inclusive design information among subproblems, without prematurely excluding potentially useful options. The latter cannot be guaranteed, as in the case of direct approaches. However, an approach that starts from subproblems of high combinatorial complexity using low-fidelity models and gradually proceeds to subproblems of lower combinatorial complexity using higher-fidelity models could produce practically useful and meaningful design results. Solving these subproblems using a direct approach could be efficient and productive.

s0155 **Generate and Test Approaches**

p1320 Generate and test approaches include the enumeration of molecular structures that result from combinations of molecular groups and the subsequent evaluation of their properties. The main stages of the “generate” step include the group selection, group characterization, and molecular feasibility rules, to support the synthesis of feasible molecular structures. The main stages of the “test” step include the GC methods for property estimation, calculated properties, property constraints, and evaluation (performance indices).¹⁸⁰ The procedure results in the generation of a ranked list of molecular structures.

p1325 The approach is first proposed by Gani and Brignole,² where they highlight the issues that should be taken into account, especially with respect to the problem combinatorial complexity. Brignole et al.¹⁸¹ provide a formalization of the proposed

framework with improved procedures for solvent synthesis, solvent evaluation, and solvent screening. Pretel et al.¹⁸² propose a framework for addressing combinatorial complexity through group reduction rules. These include consideration of interaction parameter availability, elimination of structural and functional isomers, and elimination of groups that affect the chemical or physical state of the solvents. The work is later extended by Cismondi and Brignole¹⁸³ to account for branched molecules. Joback¹⁸⁴ and Joback and Stephanopoulos¹⁸⁵ also propose procedures for reduction of combinatorial complexity through chemical feasibility constraints. Derringer and Markham¹⁸⁶ propose a procedure based on enumeration and constraint satisfaction for selection of polymer molecules.

p1330 Gani et al.¹⁸⁷ propose several steps similar to their previous works but allow for the use of optimization after feasible structures are previously identified through enumeration. Important equilibrium-based constraints are also used in this work regarding the formation of binary azeotropes, the determination of partial miscibility, and the location of solubility limits. Constantinou et al.¹⁸⁸ further extend this approach to systems exhibiting vapor–liquid and solid–liquid equilibria, with applications in polymer design and foam blowing and cleaning agents for textiles. Harper et al.¹⁸⁹ propose a multilevel molecular design approach, where a set of feasible structures is generated and screened based on first- and second-order groups. After the first two steps, the remaining structures are represented through atomic connection matrices which include atoms and atom connectivity. Structures that remain after the previous screening level are passed onto the final level where QM models are used. This approach is included in the CAMD step of a broader framework proposed by Harper and Gani,¹⁹⁰ which also includes predesign and postdesign steps. The predesign step supports the problem formulation through a knowledge base. In this step, the user only specifies the desired task (e.g., separation of a mixture) and the knowledge base supports the identification of different ways (e.g., processes) to perform the task. Each available option corresponds to a different CAMD problem formulation, which is proposed by the knowledge base. The postdesign step includes search in databases for more detailed information regarding the selected molecules. Conte et al.¹⁹¹ expand the generate and test approach into a generic multistep methodology for the design of formulated products. They propose an integrated experiment-modeling approach which supports the decomposition of the design problem into a hierarchical sequence of subproblems (levels). At the outer level, predictive models with wide application are employed to screen out options which are reduced toward the inner levels. The latter employ special correlations, rigorous models, and/or experiments. The developed workflows are used in the Virtual Product-Process Design (vPPD) laboratory, a software proposed and developed by Conte et al.¹⁹² with the aim to be able to propose product-process designs and then determine the experiments needed to verify them. Applications include water-based insect repellent and waterproof sunscreen. The framework is extended by Mattei et al.¹⁹³ to emulsion-based formulated products.

p1335 Yang and Song¹⁹⁴ propose an approach called “classified enumeration” to address the combinatorial complexity in generate and test approaches. The authors define skeleton groups which are used to generate alkane-based structures and function groups which substitute the alkane-based groups. By implementing constraints, the skeleton-based structures are pruned and gradually replaced by skeleton-function containing structures.

s0160 **Deterministic Optimization Algorithms**

s0165 **Overview of main algorithms**

p1340 Molecular design optimization problems often include continuous and integer variables in combinations with linear or nonlinear objective functions and/or constraints. Nonlinear programming (NLP) approaches address the solution of optimization problems with continuous variables, where the objective function is nonlinear, and/or the feasible region is determined by nonlinear constraints. The combination of a linear objective function and constraints with continuous and integer optimization variables results in mixed-integer linear programming (MILP) problem formulations. The presence of integer variables in nonlinear problems results in mixed-integer nonlinear programming (MINLP) problem formulations. Integer variables indicate discrete decisions, pertaining to the existence of groups in molecular structures, which are represented by binary variables in MILP and MINLP formulations.

p1345 The development and application of deterministic algorithms for the solution of the optimization problem depends on the analytical properties of the mathematical model (e.g., monotonicity and convexity). The combinatorial complexity increases with the addition of more binary variables, whereas the use of detailed models often results in nonlinearities. When the nonlinear constraints are convex, and the objective function is convex (for a minimization problem), the globally optimum solution may be identified even for problems of large size. When the objective function or any of the constraints are nonconvex, the optimization problem may have multiple feasible regions and multiple locally optimum solutions within each region, leading to significant challenges in identifying the global optimum.

p1350 According to Adjiman et al.,¹⁹⁵ deterministic algorithms generally approach the optimization problem through the formulation of subproblems which are easier to solve than the original problem. Subproblems are generated by strategies such as fixing of variables, relaxing of constraints, and linearizing of nonlinear functions, to name but a few. The algorithms iterate through the solution of the subproblems to identify upper and lower solution bounds on the optimal solution of the original problem. Algorithms have been developed to identify globally optimum solutions in both convex and nonconvex MINLP formulations. While deterministic optimization algorithms are often based on the calculation of derivatives, derivative-free optimization (DFO) algorithms also exist.¹⁹⁶ The main classes of deterministic optimization algorithms and approaches used in computer-aided molecular design are briefly summarized here. A detailed discussion and classification of derivative-based, deterministic

optimization algorithms is given in Adjiman et al.¹⁹⁵ and Floudas and Gounaris,¹⁹⁷ together with a review of more advanced algorithms. Boukouvala et al.¹⁹⁸ and Rios and Sahinidis¹⁹⁶ provide reviews of DFO algorithms.

s0170 Branch and bound

p1355 Branch and bound (BB)¹⁹⁹ involves a binary tree where all binary variables are relaxed between 0 and 1 at the initial node. The NLP problem is then solved, giving a lower bound in the objective function (for a minimization problem). The node is then separated into two branches representing the discrete values for one binary variable. NLP problems are solved in both nodes and the node with the best objective is the best option for continuing. The procedure continues, and as new layers are added the final layer represents the feasible solutions and provides the upper bound to the problem. Nodes can be eliminated if their lower bounds exceed the current upper bound or if the relaxation is infeasible. The optimum is the solution in the final unrelaxed layer. The aim is to explore as few nodes as possible. BB principles are used in advanced approaches, such as the SMIN- α BB algorithm^{195,200} and the Branch-and-Reduce approach²⁰¹ used in the BARON algorithm. Both are global optimization algorithms for nonconvex problems, used in molecular design.

s0175 Generalized Benders decomposition

p1360 Generalized Benders decomposition (GBD)²⁰² decomposes the problem into primal and master subproblems and the solution is identified by iterating between the two. The primal problem solves the NLP generating an upper bound to the optimal solution. The master problem determines the next integer combination and provides a lower bound on the solution. Gradually, the upper bound decreases and the lower bound increases, hence the algorithm terminates when the two bounds are sufficiently close. An illustration of how the method is applied to CAMD problems is available in Buxton et al.²⁰³

s0180 Outer approximation

p1365 Outer approximation (OA)²⁰⁴ alternates between master and primal problems to identify upper and lower bounds on the solution to the MINLP. The upper bound results from the solution of a primal problem which is formulated identically to the primal problem for GBD. The lower bound is determined by solving a master problem which is an outer linearization of the problem around the primal solution. Outer approximation with equality relaxation (OA/ER) is proposed by Kocis and Grossmann²⁰⁵ to allow for nonlinear equality constraints, which are relaxed into inequality constraints. Outer approximation with equality relaxation with augmented penalty (OA/ER/AP)²⁰⁴ uses a relaxation of the linearizations of the master problem in order to expand the feasible region. Through this expansion, the probability of cutting part of the feasible region due to one of the linearizations is reduced. The aim of the algorithm is to avoid convexity assumptions required to find the optimum solution, but this cannot be guaranteed because part of the feasible region may be eliminated.

s0185 Generalized disjunctive programming

p1370 Generalized disjunctive programming (GDP)²⁰⁶ is an approach that involves Boolean and continuous variables associated through disjunctions, algebraic equations, and logical propositions. In the general GDP formulation, the objective is a function of continuous variables, while there are sets of inequality constraints which represent general constraints that must hold regardless of the discrete choices. Additionally, there are nonlinear and linear conditional constraints that depend on the discrete decisions, as represented by the Boolean variables. Conditional constraints are included inside disjunctions and linked with the OR operator. The disjunctive terms are assigned Boolean variables, whereas each term is active when the corresponding Boolean variable is true. Propositional logic is used to express logic relations for the Boolean variables.²⁰⁷ GDP formulations are converted into MINLP through the Big-M (BM) approach²⁰⁸ or the Hull reformulation (HR).²⁰⁹

s0190 Interval analysis

p1375 Interval analysis (IA)-based algorithms²¹⁰ are classified by Adjiman et al.¹⁹⁵ as part of algorithms used to target globally optimum solutions in nonconvex MINLP formulations. The underlying principle is similar to BB regarding the successive partitioning of the solution space and bounding of the objective function within each domain, with branching performed on the discrete and continuous variables. However, bounds on the problem solution in a given domain are not obtained by optimization.¹⁹⁵ Instead, interval arithmetic is used to compute the range of the objective function in the considered domain. The idea is to continuously delete parts of the search space with the objective of maintaining a final box of any desired width which contains the global solution.²¹⁰

s0195 Deterministic derivative-free optimization

p1380 Deterministic DFO is a class of algorithms which avoid the calculation of derivatives.¹⁹⁶ In several problems, the derivative information is unavailable, unreliable, or impractical to obtain. In such cases, DFO algorithms generally approach the nonlinear optimization problem directly or indirectly, based on function evaluations. In the direct approach, mathematical analysis has established theoretical foundations which guarantee convergence to local first- or second-order stationary points using samples.¹⁹⁸ The indirect approach can generally exploit theoretical developments for deterministic MINLP problems. DFO algorithms are generally classified into local search methods and global search algorithms. Local search methods include direct local search methods and local model-based search algorithms. Global search methods include deterministic global search algorithms and global model-based search algorithms. Details are reported in Rios and Sahinidis.¹⁹⁶

s0200 **Deterministic optimization approaches in molecular design**

p1385 This section summarizes contributions where deterministic optimization approaches are used in molecular design. The review focuses on works that proposed deterministic optimization approaches and algorithms. Cases where such algorithms are implemented in various problem classes or applications are further reviewed in subsequent sections. Table 5 provides a summary of the main features discussed here.

p1390 The original use of optimization approaches traces back to the work of Macchietto et al.,¹⁷³ who use a continuous optimization formulation which results in a relaxed solution, representing the molecule through a noninteger number of groups. Integer forcing constraints have to be added, but the resulting MINLP cannot be solved rigorously, hence the solvent optimality cannot be guaranteed. This is addressed by Odele and Macchietto,¹⁶⁰ who use an MINLP solver based on the OA/AP approach.²⁰⁴ Meanwhile, Naser and Fournier¹⁷⁴ propose a continuous optimization representation of the CAMD problem, based on the concept of the “theoretical molecule” that results from the continuous representation of the molecular composition. They also propose different procedures to address (a) the molecular feasibility constraints in view of the continuous molecular representation and (b) the possibility of having discontinuities due to the appearance of multiple phases in the liquid–liquid equilibrium problem that they investigate. In case (a), they generate five classes of groups and assign different UNIFAC groups to them. They further deduce numeric inequality constraints which provide a region of feasibility whose nonzero width provides fuzzy feasibility and whose zero width provides strict feasibility. In case (b), they propose the “distance from a target” concept. They define an ideal theoretical separation target and minimize the distance from this target, as their optimization objective. In doing so, they make the equilibrium relationship dependent only on the molecular structure. The problem is solved through successive quadratic programming (SQP) based on Schittkowski.²¹¹

p1395 Klein et al.²¹² propose a mathematical CAMD formulation, including multiple linear constraints and one nonlinear constraint. The discontinuities observed in the derivatives due to having the number of solvents as a design variable could not be addressed through standard nonlinear optimization at the time. The authors use the successive regression and linear programming (SRLP)²¹³ approach to overcome this issue. The blend design problem is also addressed, while the method is then implemented as the fourth part of a four-stage approach²¹⁴ which first reduces the design space by identifying fewer feasible candidates and then addresses the identification of the optimum blend.

p1400 To address problems due to nonconvexities, Maranas²¹⁵ proposes to transform nonlinear structure–property relations into an equivalent MILP problem. This is approached through the expression of integer variables as a linear combination of binary variables and the replacement of continuous and binary variable products with linear inequality constraints. The approach is extended to

t0030 **Table 5** Main features of deterministic optimization approaches used in molecular design problems

Authors	Optimization methods	Approach
Macchietto et al. ¹⁷³	NLP	Decomposition
Naser and Fournier ¹⁷⁴	NLP, SQP	Direct
Odele and Macchietto ¹⁶⁰	MINLP, OA/AP	Direct
Klein et al. ²¹² and Gani and Fredenslund ²¹⁴	SRLP	Direct
Maranas ²¹⁵ and Raman and Maranas ¹⁰	Transformation of MINLP into an equivalent MILP problem	Direct
Camarda and Maranas ²¹⁶ and Maranas ¹⁶⁸	Convexification of nonlinear terms, MINLP, OA/AP	Direct
Duvedi and Achenie, ^{217,218} Churi and Achenie, ²¹⁹ Wang and Achenie ²²⁰	MINLP, OA/AP	Direct
Sinha et al. ²²¹ and Ostrovsky et al. ²²²	IA global optimization for nonconvex problems through BB with linear underestimators and splitting functions for branching reduction	Direct
Friedler et al. ²²³	MINLP, BB	Direct
Achenie and Sinha ²²⁴ and Sinha et al. ²²¹	IA global optimization, LIBRA algorithm	Direct
Karunanithi et al. ^{226–228}	General framework for MINLP formulations	Decomposition
Vaidyanathan and El-Halwagi ²²⁹	IA global optimization for nonconvex problems with lower bound tests and distrust-region method, GINO software as local optimizer	Direct
Hostrup et al. ²³¹	Hybrid approach, MINLP, inner NLP, outer MILP	Decomposition
Zhang et al. ²³²	General MILP/MINLP formulation	Direct
Buxton et al. ²⁰³ and Giovanoglou et al. ²³³	MINLP, MIDO, GBD	Decomposition
Sheldon et al. ²³⁴	MINLP, OA/ER	Direct
Sioukrou ¹⁵⁹	MINLP, SMIN- α BB	Direct
Burger et al. ²³⁷	MINLP, OA/ER/AP	Decomposition, direct solution in each of two stages
Jonuzaj et al. ^{238,239}	MINLP, GDP, BM	Direct
Jonuzaj and Adjiman ²⁴⁰	MINLP, GDP, HR	Direct
Gopinath et al. ²⁴¹ and Gopinath ¹⁷⁹	MINLP, modified OA	Decomposition
Sahinidis et al. ²⁴²	Branch-and-reduce global optimization, BARON	Direct
Samudra and Sahinidis ^{158,243}	MILP	Decomposition
Austin et al. ¹⁴⁹	DFO, MILP, and BB (BARON) for continuous nonconvex problem	Direct
Cheng and Wang ²⁴⁴	MIHDE and MISQP	Direct

- more complex molecular representations, considering nonlinearities introduced by topological indices.¹⁰ Camarda and Maranas²¹⁶ employ the topological index representation and propose a MINLP CAMD formulation which includes only linear and convex nonlinear relations. The problem is solved using the OA/AP approach. Maranas¹⁶⁸ addresses the CAM^UD problem by introducing the deterministic equivalent chance constraint representation to transform the original nonlinear stochastic formulation into a deterministic MINLP problem with linear binary and convex continuous parts.
- p1405 Duvedi and Achenie²¹⁷ propose a MINLP formulation which includes nonlinear constraints with respect to integer variables. This is addressed by defining new continuous variables and new constraints in which these variables are set equal to the integer variables, thereby restricting the nonlinearity to the continuous variables only. The problem is solved with OA/AP. The approach is extended to the design of refrigerant blends by Duvedi and Achenie²¹⁸ for a single evaporator cycle and dual evaporator cycle by Churi and Achenie²¹⁹ and to the design of extractive fermentation solvents by Wang and Achenie.²²⁰ Sinha et al.²²¹ and Ostrovsky et al.²²² propose modifications in a BB algorithm to identify globally optimum solutions in CAMD. They introduce splitting functions to reduce the number of branching nodes and further propose a strategy for the construction of linear underestimators relevant to solvent design problems. Friedler et al.²²³ propose a BB approach to CAMD in order to overcome limitations of the generate and test approach.
- p1410 Achenie and Sinha²²⁴ and Sinha et al.²²⁵ develop the LIBRA algorithm which is based on global optimization using IA. They propose underestimators and splitting functions to identify globally optimum solutions through a BB algorithm. Considering water as the cosolvent, the authors either find the globally optimum concentration for seven binary mixtures, using a basis set of seven candidates, or find the globally optimum water-solvent pair and concentration in a second case study. Karunanithi et al.²²⁶ propose a decomposition-based approach applied for single molecules and mixtures consisting of constraint satisfaction steps and an optimum design step which is formulated as an MINLP. Candidate solvents go through successive steps, hence a reduction of feasible solutions is achieved. The method enables the consideration of an optimum process design step as well (CAMPD). The problem is implemented to liquid-liquid extraction solvents and to pharmaceuticals, while the method is also implemented to crystallization solvent design.^{227,228}
- p1415 Vaidyanathan and El-Halwagi²²⁹ propose the use of global optimization based on IA for mixture design. The approach eliminates parts of the search space in order to obtain a domain that contains the globally optimum solution. The authors employ the algorithm of Vaidyanathan and El-Halwagi²¹⁰ which includes a lower bound test as well as a "distrust-region" method for deleting infeasible subspaces. The GINO software²³⁰ is used as a local optimizer.
- p1420 Hostrup et al.²³¹ address CAMPD and propose a stepwise procedure to systematically reduce the process and molecular search space based on the approach of Harper et al.¹⁸⁹ The process and molecular candidates in the final step are identified using an inner loop NLP and an outer loop MILP, unless too few options enable complete enumeration. For fixed binary variables, the inner NLP identifies the optimum solution of the continuous formulation. At this solution, the nonlinear equations are linearized, and the solution is communicated to the outer loop MILP to generate new values for the binary variables. The overall approach is a generate-and-test/optimization hybrid. Zhang et al.²³² develop a CAMD methodology with the aim to extend the feasible space of a standard CAMD problem by including both first- and second-order group information in the MILP/MINLP formulations.
- p1425 Buxton et al.²⁰³ address the CAM^bPD problem through a decomposition approach, algorithmically implemented in simultaneous calculations. An iterative MINLP solution scheme is used based on GBD, where in each major iteration of the algorithm a new solvent mixture is generated. The mixture is examined in a series of property and mass transfer feasibility tests. Mixtures that satisfy these tests go through process design. The approach is extended by Giovanoglou et al.²³³ to address mixed-integer dynamic optimization (MIDO) problems for simultaneous batch separation and solvent design. Sheldon et al.²³⁴ address a QM-CAMD problem. They formulate an MINLP with nonlinear equality constraints where the objective function value and its gradients are calculated through the QM software. The gradients of the nonlinear constraints and the Lagrange multipliers of the nonlinear equality constraints are calculated via SNOPT.²³⁵ The MINLP is approached through OA/ER, without the need for the AP version. The master and primal problems are solved in two different computers with automated transfer of information. Sioungkrou¹⁵⁹ addresses the QM-computer-aided molecular design for reactions (QM-CAM^RD) problem through an MINLP formulation using the SMIN- α BB algorithm, a deterministic global branch-and-bound optimization algorithm for nonconvex MINLP problems with general nonconvexities in the continuous variables and linear and bilinear participation of the binary variables. The problem addressed by Sioungkrou¹⁵⁹ is a bi-level optimization problem, where internal geometry optimization is required in addition to identifying the optimum solvent. In this case, the geometry is assumed to be fixed when moving from the gas to the liquid phase, hence the problem is transformed to single-level optimization. To address computational complexities from full QM calculations, Sioungkrou¹⁵⁹ derives and employs a Kriging-based surrogate model²³⁶ in place of the full QM model. The resulting MINLP is solved through BARON. Burger et al.²³⁷ address the CAMPD problem through the HiOpt approach where initial molecular and process solutions are derived using low-fidelity models at the process stage. These solutions are used as starting points for the subsequent CAMPD problem which employs process models of higher fidelity. First-stage solutions are extracted using MOO. The problem is solved using OA/ER/AP. Jonuzaj et al.²³⁸ propose a GDP framework to address the CAM^bD problem. In the proposed formulation, they account for specific constraint characteristics; those that do not depend on the logic conditions are formulated as general constraints, whereas those that depend on the logic conditions, such as on the assignment of compounds or on the number of components in a mixture, are formulated within the disjunctions as conditional constraints. The GDP is transformed into MINLP through the BM formulation. This work initially addresses blends, but is then extended to mixtures by Jonuzaj et al.²³⁹ Jonuzaj and Adjiman²⁴⁰ extend the blend design work by considering the HR approach for the transformation of the GDP to an MINLP and find that it leads to bounds that are at least as tight as or tighter than the BM approach, at the expense of introducing new variables and

constraints. Gopinath et al.²⁴¹ address the CAMPD problem by proposing a modified OA algorithm. Tests that remove infeasible regions from both the process and molecular domains are embedded within the OA framework. Four tests are developed to remove subdomains where constraints on phase behavior that are implicit in process models or explicit process (design) constraints are violated. The work is extended to cases including both continuous and discrete variables in the process synthesis domain.¹⁷⁹

p1430 Sahinidis et al.²⁴² propose a global optimization approach to CAMD based on an algorithm that provides all feasible solutions to this formulation through the implicit enumeration of a single branch-and-reduce tree. Samudra and Sahinidis¹⁵⁸ propose a multistage CAMD framework. The first stage employs approximate GC models to obtain molecular structures that fit desired design targets. The second stage identifies isomers for these structures, with structure–property relations providing more accurate property predictions. The third stage uses problem-specific models to further refine predictions. The proposed decomposition allows the formulation and solution of MILP subproblems, hence avoiding MINLP formulations. Additional features include the incorporation of several graph theoretic and linear integer programming models, which are developed to address the systematic generation of isomers and avoid redundancy in molecule graphs. The method is applied to the design of heat-transfer fluids for refrigeration.²⁴³ Austin et al.¹⁴⁹ propose a multistage approach for CAM^bD using DFO. Initially, the DFO produces trial points which are used to generate upper and lower property bounds. Molecular compositions (i.e., first-order groups and number of appearances of each group in the composition) are proposed through CAMD within these property bounds based on minimization of their distance from the trial point. These compositions are then transformed into structures using higher-order models and the one with the minimum distance from the trial point is selected. In the next step, the selected structure is used as the component that is designed within the mixture and the optimum concentration is specified for all mixture components. This step is a continuous, nonconvex NLP problem. The obtained objective function value is communicated to the DFO, where if the termination criteria are not satisfied the DFO proposes a new trial point. The advantage of this approach is that it solves CAMD problems within small search spaces using MILP technology, while the continuous nonconvex mixture concentration problem is decoupled.

p1435 Cheng and Wang²⁴⁴ address the CAMPD problem by proposing an approach where mixed-integer hybrid differential evolution (MIHDE)²⁴⁵ is firstly used to find a feasible solution, which is then used as the initial starting point for a trust region sequential quadratic programming algorithm (MISQP).²⁴⁶

s0205 Stochastic Optimization Algorithms

s0210 Overview of main algorithmic concepts

p1440 Stochastic optimization algorithms are used to solve MI(N)LP problems without considering derivative information. They emulate natural processes in order to explore the solution space of a given problem. The optimum solution is identified after a series of biased probabilistic transformations of an initial problem instance. The introduced bias is often cast as descent bias (i.e., based on the objective function), memory bias (i.e., based on previously made decisions), or experience bias (i.e., based on prior performance).²⁴⁷

p1445 Stochastic optimization algorithms generally require little or no assumptions on the optimization problem characteristics and structure (e.g., monotonicity and convexity), while they can even be applied to ill-structured problems for which no efficient local search procedures exist.²⁴⁸ These are useful features for molecular design problems because the search space may be discontinuous (e.g., due to the appearance and disappearance of phases as molecular composition varies) and nonconvex, which represent significant challenges for deterministic algorithms. Furthermore, they avoid locally optimum solutions through mechanisms which are based on the probabilistic acceptance of solutions that fail to improve the objective function. This feature is called diversification and enables efficient exploration of the search space. They also include mechanisms to intensify the search around promising solutions as the optimization search proceeds. Balancing diversification and intensification is one of the important challenges in such algorithms,²⁴⁷ together with the lack of guarantees regarding the optimality of the obtained solutions. The latter is evaluated using statistical procedures.²⁴⁹

s0215 Genetic algorithms (GA)

p1450 GA²⁵⁰ are a metaheuristic used in several molecular design implementations, with contributions summarized in Table 6. It is based on the Darwinian model of natural selection and evolution.²⁵¹ Chromosomes are used to represent molecular structures, comprised of genes which represent functional groups. The mapping of structural molecular information in a format that can be used in GA operations is called molecular encoding. Efficient encoding should facilitate new individuals via genetic operations while also favoring the decoding procedure where individual fitness is evaluated.²⁵² An initial population of chromosomes representing the first generation is evolved randomly to successive generations through genetic operators which serve different purposes. Selection is an operator where pairs are randomly selected from the population and only the fittest of each pair are allowed to mate. Each pair of parents creates one or more children that have some of the characteristics of each one of the parents. Crossover is another operator where the genetic material of the fittest individuals is randomly combined. Mutation is also used to introduce new solutions into the search, while elitism is used to enhance intensification around promising solutions by passing them to the subsequent generation. There are several different genetic operators, discussed in Patkar and Venkatasubramanian.²⁵¹

p1455 Venkatasubramanian et al.²⁵³ present a GA approach for CAMD, addressing the case of polymer design where there is a need of repeating monomer units. They employ a symbolic molecular encoding scheme where elemental, substructural, and monomer units are represented by genes. Van Dyk and Nieuwoudt²⁵⁴ employ GA to design solvents for extractive distillation. They use UNIFAC groups and note that genes are linearly combined to construct chromosomes, but they need not be linear themselves. This

10035 **Table 6** Main features of GA approaches used in molecular design

Author	Encoding	Operators
Venkatasubramanian et al. ^{253,264} and Pakar and Venkatasubramanian ²⁵¹	String of symbols, double-valence groups are main-chain blocks, single-valence groups are side-chain blocks	Single-point crossover, main-chain and side-chain mutation, insertion and deletion, blending, hop-mutation
Van Dyk and Nieuwoudt ²⁵⁴	UNIFAC groups, single-, double-, and zero-valence blocks	Point mutation, crossover, insertion and deletion
Lehmann and Maranas ¹⁷²	Bit strings as one-dimensional fields of binary elements	Selection, crossover, mutation
Xu and Diwekar ^{255,256}	Element vectors with a predefined maximum number of allowed groups	Tournament selection, uniform crossover, jump mutation
Heintz et al. ¹⁵³	Graph-based encoding through adjacency matrix	Mutation, crossover, insertion, deletion, substitution
Herring and Eden ²⁵⁷	Graph-based encoding	Crossover, reduction, insertion, fragment mutation
Zhou et al. ²⁵²	Tree structures stored as a dynamic list with UNIFAC groups as the tree nodes (building blocks)	Crossover, mutation, insertion, deletion
Scheffczyk et al. ²⁵⁹	Molecules as a set of fragments using a SMILES-based representation	Crossover, mutation
Liu et al. ²⁵⁸	Similar to Xu and Diwekar ²⁵⁵	Selection, crossover, mutation, repair

is a useful feature to represent branched or cyclic compounds. They use three types of genes, namely start/end genes with only one free bond, middle genes with two free bonds, and genes without free bonds. Lehmann and Maranas¹⁷² use GA in an approach that combines QM models with GC. Xu and Diwekar^{255,256} employ GA for solvent design under uncertainty. They develop strategies to ensure that when crossover is applied to molecules with different lengths, longer structures do not remain without changes, hence the search remains wide and efficient. Heintz et al.¹⁵³ used GA for mixture design, based on a graph representation of the molecules, in the form of an adjacency matrix. Herring and Eden²⁵⁷ employ spatial fragment-based descriptors to generate candidate solutions within a population, evolved, and evaluated using GA operations. A graph-based molecular encoding is used, hence customized crossover and mutation operators are developed to be compatible with this representation. Zhou et al.²⁵² use GA for the design of solvents for liquid-phase reactions. They propose a novel-encoding scheme where all the important structural and group information for the molecule is stored as a dynamic list. This facilitates the genetic operations and the subsequent decoding. Liu et al.²⁵⁸ propose a hybrid formulation between GA and simulated annealing (SA). The aim of this work is to exploit the intense diversification accomplished through GA with the intensification capabilities of SA. The encoding of the molecules is similar to Xu and Diwekar.²⁵⁵ Scheffczyk et al.²⁵⁹ use GA to integrate QM information into CAMD. A fragment-based molecular description is used, where molecular structures are decomposed into basic fragments and each fragment is described by a unique SMILES-based representation.²⁶⁰ The GA is based on the LEA²⁶¹ and the LEA3D²⁶² algorithms used for drug design. It is worth noting that computer-aided drug design (CADD) uses similar approaches and algorithms to CAMD (e.g., reverse design of molecules, evolutionary algorithms, and MOO—reviewed by Devi et al.²⁶³) but evolved independently of CAMD. In CADD problems, property prediction models and requirements are often different to those used in CAMD, hence they are not reviewed here.

s0220 **Simulated annealing**

p1460 SA²⁶⁵ is an algorithm that emulates the annealing of metals. Algorithmically, it is a statistical cooling optimization technique that generates a biased random search and employs Monte Carlo simulations under a variable probability schedule.²⁴⁹ The algorithm starts from an initial problem state (molecular structure) at a high temperature. The initial structure is randomly transformed into a number of successive structures within each temperature interval, always using the previously accepted structure as a starting point. Each structure enables the calculation of a molecular or other index which is used as an objective function. The latter is evaluated using the Metropolis criterion,²⁶⁶ whereby structures that improve the objective function are directly accepted, while structures that deteriorate the objective function are accepted with a certain probability. This operation supports diversification and facilitates the avoidance of local optima. After a predetermined number of iterations (Markov chain length), the temperature decreases. This procedure is repeated until certain termination criteria are satisfied. The acceptance rate of each newly generated structure is gradually reduced at lower temperatures, giving rise to increasingly intensified search around solutions that improve the objective function.

p1465 SA-based CAMD is proposed by Marcoulaki and Kokossis^{267,268} and applied to solvent and refrigerant design by Marcoulaki and Kokossis²⁶⁹ and Marcoulaki et al.²⁷⁰ Molecular structures are defined as the product of a vector indicating the type of UNIFAC functional groups and a diagonal matrix indicating the number of appearances of the group in the molecule. The authors provide a systematic approach for the evolution of the molecular structures during optimization in view of different connectivity constraints. Operators for new structures include expansion, contraction, and alteration. Papadopoulos and Linke¹⁶¹ extend the algorithm of Marcoulaki and Kokossis²⁷⁰ into MOO SA with applications in various fields.¹⁵⁰ Ourique and Telles²⁷¹ also use SA for CAMD using a molecular graph representation in the form of a structure–composition matrix, similar to the one employed by Heintz et al.¹⁵³ (see “Genetic algorithms” section). Kim and Diwekar²⁷² apply SA in CAM¹D in the form of the stochastic annealing algorithm.¹⁶⁷ The work is also extended in the case of the CAM¹PD problem.²⁷³ SA is also used by Song and Song²⁷⁴ and Song et al.²⁷⁵ for the design of separation solvents.

s0225 **Tabu search**

p1470 Tabu search (TS)²⁷⁵ does not mimic a physical system like the previous algorithms. In principle, the algorithm applies a best-improvement, local search and uses a short-term memory to escape from local minima.²⁴⁷ The short-term memory is implemented as a tabu list that registers recent solutions (molecular structures) which should not be revisited for desired number of iterations. At each iteration, the best molecular structure is chosen as the new current structure. Additionally, this structure is added to the tabu list, and one of the structures that were already in the tabu list is removed. Aspiration criteria allow the tabu list to be overridden if either the new solution is the best solution so far or all neighborhood solutions trialed are on the tabu list.²⁷⁶

p1475 Chavali et al.²⁷⁷ use TS for the design of transition metal catalysts. The properties of the catalysts are predicted using connectivity index descriptors which are used within specifically derived structure–property correlations. Lin et al.²⁷⁸ propose a generic framework for CAMD using TS where molecules are represented through connectivity indices and neighbor generation operators are defined, including operators such as replace, insert, delete, swap, and move. The proposed TS approach is applied to the design of cross-linked polymer networks,²⁷⁹ ionic liquids,²⁸⁰ and visible light photosensitizers for dental adhesive.²⁸¹

s0230 **Ant colony optimization**

p1480 Ant colony optimization (ACO) is an algorithm inspired by the foraging behavior of real ants.²⁴⁷ The underlying idea is that ants try to find the shortest paths between food sources and their nest. The paths are tracked by ants through the pheromone they lay as they move. Increasing concentrations of pheromone indicate paths worth following, as they are likely to lead to food sources. This collective behavior leads to the probabilistic emergence of the minimum length of paths. Artificial ants used in optimization choose their destination with a probability that is a function of the associated pheromone value. They are further forced to move to regions they have not visited before (e.g., through the use of tabu lists), while after every iteration is completed they lay an amount of pheromone that is a function of the improvement of the fitness of the solution. Diversification can be enhanced through trail diffusion operations in order to generate new regions out of the existing ones.²⁴⁹

p1485 ACO has been first proposed for CAMD by Gebreslassie and Diwekar²⁸² in the form of the efficient ACO (EACO) algorithm.²⁸³ In EACO, the random samples required by the algorithmic operations are drawn using Hammersley²⁸⁴ sampling, a method that enables the derivation of a uniformly distributed sampling pattern. The work is generalized by Gebreslassie and Diwekar,²⁸⁵ who develop a general multiagent optimization framework which can consider various different algorithms. Further applications of EACO have been reported by Mukherjee et al.²⁸⁶ and Doshi et al.²⁸⁷ in the design of adsorbents for metal ion removal from water and by Benavides et al.^{288,289} in the design of adsorbents for adsorption of natural occurring radioactive material (NORM) of natural gas fracking waste.

s0235 **Property Clustering Approaches**

p1490 Property clustering has been first proposed by Shelley and El-Halwagi²⁹⁰ as a tool for process design. The underlying idea is that tracking of components is not necessary in many processes, especially in those where there are numerous different species that vary in composition and concentration among different streams. Instead, it is both necessary and easier to track component properties for measuring process performance. Developing design procedures that account for properties is not as straightforward as using components; properties are not conserved whereas components are. Shelley and El-Halwagi²⁹⁰ propose the concept of clusters which are formal conserved quantities that are related to nonconserved properties. Essentially, clusters support the conserved “mixing” of properties calculated from the mixing of process streams. Intrastream clusters are weighted sums of the component properties contained in a stream. The weights are the fractional contributions of a stream into the total flowrate of the mixture. Conservation laws apply both at intrastream and interstream property mixing. The interstream conservation allows for lever-arm rules to apply, so the clusters may be summed based on their contribution when two streams are mixed. Up to three clusters may be plotted in a ternary diagram, enabling visual tracking of properties.²⁹¹ This idea applies in sink-source mapping at the process level.

p1495 Eden et al.²⁹² propose this approach for use in CAMPD, where the property clustering is used for the process targeting and design part of CAMPD (details are reported in “Property Clustering and/or Reverse Formulations” section). Eljack and Eden²⁹³ built this concept in order to exploit it in GC-based CAMD. Clusters represent the “mixing” of properties which are calculated by combining functional groups. The development of clusters is described as follows:

$$\psi_j^M(P_j) = \sum_{g=1}^{N_g} n_g P_{jg} \quad (37)$$

$$\Omega_{ji}^M = \frac{\psi_j^M(P_{ji})}{\psi_j^{ref}(P_j)} \quad (38)$$

$$AUP^M = \sum_{j=1}^{NP} \Omega_j^M \quad (39)$$

$$C_j = \frac{\Omega_j^M}{AUP^M} \quad (40)$$

where $\psi_j^M(P_j)$ is the molecular property operator of the j th property, n_g is the number of type g groups in the molecule, P_{jg} is the contribution of group g in property j , Ω_{ji}^M is a normalized value of the operators, using a reference (*ref*) value for each property for

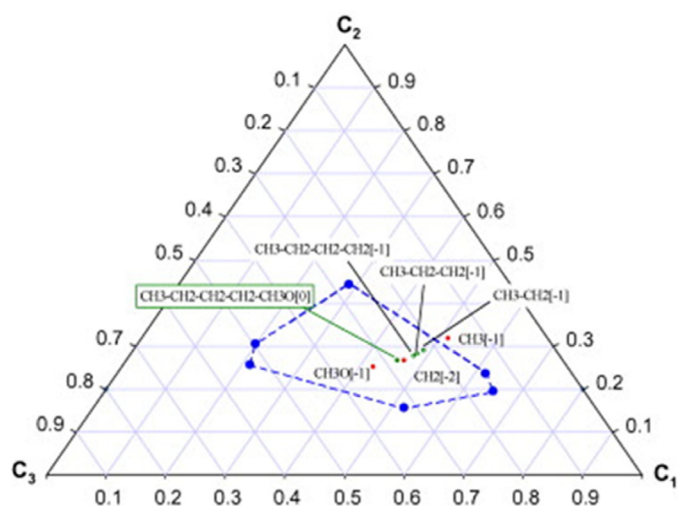


Fig. 4 Example of visual CAMD using the property clustering approach.²⁹³ Reproduced from Eljack, F.T.; Eden, M.R., *Comput. Chem. Eng.* **2008**, *32*(12), 3002–3010 by permission of Elsevier.

scaling purposes, AUP^M is an augmented property index for each molecule, and C_j is the cluster for property j . Eq. (37) is an expression of the right-hand side of Eq. (1) (assuming only first-order groups). This means that the left-hand side may still be nonlinear, although the clusters are based on linear additive rules. This is a key feature of the property clustering approach. Intra- and intermolecular conservation relationships also apply, as well as rules (e.g., for adding groups and chemical feasibility) detailed in Eljack²⁹⁴ and Eljack and Eden.²⁹³ Fig. 4 shows an example of CAMD using a visual clustering representation. The blue-dashed lines indicate feasibility limits based on desired constraints. Red dots indicate functional groups, whereas green dots indicate larger molecular fragments (and eventually a complete molecule) that result from linear addition of functional groups. The location of the fragments in the diagram is deduced from lever-arm calculations. Numbers in brackets indicate valences.

Research work pertaining to the use of the method in CAMPD and CAM^bD is reported in “Property Clustering and/or Reverse Formulations” and “Mixture and Blend Design” sections, respectively. Chemmangattuvalappil et al.²⁹⁵ propose the inclusion of second-order groups in molecular property clustering CAMD. They develop rules which determine the consideration of first-order groups as building blocks to second-order ones. They also propose an algebraic approach for solving molecular design problems in the cluster domain and introduce a systematic way to generate all possible molecular structures with a given set of property constraints. Chemmangattuvalappil et al.²⁹⁶ introduce molecular signature descriptors in the molecular property clustering framework. They develop an algorithm to account for different quantitative structure–activity relationship (QSAR)/QSPR expressions based on multiple topological indices for molecular design. In this new algorithm, GC methods are coupled with expressions of topological indices on a common property platform. Chemmangattuvalappil and Eden²⁹⁷ further extend their work to account for the simultaneous use of topological indices with different levels of structural information. A new CAMD algorithm is further proposed to simultaneously consider GC models with higher-order contributions along with different varieties of topological indices. The algorithm is able to track changes in the properties that occur as a result of reactions. Hada et al.²⁹⁸ use data obtained from infrared spectroscopy to develop latent variable property models which are used for the design of biofuel additives. The models are developed using techniques based on principal components and are transformed into property clustering operators, which are used in CAMD. This work is extended by Hada et al.²⁹⁹ by introducing QM-based calculations to supplement experimental IR data. Hada et al.³⁰⁰ further use chemometric techniques and the property clustering approach for blend and mixture design purposes. Latent variable property models are developed that simultaneously incorporate the properties of the raw materials, their blend ratios, and the process conditions. The identified solution is used as the target in the property clustering step to solve the mixture formulation problem in a reduced dimensional property space. Woo et al.³⁰¹ develop a method for property model derivation based on factorial design to estimate the relationships between properties and to determine the underlying interactions between more than two factors with the target property. The method is implemented using the approach of Chemmangattuvalappil et al.²⁹⁶

s0240 Merits and Shortcomings of Solution Approaches

p1505 Table 7 provides a summary of the merits and shortcomings of different solution approaches used in molecular design.

p1510 *Generate and test* approaches provide the opportunity to incorporate expert knowledge in order to rationally reduce the space of molecular candidates. This reduction allows the solution of easier optimization problems in the remaining space, while it also supports the identification of practically realizable solutions. When fewer candidates are available, they can be evaluated easier

t0040 **Table 7** Summary of merits and shortcomings of CAMD solution approaches

Approach	Merits	Shortcomings
Generate and test	Easy incorporation of expert knowledge and identification of practical solutions	Challenging enumeration of large molecular design spaces
Deterministic optimization	Fewer function evaluations than stochastic methods to reach optimum, analytical mathematical determination of local or global optimum, confidence results in problems where global optimum can be guaranteed	Derivative transformations, difficult simulation initialization in nonconvex models, mechanisms to avoid local optima is an active research field, locally optimum solutions may differ significantly, require knowledge of analytical problem properties (e.g., convexity and monotonicity)
Stochastic optimization	Include mechanisms to target globally optimum solution, no need for derivative calculations, able to handle discontinuities and nonconvex problems, no need for knowledge of problem structure, rich design insights from close but different optimum solutions	Larger number of function evaluations to identify optimum solution, statistical assessment of solution optimality
Property clustering	Insightful visualization of problem and solutions, easy avoidance of infeasible solutions, direct links with visualization approach applied in process design, algebraic representations available for more than three design performance criteria	Visualization advantage available for up to three performance criteria, efficient enumeration of candidate structures required when multiple molecular building blocks are considered

using higher-fidelity molecular or process models. On the other hand, the complete enumeration of large molecular spaces is computationally challenging.

p1515 *Deterministic* optimization methods are based on analytical properties (e.g., convexity and monotonicity) of the problem to generate a deterministic sequence of points converging to a local or global optimum.²⁴⁸ They provide insights regarding the local or global optimality of a solution through analytical mathematical conditions. Local optimization approaches require a lower number of objective function evaluations to reach an optimum solution than stochastic methods. Global optimization approaches for nonconvex problems require computations of similar intensity to stochastic methods. For example, although a BB-based approach is effective in locating the global solution in MINLP problems, it can be computationally expensive, because all binary variables are used as branching variables, and their number can reach several hundred.¹⁵⁹ General limitations involve the computationally intensive use of derivative transformations and difficulties in the initialization of simulations when complex models are considered. A fundamental issue of deterministic methods is to transcend local optimality,²⁴⁸ hence the development of mechanisms to prevent the convergence in local optima in highly nonconvex problems is also a very active research field.¹⁹⁷ SMIN-*a*BB and BARON are notable examples of deterministic algorithms that target the global optimum of nonconvex problems.^{195,302} Recent developments in DFO approaches reduce or remove the need for challenging derivative transformation and provide an avenue for convergence of research efforts between deterministic and stochastic algorithms.¹⁹⁸

p1520 *Stochastic algorithms* can be classified into trajectory- and population-based methods. SA follows a single-point search approach whereby an initial solution is evolved through successive alterations during optimization. TS allows the generation and assessment of multiple candidate solutions in the same iteration. Both SA and TS are classified as trajectory methods, meaning that the algorithm starts from an initial state and describes a trajectory in the state space.²⁴⁷ Successor solutions are generally implemented as small random increments of the preceding solutions, making the generated trajectories easy to track, review, and manipulate through the algorithmic control parameters. The latter are few; the Markov chain length and the cooling rate are the main control parameters in SA, whereas the tabu tenure (list length) and type of aspiration criteria are the corresponding parameters in TS.

p1525 GA and ACO are population-based methods where successor solutions are often weighted averages (e.g., through crossover and trail diffusion) or random instances (e.g., through mutation) of preceding solutions. Solution tracking is therefore difficult, while the control parameters are more than SA and TS. GA control parameters include the population number, the operator type (e.g., crossover, jump mutation, and creep mutation) and rate, and the population and solution management schemes (e.g., elitism and niching).²⁴⁹ ACO also includes the population number, the operator type (e.g., pheromone evaporation and trail diffusion) and rates, and the tabu tenure. In process synthesis experiments including both discrete and continuous variables, Papadopoulos and Linke²⁴⁹ found that SA and ACO have an equivalent performance in terms of solution quality, with ACO being faster. GA is always able to identify promising solution paths early-on in the search, but intensification (local search) around them is challenging and often more time-consuming than ACO and SA. A comparison between TS and SA in complex process synthesis problems¹⁷⁸ shows that both algorithms reach solutions of similar quality, but TS is faster.

p1530 *Property clustering* methods provide an insightful representation of the CAMD problem and of the obtained solutions. Their visualization capabilities support the exploration of multiple different options, other than the ones indicated by the design criteria. This allows the introduction of expert knowledge and the refinement of the design/selection criteria or property prediction models even during the search. The clear visualization of the feasibility space is also important. The combination of a clustering approach for the identification of desired process options for each investigated molecule is also straightforward as the two approaches are linked directly through component properties. On the other hand, visualization diagrams are limited to three performance criteria. The consideration of more criteria is often necessary to identify optimum CAMD solutions. Although algebraic representation

approaches of the method address this issue, the visualization advantages are removed. Furthermore, the visualization of the feasibility space and the lever-arm approach to molecular design reduce the investigation of redundant candidates, but the availability of multiple functional groups as molecular building blocks still requires the use of an efficient enumeration approach.

s0245 **Methodologies for Solution of CAMPD Problems**

p1535 The CAMPD problem has been approached using several different methodologies which employ all the solution approaches discussed in the previous section. CAMPD research contributions are summarized in [Tables 8 and 9](#). Examples of applications are given in certain cases to highlight key features of the design problem. The main aim of this section is to discuss major methodological contributions, hence further details regarding the applications or other developments are discussed in “[Methodologies for Various Problem Classes](#),” “[Methodologies for Reactive Systems or QM-Based Approaches](#),” and “[Applications](#)” sections. Methodologies for CAMPD of reactive systems or implementations that use QM models are discussed separately in “[Methodologies for Reactive Systems or QM-Based Approaches](#)” section.

s0250 **Generate and Test Approaches**

p1540 Jaksland and Gani³⁰³ propose a generic approach for CAMPD problems pertaining to solvent and separation process selection and design. The aim of the work is mainly to reduce the solvent and process options into a smaller set of feasible candidates based on thermodynamic insights. The separation technology itself is considered a decision parameter. Level 1 of the proposed approach employs pure component properties to identify feasible separation technologies. Level 2 employs pure component and mixture properties for separation task selection/sequencing, separation technology validation, and process/product design. After applying both levels of the methodology, a feasible flow sheet is generated together with alternatives. The flow sheet includes selected and sequenced separation tasks with corresponding separation techniques, estimates of the conditions of operation, and a selected solvent. The flow sheet is validated through process simulation.

p1545 Hostrup et al.²³¹ propose a multilevel procedure which includes seven main levels. An initial pool of separation technologies is gradually refined based on pure component and binary mixture tests up to level 4. In levels 5 and 6, solvents are designed using the CAMD algorithm of Harper et al.¹⁸⁹ and further mixture tests reduce potential solvent and process flow sheet options. The remaining process and solvent options are solved in level 7 as an optimization problem using an inner loop NLP and an outer loop MILP. These ideas are all expanded and used in the vPPD laboratory.^{191,192}

s0255 **Deterministic Optimization Approaches**

p1550 Hamad and El-Halwagi³⁰⁴ propose the simultaneous design of solvents and mass separation networks. The optimization of the mass exchange networks is based on process insights gained by application of mass integration principles. The resulting nonlinear thermodynamic feasibility constraints and overall material balance of the separation network as well as the material property estimation models are linearized using various different techniques. The study demonstrates that the optimal solvents are not the ones with the optimal properties or cost but instead compromises in both criteria. Pistikopoulos and Stefanis³⁰⁵ propose the selection of an optimal solvent candidate based on global plant-wide process and environmental constraints and apply their methodology to the design of absorption processes in the course of solvent design. Buxton et al.²⁰³ generalize this approach in a stepwise formulation where solvent mixtures are designed through CAM^bPD embedding tests to reduce the potential options using equilibrium-based assessment. Few feasible options are then considered with respect to optimum process operation for a fixed process structure. This methodology is used by Giovanoglou et al.²³³ for design of batch separation processes. Burger et al.²³⁷ address CAMPD through the HiOpt method discussed previously. The solvent design options are generated and evaluated through GCs using the SAFT- γ -Mie EoS for prediction of vapor-liquid equilibria. The process design space includes operating parameters for a fixed process structure. The embedding of tests into CAMPD is proposed by Gopinath et al.²⁴¹ and Gopinath,¹⁷⁹ who develop inlet stream phase stability tests after isothermal expansion, solvent handling, separation, and purification feasibility tests to eliminate infeasible areas of the CAMPD space. The approach is extended to molecular design and process synthesis problems where process structural decisions are also considered. Duvedi and Achenie²¹⁷, Churi and Achenie²¹⁹, Wang and Achenie,²²⁰ Achenie and Sinha,²²⁴ Karunanithi et al.,^{226,227} as well as Cheng and Wang^{244,306,307} have used advanced optimization algorithms considering process constraints in solvent design. Details are reported in “[Deterministic optimization approaches in molecular design](#)” section.

p1555 Martin and Martinez³⁰⁸ address the problem of identifying optimum blends and manufacturing conditions for development of detergents. The problem includes product performance constraints through a model regressed from empirical observations, environmental constraints, as well as manufacturing process constraints. An MINLP formulation is used and solved using the BARON algorithm. Fung et al.³⁰⁹ propose a generic CAMPD approach for product design which goes beyond the production process boundaries and accounts for product-use conditions, product quality, price and market share, product cost, manufacturing cost, economic analysis, and product commercialization. Appropriate models are considered for all these aspects and software is proposed which could perform such calculations. One of the case studies for die attach adhesive design includes a complete CAMPD MINLP formulation solved using BARON.

t0045 **Table 8** Overview of CAMPD approaches using generate and test methods or deterministic optimization algorithms

References	Key features
<i>Generate and test approaches</i>	
Jakslund and Gani, ³⁰³ Hostrup et al. ²³²	Multilevel problem decomposition to gradually reduce options, feasible separation process tasks identified through knowledge-based support, reduced superstructure of potential processing options, generate and test approach for molecular design, final solvent–process flow sheet set optimized
Conte et al. ^{191,192}	vPPD laboratory, multistep methodology to screen and verify options in formulated product design
<i>Deterministic optimization approaches</i>	
Hamad and El-Halwagi ³⁰⁴	Synthesis of solvents and mass exchange networks, linearization of the nonlinear problem, determination of optimum solvent structures and process characteristics, direct problem solution
Pistikopoulos and Stefanis, ³⁰⁵ Buxton et al., ²⁰³ Giovanoglou et al. ²³³	Decomposition approach to gradually identify feasible solvent mixtures which are then tested in process design, environmental constraints considered, application of approach in batch separation processes using dynamic models
Burger et al. ²³⁷	Solution of complete CAMPD problem using initial guesses from a previous stage solving CAMPD using low-fidelity models, each stage is solved directly, use of SAFT- γ -Mie EoS, physical CO ₂ capture solvents and processes
Gopinath et al. ²⁴¹ and Gopinath ¹⁷⁹	Infeasible areas are removed in the course of CAMPD using equilibrium-based tests, process and flow sheet structural decisions are also considered, use of SAFT- γ -Mie EoS, physical CO ₂ capture solvents and processes
Duvedi and Achenie, ²¹⁷ Churi and Achenie, ²¹⁹ Wang and Achenie, ²²⁰ Achenie and Sinha, ²²⁴ Karunanithi et al. ^{226,227}	Reformulation of nonlinear constraints and new continuous variables to restrict nonlinearities to continuous variables, use of OA/AP or IA global optimization algorithms, decomposition-based approach consisting of constraint satisfaction steps, few feasible molecules introduced in process design step, application to refrigeration, separation, extractive fermentation, blanket wash and crystallization systems
Cheng and Wang ^{244,306,307}	MIHDE algorithm, fuzzy programming, combination of MISQP, direct problem solution, applied to extractive fermentation systems
Martin and Martinez ³⁰⁸	Blend problem, direct formulation, product performance evaluated through empirical model, environmental and manufacturing process constraints, direct MINLP using BARON
Fung et al. ³⁰⁹	Grand model for chemical product design, inclusion of process, material, quality, pricing, and economic analysis models to bring products to commercialization, case studies as direct MINLP
Roughton et al. ³¹⁰	Ionic liquid and distillation process design, direct MILP
Palma-Flores et al., ³¹¹ Molina-Thierry and Flores-Tlacuahuac, ³¹² Santos-Rodriguez et al., ³¹³ Thierry et al. ³¹⁹	Working fluid and ORC process design, decomposed or direct (simultaneous) CAMPD, NLP, or MINLP formulations, single fluid or blend design, design under uncertainty
White et al. ³¹⁴	Single ORC working fluids and processes, SAFT- γ -Mie, direct solution of MINLP with OA/ER/AP
Cignitti et al. ^{315,317}	General framework for decomposition-based product-process design, applying optimization to working fluid and ORC design, supports gradual addition of modeling and decision-making details, other applications include refrigeration and liquid–liquid extraction systems
Stijepovic et al. ³¹⁸	Generic framework for the simultaneous selection of ORC working fluids and design of complex ORC flow sheets to draw heat from different heat sources, potential to select different working fluids in heat-sharing ORC cascades, MINLP formulation solved directly with BARON

p1560 Roughton et al.³¹⁰ propose the design of ionic liquids as solvents to break azeotropes in distillation. Ionic liquids are first designed using pure component and mixture properties as design criteria. The best options are then tested in the extractive distillation column design using a driving force method with a specifically proposed feed stage scaling to minimize energy inputs. Along with the distillation column, an ionic liquid recovery stage is designed, and simulations are used to determine the overall heat duty for the entire process for the best ionic liquid candidates.

p1565 Palma-Flores et al.³¹¹ propose a two-stage approach for the design of working fluids for Organic Rankine cycles (ORC). CAMD is addressed in the first stage as MINLP formulation using SOO, whereas selected highly performing molecules are subsequently introduced into an ORC process model to evaluate their performance in terms of thermal efficiency. Molina-Thierry and Flores-Tlacuahuac³¹² propose the simultaneous blend and ORC process design in an NLP formulation testing several different objective functions in independent SOO problems. A detailed overview regarding the two last papers can be found in Linke et al.¹⁷⁷ Santos-Rodriguez et al.³¹³ extend the previous works into the design of working fluids and ORC under uncertainty using deterministic stochastic programming. White et al.³¹⁴ implement a direct CAMPD approach in the design of working fluids and ORC processes

10050 **Table 9** Overview of CAMPD approaches using stochastic optimization algorithms, continuous molecular representation or property clustering approaches

References	Key features
<i>Approaches using stochastic optimization</i>	
Marcoulaki and Kokossis, ^{268,269} Linke and Kokossis ³²⁰	Simultaneous solvent and separation/reactive separation systems design, superstructure-based process synthesis, stochastic optimization, direct solution, SA
Papadopoulos and Linke, ^{161,163,321,322,326} Papadopoulos et al. ^{162,323,324}	MOO, superstructure-based process synthesis, data mining, grid computing separation and reactive separation process synthesis, single and mixed working fluids and ORC process design, integrated solvent, process and control design, decomposed approaches, SA
Qadir et al. ³²⁹	Solvent selection or blend design and process design, PPC-SAFT, prespecified solvent options, GA with internal loop for continuous parameters, precombustion carbon capture
Kim and Diwekar, ²⁷³ Kim et al., ³²⁸ Xu and Diwekar, ²⁵⁶ Ulas and Diwekar ³⁴⁴	Systematic uncertainty quantification methods, MOO, HSTA, GA, batch separation processes, direct approach, assessment of process controllability in view of different solvents, decomposed approach
Zhou et al. ³³³	Hybrid methods combining GA and NLP for direct solvent-process design
<i>CAMPD using continuous molecular representation</i>	
Bardow et al., ¹⁷⁵ Stavrou et al., ³³⁴ Lampe et al., ^{335,336} Schilling et al. ^{338,339}	Two- and one-stage CoMT-CAMD, use of PCP-SAFT or GPC-SAFT EoS, NLP or MINLP (in one-stage formulation) approach, postdesign property uncertainty analysis for selected molecules, physical CO ₂ capture solvents and processes, working fluids and ORC process design, incorporation of detailed heat exchanger models, and sizing of equipment
Wang and Lakerveld ³⁴⁰	PCP-SAFT using the approach of Bardow et al. ¹⁷⁵ for crystallization solvent design in view of process operating condition constraints
Roskosch and Atakan ^{176,341}	Approach similar to Bardow et al. ¹⁷⁵ with cubic EoS where molecule is represented continuously through parameters such as critical temperature, pressure, and acentric factor, application to heat pump and ORC fluids and processes
Pereira et al., ¹⁵² Mac Dowell et al., ³⁴³ Mac Dowell ³⁴⁵	SAFT-VR for mixture or blend design considering process operating optimization, direct approach, equilibrium and rate-based column models, ³⁴⁶ physical and chemical CO ₂ capture
Frutiger ³⁴²	Continuous molecular representation for CAMPD using cubic EoS and mapping of fluids considering uncertainty, details in "Molecular Design Under Uncertainty" section
<i>Property clustering approaches and reverse problem formulations</i>	
Eden et al., ²⁹² Kazantzi et al., ³⁴⁷ Eljack et al., ³⁴⁸ Chemmangattuvalappil et al., ²⁹⁶ Bommarreddy et al., ^{349,350} Chemmangattuvalappil and Eden, ²⁹⁷ Eljack, ³⁵¹ Kheireddine et al. ³⁵²	Property clusters for process and molecular design, synthesis of molecules using property clustering, advanced molecular representations, signature molecular descriptors and higher-order property prediction models, visual and algebraic CAMPD solution approaches
Ng et al. ^{353,354}	Design of biochemicals and mixtures using signature molecular descriptors and property operators, design problem posed as MILP, design of biorefinery to produce these biochemicals using superstructure representation. Mixture design approach derives additive components using property operators, which are mixed in a fuzzy MOO formulation

using the GC SAFT- γ -Mie EoS. They consider linear, saturated, and unsaturated aliphatic hydrocarbon families and solve the MINLP problem using OA/ER/AP.

p1570 Cignitti et al.³¹⁵ propose the simultaneous design of single working fluids and a standard ORC process. They investigate one simultaneous design case in an MINLP formulation using a global solver. A decomposed case is also investigated using BARON (global optimizer) to handle discrete variables (i.e., molecular structure) in the MILP CAMD and CONOPT³¹⁶ to address the process NLP that corresponds to each fluid. This work applies the general framework for simultaneous product and process design proposed by Cignitti et al.,³¹⁷ where a hierarchical procedure is developed and implemented for the gradual addition of modeling and decision-making details in the overall problem.

p1575 Stijepovic et al.³¹⁸ propose the simultaneous selection of working fluid mixtures and the design of ORC process flow sheets. The design problem is approached in an MINLP formulation where an inclusive and flexible ORC model is automatically evolved by BARON. The basic building block of the model is the ORC cascade which consists of a heat extraction, a power generation, a condensation, and a liquid pressurization section. The aim of the optimization is to determine the optimum number of ORC cascades, the structure of the heat exchanger network shared among different cascades, the operating conditions, and the working

fluid used in each cascade in order to identify an overall ORC structure that maximizes the power output. A unique feature is that different working fluids may be used in ORC cascades. The working fluids are selected from a prespecified set of options. In the area of working fluid and flow sheet design, Thierry et al.³¹⁹ extend their blend design work by considering two different ORC flow sheets; the two-stage series ORC where a heat source is shared between the vaporizers and the two-stage cascade where the condenser of the first ORC acts as the vaporizer for the second. They formulate an MINLP problem where complex constraints are represented using GDP.

s0260 Stochastic Optimization Approaches

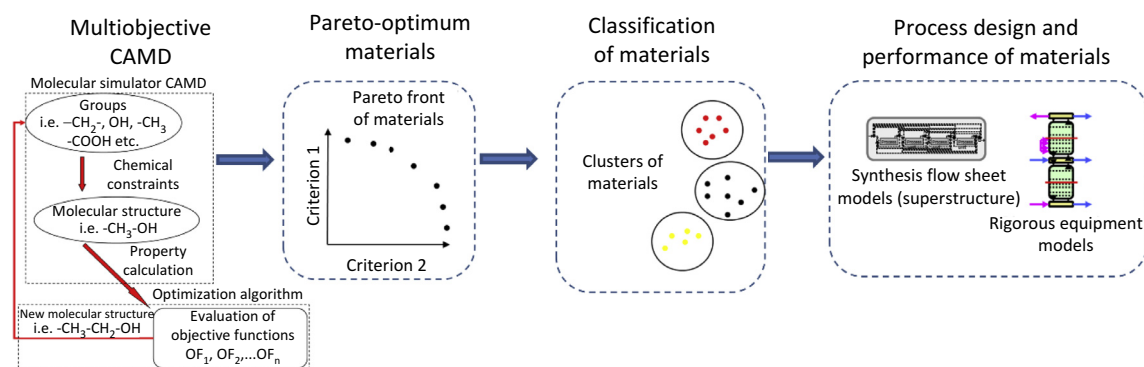
p1580 The use of stochastic optimization approaches in CAMPD generally enables the synthesis of wide flow sheets in integrated molecular and superstructure-based design approaches. An early pilot effort searched molecular design options and process configurations simultaneously.³²⁰ While the model complexity and combinatorial nature of the integrated model proved computationally challenging, the results indicated the importance of developing integrated CAMPD solutions rather than employ CAMD and process synthesis sequentially. A computationally much more efficient approach to CAMPD was later proposed by Papadopoulos and Linke,^{161,163} in which the CAMPD problem is decomposed into a CAMD and a process synthesis problem, the latter screening only those potentially optimal candidate molecules alongside the process design and operational variables. Since multiple molecular properties impact process performance and the desired directions for those properties are generally understood from a process design perspective, that is, whether minimization or maximization of a property value would lead to enhanced process performance, they propose a decomposed CAMPD approach employing MOO at the CAMD stage in order to identify a set of Pareto-optimum molecules considering multiple properties relevant to process performance as objectives. The identified set of Pareto-optimal molecules can then be considered into a process design optimization stage (overview of main steps in Fig. 5). The multiple objectives considered at the CAMD stage may include pure component properties calculated directly by GC models, mixture equilibrium properties calculated using EoS, and/or activity coefficient models and process models of lower fidelity. The design information included in the obtained Pareto-optimal set is then systematically exploited in a process design stage which may employ models of higher fidelity or models which are able to capture a very wide range of potential design options (e.g., superstructure-based synthesis models). The introduction of the optimal set of molecules into the process design stage capitalizes on the available design information through the formation of molecular clusters,³²¹ thus partitioning the molecular set into smaller compact groups of similar molecules. A representative molecule from each cluster is introduced into the process design stage as a discrete option. As a result, only a limited number of molecules are selected from the Pareto front and introduced into process design until a cluster is identified which contains very few molecules of higher process performance than all the other available options. SA is used as the optimization algorithm, but the approach could readily be implemented with other stochastic search schemes such as TS or ACO.

p1585 The approach of Papadopoulos and Linke^{161,321} is extended to CAMPD design of reactive separation solvents and flow sheets (extractive fermentation)^{163,322} and to solvent and flow sheet design of separation systems where multiple different solvents, separation tasks, and technologies are decision options.³²² The approach is further applied in the design of pure and mixed working fluids and ORC processes^{323,324} as well as in the design of chemisorption CO₂ capture solvents and absorption/desorption systems.³²⁵ A parallel implementation is also developed by Papadopoulos and Linke³²⁶ for high-performance computing environments,³²⁷ including automations and services required for unattended submission and execution of complex workflows. Papadopoulos et al.¹⁶² extend the framework to account for solvent-process design in view of variability in process operation and to propose optimum control strategies on top of optimum solvents and process characteristics (see also “Molecular, Process and Control Design” section). The work of Papadopoulos and Linke^{161,321} adopts the molecular representation of Marcoulaki and Kokossis^{268,269} which enables optimal CAMD by SA. It proposes a generalized decomposition-based CAMPD approach motivated by the work of Linke and Kokossis.³²⁰ Other works have attempted parts of the above CAMPD problem. These are summarized later. It should be noted that structural process optimization (process synthesis) is rarely attempted as part of CAMPD approaches.

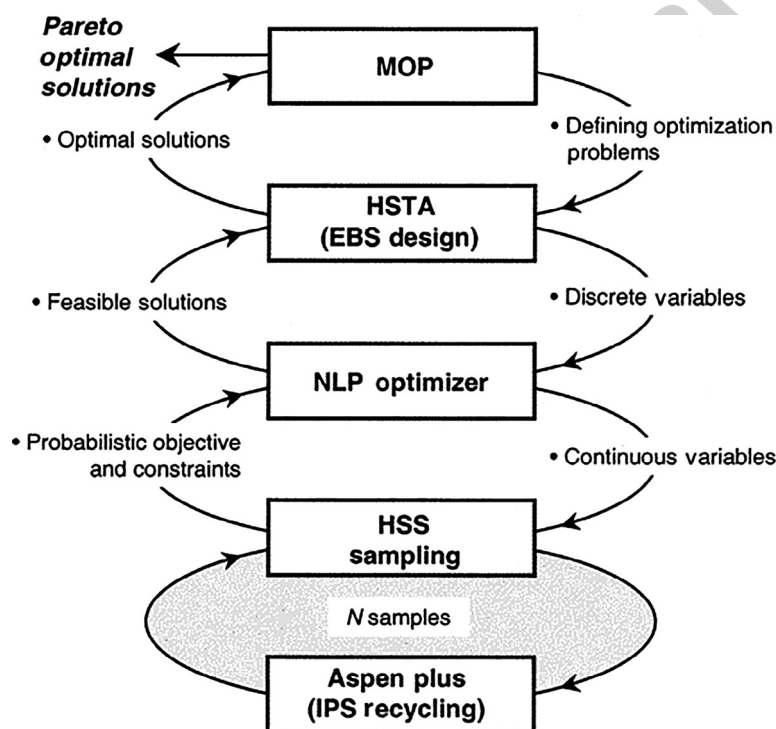
p1590 Kim and Diwekar²⁷³ propose a CAMPD framework consisting of several different stages, shown in Fig. 6. The overall CAMPD problem is formulated as multiobjective programming within an outer loop employing the Hammersley stochastic annealing (HSTA) algorithm.¹⁶⁷ The outer loop handles discrete optimization parameters (e.g., molecular groups), while an inner loop handles continuous optimization parameters (e.g., process operating parameters) through an NLP solver. Uncertainty is also considered through the Hammersley sequence sampling (HSS) method and propagated into a process simulator which simulates the given values for the discrete, continuous, and uncertain parameters of the optimization problem. It is implemented on the simultaneous integration of environmentally benign solvent (EBS) selection and in-process solvent recycling. The CAMPD approach is further extended to account for process synthesis, considering different batch distillation column configurations together with solvent design.³²⁸ The approach is further adapted to a GA by Xu and Diwekar.²⁵⁶

p1595 Qadir et al.³²⁹ use PPC-SAFI³³⁰ for simultaneous solvent selection or blend design and process design. An outer loop employs MATLAB[®]³³¹ to select solvent properties and process conditions which are introduced into an inner loop employing ASPEN Plus[®]³³² where the solvent flowrate is optimized. The solvents are selected from a prespecified list of few options. The outer loop employs GA whereas the inner loop optimization is performed using SQP.

p1600 Zhou et al.³³³ propose a hybrid optimization approach for CAMPD that combines GA with a deterministic solver. The GA optimizes the discrete molecular variables and a gradient-based deterministic algorithm solves the continuous NLP of the process at fixed molecular variables as proposed by the GA. The method is implemented to the design of solvents for separation of acetone from air.



0030 **Fig. 5** Main steps of the decomposition-based CAMPD approach proposed by Papadopoulos and Linke.^{161,321} Reproduced from Papadopoulos, A.I.; Linke, P.; Seferlis, P., *Integrated Multiobjective Molecular and Process Design: Operational and Computational Frontiers*. In *Tools for Chemical Product Design: From Consumer Products to Biomedicine*. Computer Aided Chemical Engineering. Martin, M., Eden, M. and Chemmangattuvalappil, N. Eds.; Elsevier: Amsterdam, **2016a**, Vol. 39, Chapter 11, pp. 269–313 by permission of Elsevier.



0035 **Fig. 6** The CAPMD approach proposed by Kim and Diwekar.²⁷³ Reproduced from Kim K.J.; Diwekar U.M., *Ind. Eng. Chem. Res.* **2002c**, 41(18), 4479 by permission of ACS.

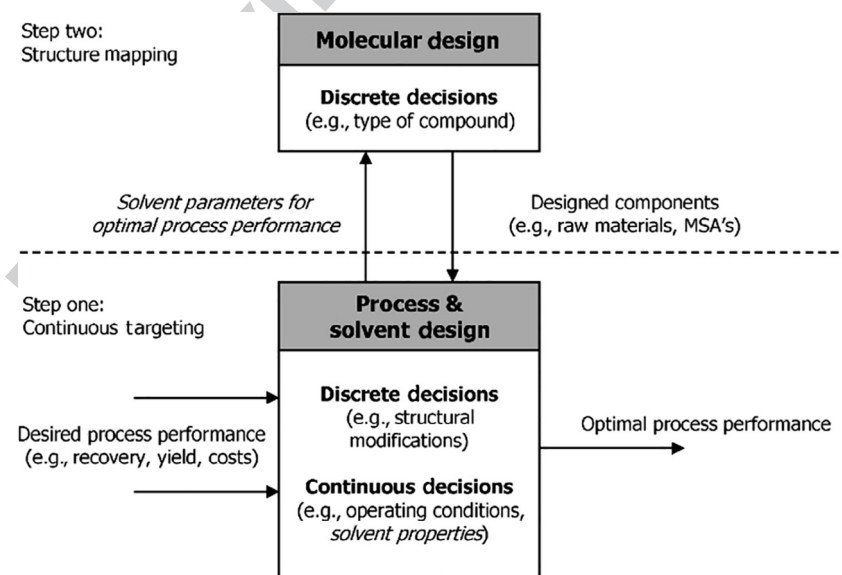
s0265 Continuous Molecular Representations

p1605 A continuous molecular representation approach was proposed in the first-recorded efforts to address CAMD through optimization.^{173,174} However, it was hindered by the inherently discrete nature of UNIFAC groups. The inherent utilization in SAFT-based EoS of molecular characteristics such as molecular segments, diameters, etc. greatly facilitated the use of continuous molecular representations in CAMPD. Bardow et al.¹⁷⁵ propose the CoMT-CAMD formulation where the problem is decomposed in two stages, namely the continuous targeting stage and the molecular mapping stage (Fig. 7). The continuous targeting stage is realized through an appropriate molecular model; it is the perturbed chain statistical associating fluid theory (PC-SAFT) EoS¹⁰⁵ which considers molecules as chains of spherical segments that interact. In this case, the molecular parameters are geometric (size and elongation), energy-related (dispersive attraction, association energy, and association volume) and polar (point dipole and

quadrupole moment). Continuous process parameters are also used as decision options. The resulting molecule is hypothetical in the sense that it does not necessarily coincide with a real fluid or satisfy chemical constraints (e.g., zero free bonds). The authors address this issue by postulating a mapping stage where the parameters of the optimum molecule are compared with the parameters of real molecules contained in a database. The proximity of the optimum and real molecules is evaluated based on the average absolute distance in parameter space, using a Taylor series approximation. A major advantage of the method is the avoidance of challenging MINLP, on top of the simultaneous solution of CAMPD. The approach is implemented in CO₂ capture using physical absorption, and extended by Stavrou et al.³³⁴ to enable more efficient mapping. Lampe et al.³³⁵ further apply the idea in the design of working fluids and ORC systems. The method is further extended by Lampe et al.³³⁶ where the mapping stage is implemented using CAMD, employing the GC GPC-SAFT EoS.³³⁷ Schilling et al.³³⁹ extend the work of Lampe et al.³³⁶ by incorporating the continuous targeting and the GC-CAMD mapping stages into one stage, calling the overall approach one-stage CoMT-CAMD. Initially, all integer variables are relaxed to calculate a hypothetical optimum by solving the resulting NLP (CoMT-CAMD). Based on an OA approach, a mixed-integer programming (MIP) master problem is then solved to identify an integer point with similar properties as the hypothetical optimum. The MIP master problem defines a tighter upper bound on the objective function. Afterward, the integer variables are fixed, and the resulting NLP is solved. The solution represents a lower bound on the objective function. Convergence of the MINLP is achieved based on OA algorithmic principles. Schilling et al.³³⁹ apply this approach in the design of working fluids and ORC using detailed process models. Particularly, heat exchangers considering different heat transfer types are modeled and sized. Transport properties needed for this purpose are modeled and predicted through PCP-SAFT.

p1610 Wang and Lakerveld³⁴⁰ use a two-stage approach, similarly to Bardow et al.¹⁷⁵ for crystallization solvent design using process operating condition constraints. Roskosch and Atakan¹⁷⁶ perform a reverse engineering design of the working fluid and a heat pump process using a cubic EoS. Fluids are represented continuously in the optimization problem through critical temperature and pressure, acentric factor, and liquid heat capacity. The problem is solved using an NLP solver. The resulting optimum solution is then identified based on its proximity to fluids available in a database. This approach is also used by Roskosch and Atakan³⁴¹ for ORC CAMPD. Frutiger³⁴² also employs a similar approach in terms of using a cubic EoS followed by a molecular mapping stage. The CAMPD problem is solved using a systematic sampling approach (further details are reported in "Molecular Design Under Uncertainty" section).

p1615 Pereira et al.¹⁵² use SAFT-VR¹⁰² for the design of a solvent mixture for CO₂ capture using physical absorption. The solvent design space includes linear alkanes. The problem is reduced to the continuous domain by optimizing over the average number of carbon atoms in the blend. Gopinath¹⁷⁹ provides an extensive discussion regarding merits and shortcomings of the approach. Mac Dowell et al.³⁴³ propose the design of a solvent blend for CO₂ capture together with process characteristics in a chemical absorption process. A rate-based absorber model is used, considering inlet temperature, solvent concentrations, and lean solvent flowrate as design variables.



r0040 **Fig. 7** The CoMT-CAMD approach for integrated molecular and process design.¹⁷⁵ This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for noncommercial purposes. Reproduced from Bardow, A.; Steur, K.; Gross, J., *Ind. Eng. Chem. Res.* **2010**, *49*(6), 2834–2840.

s0270 Property Clustering and/or Reverse Formulations

p1620 Eden et al.²⁹² proposed a two-stage approach, whereby constitutive variables (e.g., mass transfer rates, equilibrium constants, and reaction rates) are decoupled from process models (approach depicted in Fig. 8). The first stage uses the constitutive variables as design variables in a reverse problem formulation where their values are determined based on a desired performance target. In the second stage, the values of the constitutive variables are given in order to determine the optimum molecular and process characteristics. The property clustering approach is used to determine the constitutive variables which affect process operation and can easily be used as targets in a CAMPD problem. In this work, the molecules are designed using a generate and test approach.

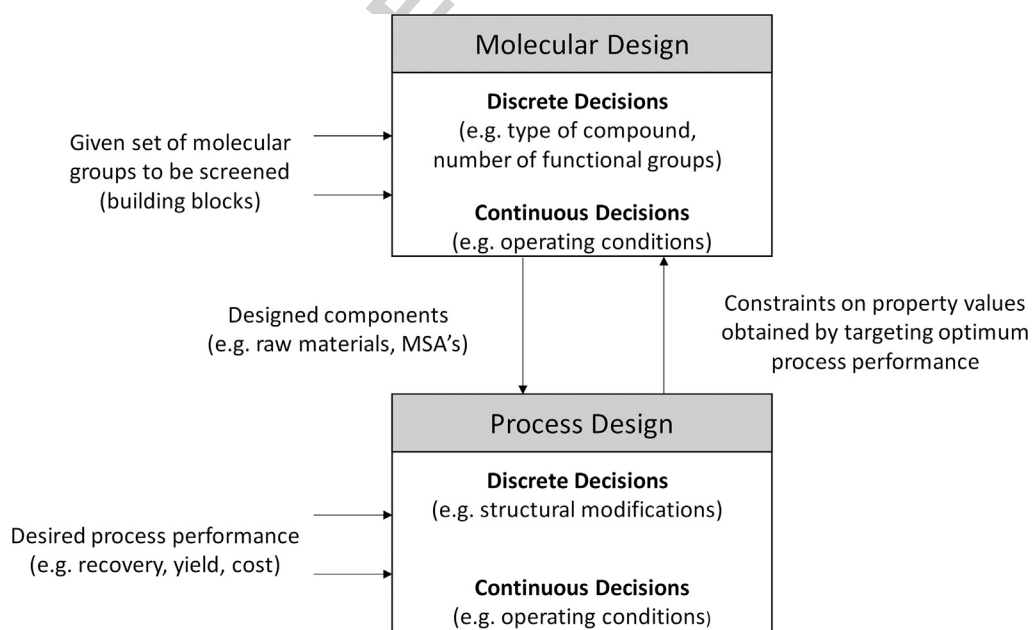
p1625 Kazantzi et al.³⁴⁷ extend the work of Eden et al.²⁹² by proposing the mapping of molecular structures in the form of functional groups on property cluster diagrams. They use GC models to calculate properties of new structures that result from small modifications of existing ones mapped on the diagrams. Eljack et al.³⁴⁸ and Eljack and Eden²⁹³ generalize this approach by proposing rules which can be used for the synthesis of molecules by mixing molecular fragments on property clustering diagrams. Chemmangattuvalappil et al.^{295,296} and Chemmangattuvalappil and Eden²⁹⁷ propose new approaches to represent molecular structures and to calculate their properties, which are all considered in the context of the CAMPD approach originally developed by Eden et al.²⁹². Details are provided in "Property Clustering Approaches" section. Bommareddy et al.^{349,350} use the developments proposed by Chemmangattuvalappil et al.²⁹⁶ to address the simultaneous process and molecular design problem using higher-order groups within an algebraic context. Kheireddine et al.³⁵² further use a property clustering approach to identify effective solvents or blends for a process that recycles lubricating oils.

p1630 In a slightly different line of work, Ng et al.³⁵³ use the signature-based molecular design technique of Chemmangattuvalappil et al.²⁹⁶ and the property operator concept of Shelley and El-Halwagi²⁹⁰ to design optimal biochemicals, which are then used as targets to synthesize optimum biomass conversion pathways in an integrated biorefinery. The latter is approached through a superstructure which accounts for different pathways, intermediates, and production technologies. The molecular design problem is formulated as MILP. The method is then extended by Ng et al.³⁵⁴ to CAM^bPD for biorefineries. The approach includes an original mixture design contribution which is discussed in "Mixture and Blend Design" section.

s0275 Methodologies for Various Problem Classes

s0280 Multiobjective Optimization

p1635 The main MOO methods used in molecular design include the weighting^{161,355} and the constraint method.^{356–358} The weighting method assigns weights to each objective function, whereas all weight and objective function terms are then aggregated into a sum that is optimized. The constraint method solves repeatedly a number of SOO problems. One objective is optimized each time, while the rest are turned into inequality constraints. Several different values need to be tested for the right-hand side of the inequality constraints in each optimization problem.³⁵⁷ Both methods have advantages and disadvantages. In the weighting method, the weights need to be varied otherwise the optimization will identify only specific sections of the overall Pareto front. If different



r0045 **Fig. 8** The CAMPD approach of Eden et al.²⁹² Reproduced from Eden M.R.; Jorgensen S.B.; Gani R.; El-Halwagi M.M., *Chem. Eng. Process.* **2004**, 43, 595 by permission of Elsevier.

weights are used in independent optimizations, then the computational effort will increase. This may be avoided by employing a variable weight sampling approach during optimization.¹⁶¹ The constraint method is computationally more expensive, but it can offer a more controlled probing of the Pareto front. A more detailed discussion is provided in Fu and Diwekar.³⁵⁷ Different formulations have also been proposed employing fuzzy representations,³⁵⁹ as discussed here. All developments are summarized in Table 10.

p1640 Kim and Diwekar²⁷³ employ in MOO CAM^uPD the constraint method, using the MINSOOP algorithm which minimizes the number of SOO problems.³⁵⁷ The minimization is achieved by efficient sampling of the number of SOO problems that need to be solved using the Hammersley sampling sequence. The approach is used in the context of a hybrid algorithm which combines HSTA with internal NLP.

p1645 Papadopoulos and Linke¹⁶¹ propose the use of a variable weighting method within a SA algorithm. The weights are randomly drawn from a uniform distribution and vary within the optimization in order to provide an efficient sampling of the potential objective function combinations and to identify a good approximation of the Pareto front. The algorithm keeps track of the Pareto front through a dynamically updated archive. A set of archival rules, implementing Pareto principles, updates the archive as appropriate. The approach is further used as part of the CAM^{bu}PD and CAMPCD frameworks of Papadopoulos et al.³²⁴ and Papadopoulos et al.¹⁶²

p1650 Xu and Diwekar²⁵⁶ propose the multiobjective efficient GA (MOEGA) for CAM^uPD using the weighting method. The authors employ random weights during optimization, as in Papadopoulos and Linke,¹⁶¹ but implement the Hammersley sampling sequence to support the random sampling of the weights.

p1655 Limleamthong et al.³⁶⁰ propose the MOO selection of chemicals using the data envelopment analysis (DEA) approach. A parameter called relative efficiency score is used to evaluate the relative changes caused to all molecular performance indices by assigning different weight values. As a result, weights are not fixed in DEA and the resulting relative efficiency scores indicate the efficient (Pareto-optimum) molecules. This is approached by solving a linear programming model (the primal problem) which maximizes the scores. An additional feature of DEA is the subsequent solution of the dual linear programming problem. The latter provides targets for improvements necessary so that suboptimal solutions may enter the Pareto front.

p1660 Ng et al.³⁵⁹ propose MOO CAMD using a fuzzy representation. A variable called "degree of satisfaction" is defined which takes values in the range [0, 1] based on satisfaction of upper and lower limits for each objective function. The degree of satisfaction is calculated as a function of these limits for each objective function. When a property to be minimized approaches its lower limit, the degree of satisfaction approaches 0 and vice versa. The degree of satisfaction therefore takes a value which indicates the distance from the upper and lower limits for each objective function. Since it also takes values within [0, 1], the optimization problem is formulated as a maximization or minimization of the degree of satisfaction subject to constraints which account for the way that it varies for each objective function. This is called a Max–Min aggregation approach and drives the optimization to identify solutions based on the weakest or worse objective, for the case of a maximization problem. This means that the other objectives may be overly relaxed, and/or it is likely to obtain two designs that have the same degree of satisfaction for one of the objectives, that is, that have the same distance from the limits of the corresponding objective. To address this limitation, a second optimization step is introduced, called the two-phase approach, where the objective is to optimize for the sum of the degrees of satisfaction over all objectives, subject to an additional constraint on the degree of satisfaction of the least satisfied objective identified in the Min–Max phase. While these are the two main steps of the multiobjective problem formulation, a prior step involves the solution of an

t0055 **Table 10** Key features of molecular design approaches employing MOO

References	Problem class and method	Optimization approach
Kim and Diwekar ²⁷³	CAM ^u PD, constraint method	Hybrid of HSTA and NLP
Papadopoulos and Linke, ¹⁶¹ Papadopoulos et al. ^{162,324}	CAMD, weighting method, also used in CAM ^{bu} D and CAMPCD frameworks	SA
Xu and Diwekar ²⁵⁶	CAM ^u PD, weighting method	GA (MOEGA)
Limleamthong et al. ³⁶⁰	Selection of molecules using DEA, variable weights, identification of Pareto front and of improvement targets for suboptimal molecules	Linear programming
Ng et al. ^{354,359}	CAMD, fuzzy representation method	MILP
Khor et al. ³⁶¹	CAM ^u D, fuzzy representation method	Disjunctive programming
Ooi et al. ^{362,363}	CAMD, weighting method, AHP, fuzzy representation of weights	MILP, NLP
Ten et al. ^{364,366}	CAMD, fuzzy representation, AHP for assignment of weights in postanalysis of designed molecules	MILP, disjunctive programming
Burger et al. ²³⁷	CAMPD, weighting method	MINLP, sandwiching algorithm
Dev et al., ³⁶⁸ Dev ³⁶⁹	CAM ^R D, constraint method, AUGMECON and AUGMECON2	MINLP
Schilling et al. ³³⁸	1-stage CoMT-CAMPD, constraint method	MINLP
Valencia-Marquez et al. ³⁷⁶	CAMPD, constraint method	MINLP

optimization problem to identify upper and lower limits for objective functions with unknown limits. This is necessary in order to define the degree of satisfaction. The entire algorithm is therefore a bi-level optimization problem including the identification of upper and lower limits, if necessary, prior to solving the Max–Min problem. The approach is illustrated in MILP CAMD formulations. The approach is extended by Ng et al.³⁵⁴ to the design of mixtures, discussed in “Mixture and Blend Design” section. Khor et al.³⁶¹ further use this approach for CAM^uD of alternative solvent to extract oil from palm pressed fiber. The problem includes the use of disjunctions in the optimization formulation to address the use of discontinuous penalty functions in health and safety indices used in solvent design.

p1665 Ooi et al.³⁶² propose the analytic hierarchy process (AHP) to identify the weight values required in the weighting method while smoothing the subjective judgment of the decision-maker. The AHP involves the development of a pairwise comparison matrix which includes the values describing the relative importance among objectives. The authors propose to use a 9-point scale system to assign such values. A sensitivity analysis stage is included which allows perturbation of the initially assigned values within a predefined range. The weighting factor of each property is then obtained by calculating the principal eigenvector of the pairwise comparison matrix. A consistency verification ratio is proposed which comprises the consistency ratio over the consistency index of a random pairwise comparison matrix. The consistency verification ratio measures the distance of the principal eigenvalue from the order of the matrix. This means that for higher distance, the inconsistency is higher in the judgment of the decision-maker regarding the relative importance of the objective functions.

p1670 Ooi et al.³⁶³ propose a fuzzified approach within the AHP. Instead of using the 9-point scale system to assign importance values in the pairwise comparison matrix, the authors use triangular fuzzy numbers where the lower and upper values may be calculated by assigning a value in parameters called “degree of confidence of the decision maker.” The optimal priority vectors (i.e., the weights of the MOO aggregate function) that approximate the solution ratio in the pairwise comparison matrix are then calculated through an NLP problem formulation which simultaneously determines the consistency index. The latter indicates the degree of satisfaction of all computed weight ratios that would satisfy the initial fuzzy judgments. The problem is formulated as disjunctive programming.

p1675 Ten et al.³⁶⁴ propose a MOO product design framework which considers safety and health aspects. The MOO CAMD formulation employs a fuzzy representation which is based on the same principles proposed by Ng et al.³⁵⁹ The designed molecules undergo a postanalysis which is based on evaluation of health and safety indices. Scores are assigned in several individual indices, and the final ranking is based on a weighted sum of all indices. The AHP approach proposed by Aminbakhsh et al.³⁶⁵ is used to determine the weights. Ten et al.³⁶⁶ extend the approach by introducing disjunctive programming to assist the allocation of index scores to the molecules depending on their property values.

p1680 Burger et al.²³⁷ propose a MOO CAMPD approach within an MINLP formulation. The authors use the weighting method in a deterministic sandwiching algorithm proposed by Bortz et al.,³⁶⁷ which is modified and used as a heuristic to achieve a compromise between performance and sampling quality. The authors describe a step-by-step procedure pertaining to the identification of Pareto points, noting that only the convex hull of the Pareto-optimal set is determined. Furthermore, they note that mixed-integer (combinatorial) MOO is a rather new field of research using heuristic treatment to approximate the Pareto front, while for nonconvex MOO problems involving only continuous variables there are efficient deterministic optimization approaches.

p1685 Dev et al.³⁶⁸ and Dev³⁶⁹ employ the AUGMECON³⁷⁰ and AUGMECON2³⁷¹ algorithms which are based on the augmented ϵ -constraint method. AUGMECON guarantees generation of Pareto-optimal solutions. In case of MOO MINLP problems, the guaranteed Pareto-optimal solutions may be locally optimal. But, this is better than the possibility of generating local weak Pareto-optimal using the conventional ϵ -constraint method. The authors employ the DICOPT^{372,373} solver in GAMS³⁷⁴ to solve the MINLP problems which are part of AUGMECON-based schemes.

p1690 Schilling et al.³³⁸ implement MOO CAMPD in their one-stage CoMT approach. They use the normal constraint method³⁷⁵ with a local MINLP solver. Non-Pareto-optimal solutions are also identified, hence the normal constraint method is repeated where the direction of movement of the normal constraint is changed. Both resulting Pareto fronts are superimposed and filtered according to Pareto dominance criteria to increase the accuracy of the solution. Valencia-Marquez et al.³⁷⁶ also propose a CAMPD approach for the design of CO₂ capture ionic liquids where MOO is employed using the ϵ -constraint in a deterministic MINLP formulation.

s0285 Mixture and Blend Design

p1695 The key features of all mixture and blend design approaches are illustrated in Tables 11–13. The contributions are separated based on the applications. Table 11 includes mixture and blend design approaches applied in solvent design, Table 12 includes applications in heat exchange fluids, and Table 13 includes applications in polymers, fuel and lubricant additives, and formulated products.

p1700 Klein et al.²¹² propose a blend design approach to identify the optimum number of solvents in the blend and the corresponding concentration. The solvents forming the blends are selected from a set of prespecified options. The problem is solved using the SLRP approach.²¹³ The method is further used by Gani and Fredenslund²¹⁴ in a four-stage approach where the resulting mixtures are determined again from a set of prespecified solvent options.

p1705 Buxton et al.²⁰³ propose a CAM^bPD methodology which includes the design of binary mixtures where the structure of both components is determined through optimization, together with the mixture concentration. Mixtures are considered in process design only if they satisfy specific thermodynamic performance tests, hence the problem is decomposed. The aim is to identify solvent mixtures and absorption process operating conditions that enable environmental impact minimization.

10060 **Table 11** Main features used in solvent mixture or blend design approaches

References	Problem class and approach summary	Optimization approach
Klein et al., ²¹² Gani and Fredenslund ²¹⁴	Blend design to determine optimum number of components in blend and the corresponding concentrations, blends generated as combinations from a prespecified set of solvents	SLRP optimization algorithm, direct optimization
Buxton et al. ²⁰³	CAM ^b PD, binary mixture, component structures and mixture concentration identified through optimization together with process features, decomposition approach	MINLP addressed with GBD algorithm
Achenie and Sinha ²²⁴	Blend design, selection from prespecified list and identification of optimum composition, process constraints	MINLP, global optimization with IA
Siougkrou et al. ³⁷⁷	Simultaneous identification of blend concentration and cosolvent (selected from a prespecified set) with process characteristics	MINLP formulation, complete enumeration
Zarogiannis et al., ¹⁶⁵ Papadopoulos et al. ¹⁶⁴	Identification of binary blend composition and concentration, selection of solvents from a prespecified set, consideration of uncertainty in property predictions, decomposition approach	MOO formulation, complete enumeration
Jonuzaj et al., ²³⁸ Jonuzaj and Adjiman ²⁴⁰	Blend design problem to identify optimum number of components and concentrations from a prespecified set of solvents	GDP with BM or HR approach, direct optimization
Kheireddine et al. ³⁵²	Identification of feasible solvents from prespecified options and concentration in a two-stage approach, pure components are evaluated first and are then combined to form mixtures, feasible region determined from process requirements	Property clustering
Van Dyk and Nieuwoudt ²⁵⁴	CAM ^b PD formulation for identification of optimum structure of mixture components and concentrations, the number of components in mixtures can also be identified through a sequential increase (i.e., not a degree of freedom in optimization), decomposition approach	GA
Karunanithi et al., ²²⁶ Sinha et al. ²²⁵	CAM ^b PD, partial mixture design, one component is fully specified, the structure of the other component is designed together with mixture composition, decomposition approach	MINLP, global optimization with IA
Karunanithi et al. ²²⁷	CAM ^b PD, the structure of both components needs to be determined in a binary mixture together with mixture composition, decomposition approach	MINLP
Heintz et al. ¹⁵³	CAM ^b D, partial mixture design, one component is fully specified, the structure of the other component is designed together with mixture composition	GA, direct optimization
Chong et al. ³⁷⁸	CAM ^b D, identification of optimum ionic liquid structures and concentration in a two-stage approach, pure components are generated first and are then combined to form mixtures	Property clustering approach
Austin et al. ¹⁴⁹	CAM ^b D approach to identify optimum structure of mixture components and concentrations, the problem formulation can consider an upper number of components, it is applied to identify the structure and concentration of one cosolvent in different formulations, direct optimization.	Hybrid DFO with BARON for continuous nonconvex problem, direct optimization.
Austin et al. ¹⁴⁹	QM-CAM ^b D approach to identify optimum structure of mixture components and concentrations, applied to the design of binary mixtures.	Optimization approach similar to Austin et al. ¹⁸² , direct optimization.
Jonuzaj et al. ²³⁹	CAM ^b D complete mixture design formulation to identify optimum number of components, their structures, and concentrations.	GDP formulation, direct optimization.

p1710 Karunanithi et al.²²⁶ propose a CAM^bD approach comprising several stages to gradually evaluate constraints such as chemical feasibility, pure component properties, mixture properties, and so forth. Molecules that satisfy the constraints in each stage are selected to proceed in the subsequent screening stage, resulting in a gradual reduction of the available molecular set. The presented applications consider a binary solvent blend where water is prespecified as the first solvent. The approach is used to design the cosolvent, while an optimization search is performed to determine the optimum concentration of the binary mixture. Karunanithi et al.²²⁷ extend this approach to the design of a binary mixture, where both the solvent and the antisolvent need to be designed for a crystallization process. A similar approach is also proposed by Sinha et al.²²⁵, however a more efficient algorithm, namely IA optimization, is used to address the nonconvexities of the thermodynamic prediction models.

p1715 Jonuzaj et al.²³⁸ and Jonuzaj and Adjiman²⁴⁰ propose a blend design approach to identify the optimum number of components in the blend and the corresponding concentrations. The components in the blend are selected from a list of prespecified options. Jonuzaj et al.²³⁹ address the complete CAM^bD problem where the number of components, their structures, and concentrations result from the optimization (see also see "Deterministic optimization approaches in molecular design" section).

p1720 While all the earlier-mentioned works utilize deterministic optimization methods, a GA is employed by Van Dyk and Nieuwoudt²⁵⁴ to address the design of multicomponent mixtures through a decomposed approach. The method starts by implementing CAMD for the design of single solvents and selecting the 10 best options. Each of these solvents is then enhanced with 1000 more solvents selected randomly to form a population of 10.000 binary mixtures. For each mixture, the first solvent is

10065 **Table 12** Main features used in heat exchange fluid blend or mixture design approaches

References	Problem class and approach summary	Optimization approach
Duvedi and Achenie, ²¹⁸ Churi and Achenie ²¹⁹ Lee et al. ³⁸⁰	Binary blend design to determine composition and concentration from prespecified set of options, process operating constraints	MINLP, OA/AP, direct optimization
Mavrou et al. ^{382,383}	Blend design of mixture concentrations and process features considering prespecified options from CAM ^b PD and literature, consideration of nominal and off-design operation of ORC, decomposition approach	MOO, nonlinear sensitivity analysis, complete enumeration
Molina-Thierry and Flores-Tlacuahuac ³¹²	Blend design of number of mixture components, mixture composition and concentration with process operation, prespecified set of pure fluids to form mixture combinations, application to working fluid and ORC process design	NLP formulation, CONOPT in GAMS, direct optimization
Thierry et al. ³¹⁹	Blend design of number of mixture components, mixture composition and concentration with process structure and operation, application in ORC	NLP for single and series ORC model with CONOPT, disjunctive programming for ORC cascade, extended mathematical programming (EMP) ³⁸⁴ using the BM reformulation and then solved via the SBB ³⁸⁵ solver, direct optimization
Santos-Rodriguez et al. ³¹³	Blend design of number of mixture components, mixture composition and concentration with process operation, prespecified set of pure fluids to form mixture combinations, application to working fluid and ORC process design under uncertainty	NLP formulation, CONOPT in GAMS, direct optimization
Solvason et al. ³⁸¹	CAM ^b D using property clustering techniques, generation of all possible combinations of groups in cluster space, gradual increase in number of mixtures with specification of structure of mixture components and concentration	Property clustering, complete enumeration
Papadopoulos et al. ³²⁴	CAM ^b PD, two-stage approach, screening phase to identify first mixture component, design phase to identify matching second component and concentration, ORC working fluids and processes, decomposition approach	MOO SA
Cignitti et al. ³¹⁷	CAM ^b PD for simultaneous binary mixture and refrigeration process design, decomposition approach	MINLP formulation, SBB solver in GAMS

kept fixed, while the second solvent is allowed to evolve through optimization to identify and finally select the 100 best overall binary mixtures. The method is repeated by assigning 100 randomly selected solvents to each binary mixture, hence forming 10.000 tertiary mixtures, while a fourth component is also subsequently considered in the same manner. After completion of quaternary mixtures, all solvents within each mixture are allowed to evolve freely. The method is applied in the design of binary solvent mixtures for extractive distillation processes.

p1725 Austin et al.¹⁴⁹ propose a CAM^bD approach for the design of mixtures with the aim to identify the optimum structure of each component and mixture concentration, while the approach could also consider an upper number of components in the mixture. The method combines numerous different tools, including DFO, the AMODEO framework,¹⁵⁸ and BARON to handle the nonconvex mixture concentration. The approach starts with the generation of trial points by the DFO to identify upper and lower property bounds AMODEO is used to design structures within these bounds, at a minimum distance from the trial points, and the mixture concentration is optimized for these structures. Higher-order GC models are used for improved property prediction of structures. The approach is extended in QM-CAM^{bRD}¹⁴⁹ where the properties are calculated through COSMO-RS and COSMO-SAC.

p1730 Sioukrou et al.³⁷⁷ report an approach for the selection of a cosolvent from a list of prespecified options in a blend and for the identification of the optimum concentration and process conditions.

p1735 Zargiannis et al.¹⁶⁵ and Papadopoulos et al.¹⁶⁴ propose a systematic approach for the identification of optimum solvent mixtures in postcombustion CO₂ capture. The Pareto front of pure solvents resulting from MOO CAMD is used to exhaustively generate all possible composition and concentration combinations in terms of binary blends. Screening tests are performed accounting for boiling and melting point constraints, mixture miscibility, and CO₂ dissolution in mixtures. Binaries that satisfy constraints are evaluated using a MOO approach. Uncertainty is accounted for by considering multiple different models for property prediction in Zargiannis et al.¹⁶⁵

p1740 Chong et al.³⁷⁸ propose a visual property clustering approach for the design of mixtures of ionic liquids used as solvents in CO₂ capture. The method includes the generation of pure ionic liquid candidates using property clustering based on Chong et al.³⁷⁹ The pure component options are then used in combinations to generate binary mixtures. Once the potential pairs are generated, the

10070 **Table 13** Main features used in blend or mixture design including polymers, fuel and lubricant additives, and formulated products

References	Problem class and approach summary	Optimization approach
<i>Polymers</i>		
Solvason et al. ³⁸⁶	Blend design for identification of optimum mixture concentration, approach similar to Solvason et al. ³⁸¹	Property clustering, complete enumeration
Hada et al. ³⁰⁰	Blend design of optimum mixture concentration approach similar to Solvason et al., ³⁸¹ but with advanced chemometric techniques for structure–property identification	Property clustering, complete enumeration.
Vaidyanathan and El-Halwagi ²²⁹	CAM ^b D for determination of structure and concentration in binary mixtures	IA global optimization for nonconvex problems with lower bound tests and distrust-region method, GINO software as local optimizer, direct approach
<i>Additives in fuels or lubricants</i>		
Yunus et al. ³⁸⁷	Blend design of optimum number of components and concentrations, multistage decomposition approach eliminating options	NLP optimization
Hada et al. ²⁹⁸	CAM ^b D, partial mixture design, identification of optimum structure and concentration of biofuel additive, development of latent property GC model for CAMD	Property clustering, complete enumeration
Ng et al. ³⁵⁴	CAM ^b PD, one component is designed as additive in a fuel mixture, structure of component and concentration are optimized, approach similar to Ng et al. ³⁵⁹ with the addition of mixture constraints, decomposition approach	MILP, fuzzy MOO
<i>Formulated products</i>		
Martin and Martinez ³⁰⁸	Optimum blend concentration and manufacturing conditions, process superstructure	MINLP formulation, BARON, direct optimization
Conte et al. ^{191,192}	Multilevel CAM ^b D approach for solvent binary mixtures, considering suitable active ingredients, solvent binaries are generated, and upon constraint satisfaction concentrations are optimized based on desired criteria	Generate and test approach
Mattei et al. ^{193,388}	CAM ^b D to identify few components that satisfy desired constraints and then addition in preexisting mixtures to identify their optimum concentration, multistage product design framework for emulsions, focus on surfactants design, decomposition	Optimization problem formulation

mixture properties are formulated into property clusters, boundary points are generated based on constraints, and the procedure continuous until all feasible mixtures are identified.

p1745 In the area of fluids employed in heat exchange applications, the design of environmentally benign refrigerant mixtures is considered by Duvedi and Achenie²¹⁸ using a set of prespecified molecules. The mixture candidates are used as integer decision parameters in an optimization problem formulation that considers binary mixtures of concentration that is additionally regarded as a continuous decision parameter. The employed objective function combines the compressor displacement in a refrigeration cycle with the ozone depletion potential of the mixture. A similar approach is used by Churi and Achenie²¹⁹ on a double-evaporator refrigeration cycle to obtain mixtures that maximize cooling. Lee et al.³⁸⁰ focus on addressing both the combinatorial complexity in view of different concentration requirements in multicomponent refrigerant mixtures and the increased computational effort required due to need of refrigeration models in order to assess the performance of the refrigeration system. This work considers a mixture of four refrigerants commonly used in the liquid natural gas industry to optimize the mixture concentration. The method focuses on the use of a combined mathematical/thermodynamic screening approach to evaluate the impact of different decision parameters (e.g., mixture concentration and flow rate and pressure levels) in the shaft work required in a refrigeration cascade.

p1750 Cignitti et al.³¹⁷ propose an approach for simultaneous mixture and process design. It includes a hierarchical procedure for addition of modeling and design information. The approach is used to design a binary mixture (molecular structures of components and concentration) and a refrigeration process. Solvason et al.³⁸¹ build on the works of Eden et al.²⁹² and of Eljack et al.³⁴⁸ and formulate a product and molecular design problem. The aim is to generate the structures of all possible molecules and molecular mixtures that can be built from a specific number of groups. First, product attributes are identified/specified and then a CAM^bD problem is solved to find structures and mixtures that match these targets. The attributes are mapped onto property clustering diagrams and mixtures are identified within such diagrams. Regression techniques are used to associate product attributes with molecular structures. The optimum matching of the target properties is performed in two ways: starting from known pure components, combinations are generated and their properties are calculated in the cluster space. Binary mixtures may evolve to tertiary or higher mixtures until the targets are matched. Furthermore, new components can be generated from combinations of groups and new mixtures can be developed with existing components. Hence, the structure of the components, mixture number, and mixture concentration may be determined.

p1755 Papadopoulos et al.³²⁴ propose a CAM^bPD approach for binary mixtures. It addresses the identification of the optimum structure and concentration of the two components together with the optimum process conditions. A screening phase considers

the design of a mixture comprising a structurally feasible and a structurally infeasible component. A set of mixtures results from a MOO problem formulation. For each one of the structurally feasible components, a design phase aims to identify a structurally feasible second mixture component together with the optimum mixture concentration and process conditions. The problem is addressed again in a MOO formulation. The approach enables the identification of the best possible first component in a mixture for a particular task, prior to identifying its matching second component. The method is applied in ORC working fluid and process design. A sensitivity analysis is used as a post-CAM^bPD stage to quantify the effects of uncertainty in the obtained designs. Mavrou et al.³⁸² evaluate several mixtures obtained from CAM^bPD of Papadopoulos et al.³²⁴ together with mixtures from literature to determine the optimum mixture concentration and ORC operating conditions, in ORC operating with solar heat and heat storage. The approach considers screening stages based on complete enumeration through MOO. The method is extended by Mavrou et al.³⁸³ to identify blend concentration and process performance at off-design process operating conditions. A nonlinear sensitivity analysis is used together with MOO, analyzed in “Molecular Design Under Uncertainty” section.

p1760 Molina-Thierry and Flores-Tlacuahuac³¹² optimize the number of working fluids participating in the mixture, the type of working fluids that form the mixture, and the mixture concentration together with the ORC operating conditions. The working fluids that are used to generate mixture combinations are selected from a prespecified set of 3, 11 or 6 pure fluids in the performed case studies. Starting from a maximum number of mixture components, the optimum number of components is identified based on their concentrations. Thierry et al.³¹⁹ address the same blend and process design problem with Molina-Thierry and Flores-Tlacuahuac,³¹² but the process design considers different ORC structures, including a series flow sheet model and a cascade scheme. Mixtures with up to five components are considered, hence when one concentration goes to zero after optimization the number of components is reduced. Each ORC process configuration and blend design problem is solved separately. Santos-Rodriguez et al.³¹³ consider the problem of Molina-Thierry and Flores-Tlacuahuac³¹² within an uncertainty framework. More details are discussed in “Molecular Design Under Uncertainty” section.

p1765 In the case of polymer mixture design, the development of binary polymer mixtures is addressed by Vaidyanathan and El-Halwagi²²⁹ using the structure of both components and the mixture concentration as decision parameters, while the optimization problem is addressed using an MINLP approach with IA global optimization. Solvason et al.³⁸⁶ employ a property clustering method to determine the optimum number of chemical constituents delivering a desired property. The method is applied on the identification of optimum polymer blends considering four pure polymer compounds as the mixture components. The method shares similarities with Solvason et al.³⁸¹ A similar method applied on the identification of optimum starch blends considering three components is also proposed by Hada et al.³⁰⁰ A key feature is that property–structure models used to predict properties that become targets for mixture design are developed using advanced chemometric techniques. Hada et al.²⁹⁸ propose an approach for the design of biofuel additives using property clustering. Latent variable property models are developed for prediction of properties. These property models are based on groups for which contributions are derived in the latent variable space from principal component analysis and principal component regression. These models are used to design molecules in a GC-CAMD approach. The blend design problem pertains to the identification of the additive structure and its concentration in the mixture.

p1770 A decomposition-based approach is proposed for the design of gasoline and lubricant blends by Yunus et al.³⁸⁷ The approach starts by considering many potential combinations of prespecified pure chemical compounds which are then invoked on subsequent tests with respect to their performance against important blend properties. After each test, blend candidates are eliminated, gradually reducing the available set. Finally, the blends satisfying all constraints are evaluated regarding their economic performance in the application in which they are utilized. The blend design problem consists of identifying the optimum number of components and their concentration, allowing up to quaternary blends. The method is applied considering bio-derived chemicals forming mixtures with gasoline components as well as chemicals required in lubricants.

p1775 Ng et al.³⁵⁴ propose a multistep approach for CAM^bPD where mixtures are identified as products of biorefineries and the biorefineries are designed to produce them. The mixture steps involve, among others, the structural design of additive components which are then mixed in the main (prespecified) components. The miscibility of the mixtures is tested. The aim is to identify the optimum mixture structure and concentration, together with the optimum process structure that produces it, through a superstructure approach. The problem is addressed using the fuzzy MOO approach of Ng et al.³⁵⁹

p1780 Conte et al.^{191,192} propose a solvent mixture design approach for formulated products. Prior to solvent mixture design, active ingredients are selected, and solvents are then selected based on their suitability for the particular active ingredients. Given the constraints for the mixture target properties, the algorithm identifies the possible binary combinations of solvent that match the constraints. Mixtures that satisfy constraints go into the mixture optimization stage where the optimum concentrations are identified. The identification of solvents is based on a generate and test approach.

p1785 Mattei et al.^{193,388} propose a systematic multistage product design methodology for the design of emulsions. The approach includes the problem specification for the identification of appropriate active ingredients, solvents, and surfactants. CAMD can be used in the design of solvent candidates and surfactants. The focus is on the design of surfactants which are added in the mixtures and the identification of their optimum concentration, subject to pure component and mixture constraints. The CAMD approach is used to identify surfactant candidates that satisfy desired constraints. The latter are then introduced into mixtures with the preselected components. Martin and Martinez³⁰⁸ address the problem of identifying optimum blends and manufacturing conditions for development of detergents (see “Deterministic Optimization Approaches” section for details).

s0290 Molecular Design Under Uncertainty

p1790 The main features of molecular design approaches under uncertainty are summarized in Table 14. Maranas¹⁶⁸ proposes a systematic CAM^UD approach where probability distributions are used for the quantification of uncertainty in GC parameters and chance constraints are used in the optimization problem formulation. The overall problem formulation aims to identify the maximum value of the performance target that can be met by the stochastic performance objective with at least a certain probability value, while maintaining all the property values (constraints) within their respective lower and upper bounds with probability greater than or equal to a certain value. The author addresses the problems of stochastic property matching (SPM) under uncertainty and of stochastic property optimization (SPO) under uncertainty. SPM identifies the molecular compositions for which a scaled property violation target is minimized, guaranteed to be matched by all properties under a certain probability. SPO identifies the molecular compositions that maximize a property with a certain probability, while lower and upper bounds are satisfied for the rest of the properties under a certain probability. The original nonlinear stochastic problem formulations are transformed into deterministic MINLP with linear binary and convex continuous parts. It is found that property prediction uncertainty affects very significantly the selected molecule in the SPM formulation, while the effect is small in the SPO formulation. The method is used in Raman and Maranas¹⁰ where the molecular representation is based on topological indices.

p1795 Kim and Diwekar¹⁶⁷ propose a CAM^UD approach where uncertainty is quantified for various molecular and mixture properties (e.g. solubility parameter and activity coefficients) using probability distributions. The HSTA²⁷² algorithm is then used to design and identify optimum molecules. HSTA employs a variable probability sampling schedule, starting from fewer samples at the beginning of the optimization search which are gradually increased as the search is intensified toward promising solutions. A penalty function is used to account for the sampling error. The aim is to reduce the overall number of samples required from the underlying probability distributions. The latter is supported efficiently using the HSS approach which exhibits better uniformity properties than other conventional Monte Carlo techniques. Kim and Diwekar²⁷³ further extend this approach to CAM^UPD problems where it is combined with MOO. Xu and Diwekar^{255,256} further propose representation of uncertainty using probability distributions, adapted to the HSGA and MOEGA algorithms which are modified GA to account for uncertainty and MOO.

p1800 Folic et al.³⁸⁹ develop a CAM^URD approach which considers multiple different scenarios for the potential realizations of the uncertain parameters. The optimization problem accounts for all scenarios simultaneously by considering an objective function which minimizes a weighted average of the scenarios. The scenarios result from a global sensitivity analysis that considers the effects of the uncertainty parameters on the model outcome. The authors vary up to a certain number of uncertain parameters simultaneously and the uncertain parameter space is explored using Sobol's approach.³⁹⁰ Since the problem is linear, the solvents resulting from the sensitivity analysis are clustered, hence the convex hull is calculated and is then divided into several subareas. The center of each subarea is used to determine the scenarios. The method is used to assess the uncertainty in parameters fitted for the reaction rate constant model. The method is further used in the CAM^URPD approach of Folic et al.³⁹¹

p1805 Zhou et al.³⁹² propose a CAM^URD approach where they consider uncertainty in the parameters used to fit the reaction rate constant model. They first perform a sensitivity analysis to identify the parameters that cause the highest sensitivity in the results of the CAMD problem. The parameter uncertainty region is determined by confidence intervals. They then solve a multiscenario robust optimization problem considering uncertainty in the selected parameters. They first derive probability density functions to quantify uncertainty and they then follow a Monte Carlo sampling procedure to obtain samples. The CAM^URD problem is then solved for these samples.

p1810 Diky et al.³⁹³ incorporate in the ThermoData Engine of the National Institute of Standards a method for calculation of uncertainty in property prediction models. The covariance method is used together with empirical adjustments to address known shortcomings of this method, such as the assumptions that uncertainty is based on uncorrelated random error in source data and others. The engine is enhanced with a solvent design algorithm, hence uncertainties may be provided for the properties of the designed solvents.

p1815 Papadopoulos et al.³²⁴ propose a nonlinear sensitivity analysis approach for the designs obtained from CAMD in order to assess the sensitivity of MOO CAM^bD mixtures in the uncertainty of properties. The method is based on the development of a sensitivity matrix which incorporates the derivatives of multiple process performance measures with respect to multiple operating parameters and molecules. The sensitivity matrix constitutes a measure of the variations propagated through the nonlinear (process) model under the influence of infinitesimal changes imposed on the selected parameters. It is decomposed into major directions of variability associated with the eigenvectors corresponding to the larger in magnitude eigenvalues of the sensitivity matrix. The eigenvector of the largest eigenvalue represents the dominant direction of variability for the system, causing the largest change in the performance measures. The entries in the dominant eigenvector determine the major direction of variability in the multiparametric space and indicate the impact of each parameter in this direction. Having identified this direction, it is not necessary to explore all directions of variability (i.e., combinations of parameters) arbitrarily, hence reducing the dimensionality of the sensitivity analysis problem. The dominant eigenvector direction is then considered in the exploration of the system behavior as indicated by the change of key performance indices under simultaneous, multiple, and finite parameter variations.

p1820 Papadopoulos et al.³⁹⁴ propose an uncertainty assessment approach in the analysis of results obtained from MOO CAMD by considering multiple different models for the prediction of the same property used as performance index to further evaluate and rank-order the Pareto-optimum molecules. The approach considers multiple performance indices which are predicted by different models. An average of the indices is used to rank-order the Pareto molecules, hence its calculation is based on different model combinations. As models are changed for each property index, the average takes a different value for every model combination. As a

10075 **Table 14** Main features of molecular design approaches under uncertainty

<i>References</i>	<i>Problem class and approach summary</i>	<i>Optimization approach</i>
Maranas, ¹⁶⁸ Raman and Maranas ¹⁰	CAM ^u D, probability distributions for uncertainty in molecular properties	Chance constraint method, equivalent deterministic MINLP problem formulation, transformation of MINLP to MILP, direct optimization
Kim and Diwekar ^{167,273}	CAM ^u D and CAM ^u PD, probability distributions for molecular and mixture properties	HSTA, MOO, direct optimization
Xu and Diwekar ^{255,256}	CAM ^u D and CAM ^u PD, probability distributions for molecular and mixture properties	HSGA and MOEGA algorithm, MOO, direct optimization
Folic et al. ^{389,391}	CAM ^{uR} D and CAM ^{uR} PD, multisenario approach, scenarios determined from global sensitivity analysis, uncertainty in property model descriptors	MILP and MINLP problem formulations, direct optimization
Zhou et al. ³⁹²	CAM ^{uR} D, multisenario approach, sensitivity analysis for parameters of the reaction rate constant, generation of probability density functions and sample, stochastic molecular design solved for samples	MINLP formulation, direct optimization
Diky et al. ³⁹³	Uncertainty estimation for solvent properties based on covariance method with empirical adjustments	Knowledge-based approach
Papadopoulos et al. ³²⁴	CAM ^u PD sensitivity postanalysis that identifies the dominant directions of variability in uncertain parameters in multiparametric space, uncertainty in molecular properties	Development and analysis of nonlinear sensitivity matrix in multiparametric space
Papadopoulos et al., ³⁹⁴ Zarogiannis et al. ¹⁶⁵	CAMD and mixture selection postanalysis considering multiple different models for the prediction of the same properties, rank-order of molecules, uncertainty quantification by considering distribution of ranks	Consideration and rank-ordering of molecules/mixtures based on multiple different models for prediction of property performance indices
Martin and Martinez ³⁹⁶	Blend and process design, probability distributions, internal and external sampling, scenario reduction approach	Direct optimization using BARON
Ng et al. ³⁹⁷	CAM ^u D, uncertainty quantification through standard deviation from property model regression, incorporation of extended bounds in a fuzzy representation	MOO fuzzy representation, MILP, direct optimization
Ten et al. ³⁹⁸	CAM ^u D, uncertainty quantification through standard deviation from property model regression, incorporation of extended bounds in a fuzzy representation	MOO fuzzy representation, mixed-integer formulation, disjunctive programming
Frutiger et al. ³⁹⁹	Multilevel approach for molecular selection based on process performance under molecular property uncertainty, Monte Carlo approach for uncertainty analysis at optimum process operating point for each molecule	Systematic sampling of process design space for identification of optimum process operating point for each molecule from a database
Frutiger ³⁴²	CAMPD using a continuous molecular representation approach, mapping of virtual to real fluids considering property uncertainty, post-CAMPD selection of highly performing real fluids under uncertainty based on Frutiger et al. ³⁹⁹	Continuous molecular property space sampled systematically and propagated through process model to identify virtual working fluids and corresponding process performance
Frutiger et al. ⁴⁰²	(1) GC-based CAMPD and postanalysis of uncertainty effects on selected molecules based on Frutiger et al. ³⁹⁹ (2) Systematic uncertainty quantification of GC parameters, generation of multiple scenarios through sampling, solution of optimization-based CAM ^u PD for each scenario	(1) Sampling approach for GC-based CAMPD as in Frutiger et al. ³⁹⁹ (2) Optimization-based CAM ^u PD, MINLP formulation
Santos-Rodriguez et al. ³¹³	Uncertainty in process operating parameters, (1) blend and process design under uncertainty considering different scenarios in a stochastic programming formulation, resulting in one optimum design, (2) blend and process design identifying one optimum design for each scenario separately	MINLP, direct optimization, CONOPT solver
Andres-Martinez et al. ⁴⁰⁴	(1) CAMD and postanalysis of uncertainty using Monte Carlo approach (2) CAM ^u D and MINLP with uncertainty in linear constraints	MINLP formulation, direct optimization, see also Palma-Flores et al. ³¹¹

result, the ranking of the molecules changes as different combinations are used to calculate the average. Highly performing molecules are selected based on their frequency of appearance in the top positions. Uncertainty is quantified as a distribution of molecular ranks resulting from the use of different predicted values for the same properties. The distribution of ranks is used as a sampling tool accounting for molecules which may rank at the top with one property prediction model but may also rank lower with another. This allows for the refinement and reduction of the original data set without focusing only on the top options which

10080 **Table 15** Key features of contribution employing advanced computing infrastructures

Reference	Summary of contribution	Area of contribution
Satyanarayana et al. ^{405,406} Papadopoulos and Linke, ³²⁶ Papadopoulos et al. ¹⁵⁰	Use of Grid computing software-network for regression of GC models Workflows and services for deployment and automated management of CAMPD and CAMPCD tools in massively distributed and geographically dispersed Grids and Clouds	Advanced computing infrastructures Advanced computing infrastructures and services
Gebreslassie and Diwekar ²⁸⁵	Multiagent computing framework for autonomous management of cooperative, parallel, optimization-based CAMD	Advanced optimization and decision-making architectures for parallel environments

may be biased due to uncertainty in the predictions. The approach builds on the work of Papadopoulos et al.³⁹⁵ for the systematic selection of CO₂ capture solvents. Zargiannis et al.¹⁶⁵ use the same approach for the selection of mixtures under property uncertainty.

p1825 Martin and Martinez³⁹⁶ introduce uncertainty in the blend design problem of Martin and Martinez³⁰⁸ and perform probability distribution sampling to propagate uncertainty through the underlying model. They consider external sampling where once a sample is generated, the stochastic optimization problem is solved as a deterministic one by considering a sample average approximation in the objective function. A systematic scenario reduction approach is used to reduce the computational effort. Internal sampling is also considered where sampling is performed internally in the algorithm which re-computes the updated sample set together with the samples generated in all previous iterations. To reduce the computational effort, they propose a criterion which considers the difference in objective functions over two consecutive iterations; the algorithm terminates if this difference is lower than a threshold value. This probably means that the addition of new samples does not change the objective function value significantly.

p1830 Ng et al.³⁹⁷ propose a CAM^UD approach where uncertainty is introduced in a fuzzy representation. Uncertainty in molecular properties is first quantified by considering the standard deviation of the average variation between the measured and estimated values in regression analysis for the derivation of the property prediction model. The standard deviation is used to extend the range of the desired upper and lower bounds of the considered properties. These bounds are then used as part of the approach proposed by Ng et al.³⁵⁹ and the problem is solved using the Max–Min and two-phase approaches proposed by these authors. A similar method is further used by Ten et al.³⁹⁸ for CAM^UD which includes indices for safety and occupational health. The authors use the score approach of Ten et al.³⁶⁴ (see “Multi-Objective Optimization” section), where uncertainty is implemented by considering the standard deviation of the property prediction models. The overall problem is formulated through a fuzzy representation based on Ten et al.,³⁶⁶ while the appearance of discontinuous functions is treated through disjunctive programming.

p1835 Frutiger et al.³⁹⁹ propose a systematic approach for molecular selection based on process performance under uncertainty in the molecular properties. The method is developed and applied for ORC working fluids and processes. The proposed approach first identifies the optimum process parameters for a large data set of working fluids. Instead of using optimization to identify the optimum process parameters, all working fluids are tested in terms of their process performance (net power output of ORC) within a uniformly sampled space of process parameters (turbine inlet pressure and temperature). The combination of process parameters that gives the higher net power output is selected as the optimum design point for every working fluid. For this design point, the authors employ a Monte Carlo procedure for uncertainty analysis. Fluid property uncertainty is specified based on available data and the Latin hypercube sampling (LHS) method⁴⁰⁰ is utilized for probabilistic sampling of the fluid property input space of each working fluid. The generated samples are then propagated through the ORC model to calculate the net power output (performance criterion). Confidence intervals are then calculated for the net power output with respect to the corresponding input property value and the working fluids are ranked considering the uncertainty.

p1840 Frutiger³⁴² further proposes a reverse engineering approach for CAM^UPD of working fluids for heat pumps, following a continuous representation. A set of molecular property descriptors is used as the set of continuous design variables in order to identify working fluids. This set contains all the necessary descriptors (i.e., critical temperature and pressure, acentric factor, molecular weight, and heat capacity correlation parameters) to calculate the desired process properties through a cubic EoS. The working fluid design space specified by combinations of these descriptors is systematically sampled using a Monte Carlo approach. The generated samples represent a virtual working fluid as they do not necessarily correspond to an existing molecular structure or one that can be predicted by GC methods. Each virtual working fluid is introduced into process simulation and they are all ranked based on a process performance index. The derivative-based global sensitivity analysis approach of Kucherenko et al.⁴⁰¹ is then used to identify the descriptor that has the highest sensitivity with respect to the process performance index and this information is represented through a weight. Next, the virtual fluids are compared with real fluids obtained from a database and a distance function is developed to determine the real fluid that lies closest to each virtual fluid. The distance function accounts for uncertainty in properties, hence it is calculated based on the distance of each virtual fluid property from the upper and lower bounds of the real fluid property. These bounds are calculated from the confidence intervals provided for each fluid property in the database where they were retrieved from. The previously calculated weights are also used in the distance function to ensure that properties with

higher sensitivity toward the process performance measure have a higher impact on the identification of matching real fluids. After the identification of the real fluids, the approach of Frutiger et al.³⁹⁹ is used for selection of the best fluid-process scheme under uncertainty.

p1845 Frutiger et al.⁴⁰² further propose different ways for combining CAMPD with uncertainty in molecular properties. A first approach includes the utilization of GC-based CAMPD and the application of the methodology of Frutiger et al.³⁹⁹ for the evaluation of uncertainty. A second approach includes the incorporation of uncertainty within CAMPD. The authors quantify the uncertainty of GC parameters through the systematic methodology of Frutiger et al.⁴⁰³. A Monte Carlo approach is then used to generate samples which represent scenarios implemented as constraints in the CAMPD problem. The latter is then solved for each scenario using an optimization approach. Fluids that appear most often in all scenarios are considered as the desired options. The third approach is a summary of work presented in Frutiger.³⁴²

p1850 Santos-Rodriguez et al.³¹³ address the design of working fluids for ORC under uncertainty considering either directly the stochastic problem or the perfect information problem. The former is a CAM^uPD formulation where one optimum design is identified considering all potential uncertain scenarios during optimization. The latter includes the solution of CAMPD problems for each identified scenario of the uncertain parameters and the identification of one optimum design for each problem.

p1855 Andres-Martinez et al.⁴⁰⁴ propose two approaches of investigating property uncertainty in CAMD of working fluids for ORC. The first approach involves solution of the CAMD problem and then use of a Monte Carlo methodology at the optimal solutions for quantification of property uncertainty and propagation through the property prediction models used as objective functions in CAMD for every identified molecule. The Monte Carlo methodology includes the assignment on GC parameters of probability distributions, sampling through LHS, and calculation of the properties through GC models. The second approach includes direct solution of the CAM^uD problem, cast as an MINLP with uncertainty considered in the linear constraints. Uncertainty is introduced as random, symmetrical perturbations around the nominal property values.

s0295 **Molecular, Process, and Control Design**

p1860 Jakslund and Gani³⁰³ formulate a CAMPCD problem where the identification of an optimum molecule is also part of integrated process and control design. The authors acknowledge that molecular properties affect the process flow sheet design characteristics under nominal and variable operation. The presented case studies include either integrated molecular and process design or integrated process and control design problems, hence the CAMPCD formulation is not solved.

p1865 Ulas and Diwekar³⁴⁴ propose a multistage CAMPCD framework focusing on batch separation. Firstly, candidate solvents are selected based on CAMD, whereas the optimal batch column configuration is then selected based on optimization and heuristics using MOO. Finally, the optimal operation policy is found for the best column configuration. The authors consider the static thermodynamic uncertainties at the CAMD solvent stage and dynamic uncertainties at the process design stage to address their effects on operation. The approach is exemplified through a case study on acetonitrile–water separation where two solvents and a batch column with a middle vessel result from the optimum solvent and process design stage. For this column, the authors then identify an optimum control policy for the reboiler to increase the product recovery.

p1870 Papadopoulos and Seferlis¹⁶⁹ propose the assessment of solvents resulting from a CAMPD approach based on their nominal economic performance and their performance under the influence of disturbances. In a MOO sense, they show that solvent and process schemes which are optimum under nominal operation may exhibit low economic performance at conditions different than the nominal design settings (see also discussion in “Computer-Aided Molecular, Process and Control Design” section).

p1875 Papadopoulos et al.¹⁶² present a generic CAMPCD framework, with the main stages shown in Fig. 9. The first stage is based on the work of Papadopoulos and Linke^{161,163,321} where solvents are designed in MOO CAMD. The resulting Pareto set of solvents is introduced into process flow sheet design in order to obtain a first assessment of nominal economic process performance. This stage serves to identify optimum solvents based on process performance measures, providing also insights of potentially useful process features. Solvent–process schemes of high performance are then selected, and the solvents are introduced into process design using higher-fidelity process models. This stage serves to verify the high economic performance of the selected solvents and to assess their ability to reach desired operating set-points under nominal operation. At this point, it is clear that solvents unable to reach set-points under nominal operation will also be unable to reach them in view of disturbances, hence they may be eliminated. The selected solvent–process schemes are then introduced into a stage where their performance is evaluated under disturbances, whereas control structures are proposed that ensure efficient compensation for their effects. An optimum control problem is formulated where the objective function penalizes deviations from the optimum set-points of the controlled variables and the use of resources (manipulated variables) to bring the system back to its optimum operation (set-points). The problem is approached through a nonlinear sensitivity analysis method which evaluates the impact of multiple disturbances on process performance considering simultaneously multiple controlled parameters and resources (manipulated variables) needed to restore optimum operation. The proposed approach avoids the use of dynamic process models. The nonlinear sensitivity analysis enables the identification of undesired cases where small in magnitude disturbances require large variations in the steady state of the manipulated variables. Such behavior indicates limited ability by the solvent/process design configuration to address the disturbances and implies a compromise in the achieved dynamic performance by the control system.

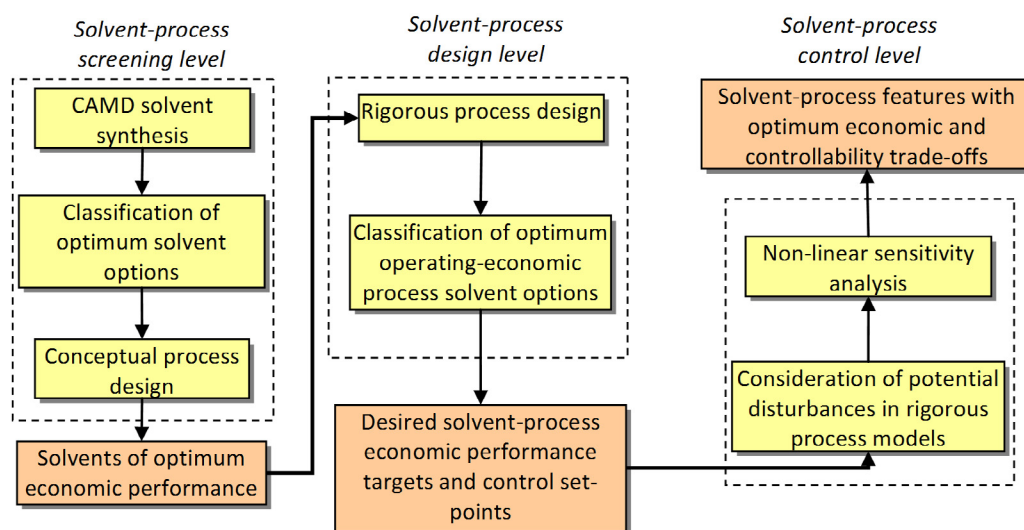


Fig. 9 Main stages of the CAMPCD approach of Papadopoulos et al.¹⁶² Reproduced from Papadopoulos, A.I.; Seferlis, P.; Linke, P., *Chem. Eng. Sci.* **2017**, *159*, 154–176 by permission of Elsevier.

s0300 Exploitation of Advanced Computing Infrastructures

p1880 The previous sections highlight very significant technological developments toward the solution of problems which include high combinatorial complexity as well as molecular and process models of high fidelity. All these developments may also be further supported by the exploitation of advanced computing infrastructures (Table 15).

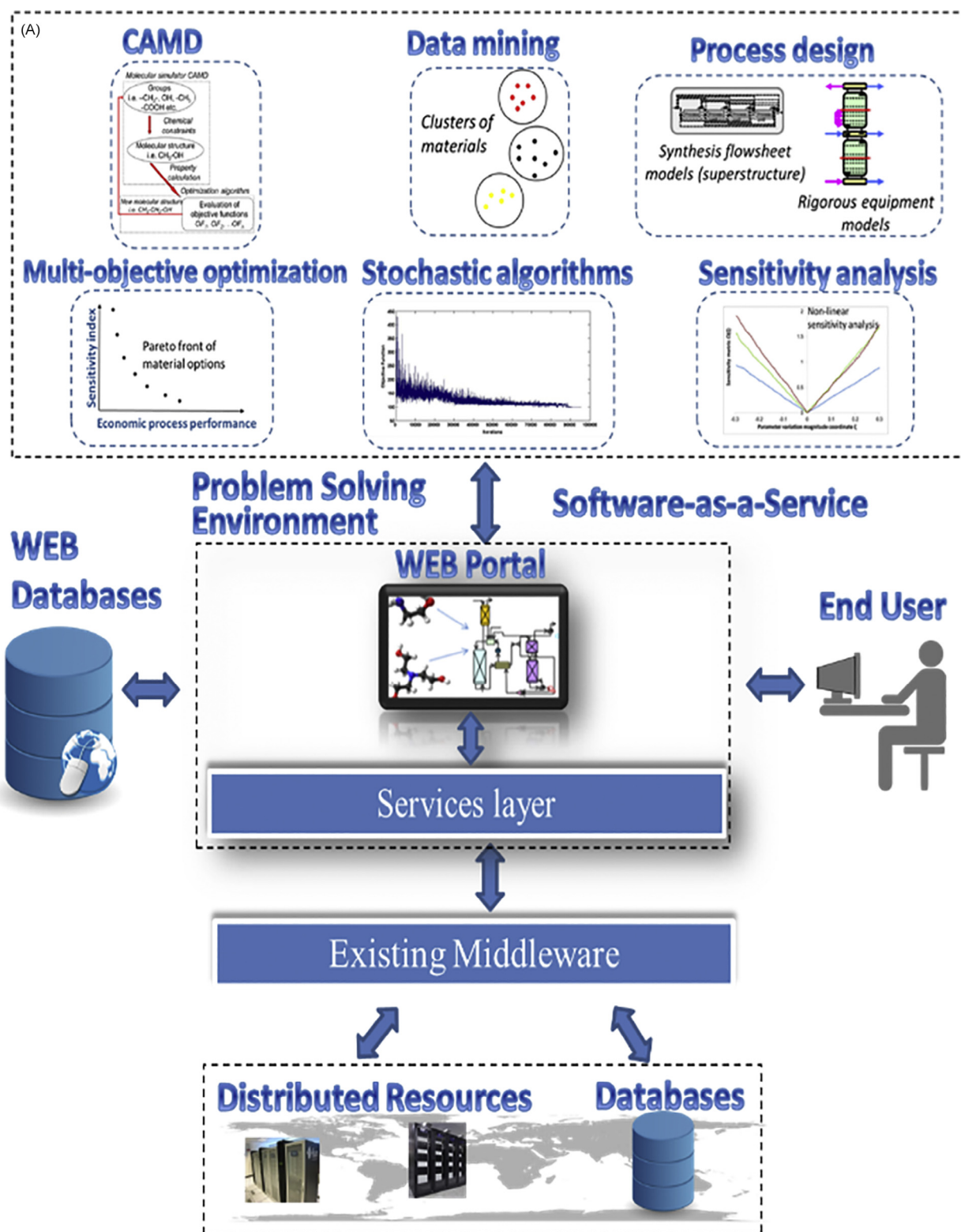
p1885 Satyanarayana et al.^{405,406} propose the use of Grid computing software for the parallelization of calculations in a computer network to regress GC models for the prediction of polymer properties.

p1890 Papadopoulos and Linke³²⁶ and Papadopoulos et al.¹⁵⁰ propose an architecture for the efficient deployment of CAMPD and CAMPCD tools and workflows in massive computing environments such as Grids and Clouds to address the high computational effort. Main components of the architecture are illustrated in Fig. 10. The authors develop a web portal to render different tools used in CAMPD and CAMPCD available to end users. The portal further serves as a connecting node between remote, geographically distributed computing resources and external databases available through web channels. Such resources involve exploitation of computer networks for distribution of the design calculations and available data repositories for storage of design-related information. The utilized computing networks include over 3000 processing units, while the data repositories have a storage capacity of over 100 TB. The geographically distributed computing resources are managed by middleware available through a third-party organization, which provides generic functionalities, is hard to use, and lacks the automation levels matching particular requirements of a problem such as CAMPD. The authors develop services that allow user-friendly access of the service functions available at the third-party middleware through the web portal. A major service includes the workflow generator, which allows the user to select from a number of developed molecular and process design workflows. The job scheduling service exploits a collection of services available at the third-party middleware to launch the workflow in the distributed resources for execution, while it dynamically manages the selection of appropriate execution resources, the storage of generated data, and the termination of the workflow execution. Data management and resource allocation services are also available independently of the job scheduling services that allow treatment of the generated data (i.e., storage, retrieval, etc.) in the available storage resources as well as querying of the status of the available storage and computing resources. Furthermore, appropriate service functions are developed that allow querying and retrieval of results in third-party web-based databases through the web portal.

p1895 Gebreslassie and Diwekar²⁸⁵ propose a homogeneous multiagent optimization framework which is applied in CAMD. The framework employs agents which include an ACO algorithm for design within different areas of the design space and from different starting points. The agents communicate solutions with each other through a global information sharing environment. The communicated solutions are exploited by agents to improve their optimization search and to avoid local optima. Each agent establishes a communication protocol with the global information sharing environment which includes all the operations for solution selection and communication, as well as scheduling and execution of the agents. The latter are implemented in parallel computing environments through services available in commercial software.

s0305 Methodologies for Reactive Systems or QM-Based Approaches

p1900 This section presents an overview of methodologies addressing reactive systems or employing QM models for reactive or nonreactive processes. See “Molecular Design” section discusses contributions for molecular design (Table 16), whereas see



10055 **Fig. 10** (A) Grid and cloud computing architecture for CAMPD and CAMPCD, (b) details of services.^{150,326} Reproduced from Papadopoulos, A.I.; Linke, P.; Seferlis, P., Integrated Multiobjective Molecular and Process Design: Operational and Computational Frontiers. In *Tools for Chemical Product Design: From Consumer Products to Biomedicine*. Computer Aided Chemical Engineering. Martin, M., Eden, M. and Chemmangattualappil, N. Eds.; Elsevier: Amsterdam, **2016**, Vol. 39, Chapter 11, pp. 269–313 by permission of Elsevier.

(Continued)

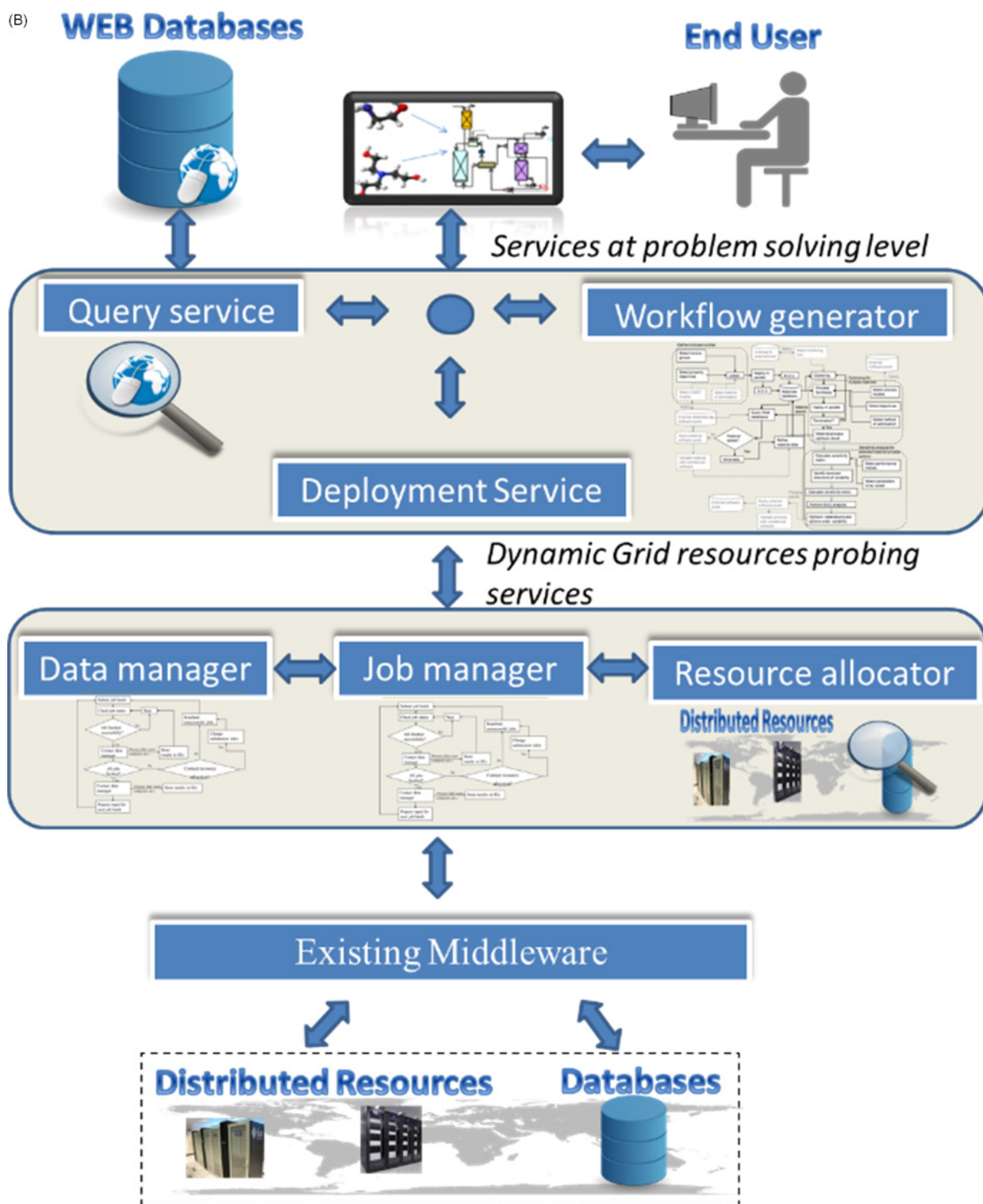


Fig. 10 (Continued)

10085 **Table 16** Overview of approaches for CAMD in reactive systems or using QM models

References	Key features
<i>CAMD in reactive systems</i>	
Buxton et al. ^{407,408}	Reaction path synthesis approach where CAMD is used for the identification of potential raw materials and stoichiometric coproducts based on the desired product characteristics
Gani et al. ^{409,411} , Folic et al. ⁴¹⁰	Knowledge-based approach for the identification of solvents that promote chemical reactions, allows the incorporation of reaction rate models or data, if available
Folic et al. ^{389,391}	Regression of solvatochromic equation coefficients for reaction rate constant prediction, investigation of deterministic and stochastic CAM RD and CAM ^{RPD} cases, global sensitivity assessment of coefficients prior to stochastic case implementation, solvolysis reaction
Papadopoulos et al. ³⁹⁴	CAMD for CO ₂ capture using chemical absorption where CO ₂ solvent solubility is predicted rigorously using SAFT-based models that simultaneously account for phase and chemical equilibrium, without the need to postulate reaction products or mechanisms
Chemmangattuvalappil and Eden, ²⁹⁷ Dev et al., ^{435,436} Dev ³⁶⁹	Systematic approach to represent molecular structures through signature descriptors used in CAMD to track changes in structures due to reactions, extension to account for nonlinear structure–property models and to modeling of the reaction rate constant in Diels–Alder reaction
Gerbaud et al. ⁴¹³	Combination of CAOS ⁴³⁷ approach with CAMD to identify molecules originating from biorenewable feedstocks, to propose sustainable synthesis pathways, and to improve their properties
<i>QM models in CAMD of reactive systems</i>	
Stanescu and Achenie, ⁴¹⁴ Stanescu et al. ⁴¹⁵	QM models used after GC-based CAMD have been completed, for the evaluation of selected solvents, DFT calculations
Hechinger et al., ⁴¹⁶ Dahmen and Marquardt ⁴¹⁸	Combination of CAMD with RNFA for the identification of fuel and blend candidates and the identification of biorenewable reaction pathways for their production, MOO, blend design, DFT calculations
Struebing et al. ^{155,419}	CAM RD is performed using a surrogate model for the reaction rate constant, the model is updated based on QM calculations for the optimum solvent found from the solution of CAM RD , DFT calculations, continuum solvation model
Sioukrou ¹⁵⁹	CAM RD is performed using QM calculations directly in MINLP, without a surrogate model, assumptions are made regarding fixed geometry from gas to liquid phase to facilitate computations. The assumptions are lifted by proposing an approach similar to Struebing et al., ¹⁵⁵ but with a Kriging approach for surrogate model development instead of the solvatochromic equation, DFT calculations
Zhou et al. ³⁹²	Derivation of surrogate model using QM prior to CAM RD and the use of the surrogate to directly solve an MINLP, testing of a deterministic and a stochastic problem, DFT calculations, COSMO model
Austin et al. ¹⁴⁹	QM-CAM ^{RPD} approach using COSMO-RS and COSMO-SAC, GC developed from a public database of σ -profiles, direct optimization combining DFO and NLP
De Vleeschouwer et al. ^{423,425}	Inverse, optimization-based evaluation of 2-naphthol derivatives using QM models based on acidity constant calculations, DFT calculations
<i>QM models in CAMD of nonreactive systems</i>	
Harper et al. ¹⁸⁹	Multilevel approach with successive use of increasingly detailed models to gradually reduce the number of potential solvent candidates, use of QM software proposed to evaluate few promising selected solvents
Lehmann and Maranas ¹⁷²	Decomposition approach, selected solvents satisfying criteria are introduced in QM model calculations in the course of CAMD, DFT calculations, continuum solvation model
Sheldon et al. ²³⁴	QM in the course of CAMD, with GC property values as input, calculation of free energy of solvation, direct solution of MINLP, continuum solvation model
Satyanarayana et al. ⁴²⁶	Polymer repeat units identified through CAMD and selected units are evaluated using molecular dynamic simulations
Hada et al. ²⁹⁹	QM-CAMD of ionic liquids using DFT to derive infrared spectra for a training set, results used for development of latent variable property prediction models, CAMD performed using property clustering technique
Farahipour et al. ⁴²⁷	Screening of ionic liquids using COSMO-RS for calculation properties related to their performance as CO ₂ capture solvents, gradual reduction of initial set based on satisfaction of constraints
Peng et al. ⁴²⁸	QM-CAMD of ionic liquids using GC-COSMO and COSMO-SAC, derivation of σ -profiles and cavity volumes for a data set before CAMD, DFT calculations
Zhang et al. ⁴²⁹	Approach similar to Peng et al. ⁴²⁸ , use of the GC method of Zhang et al. ⁴³⁰ and of a mixed SA-GA algorithm for CAMD of ionic liquids
Scheffczyk et al. ²⁵⁹	QM-CAMD using COSMO-RS in the course of molecular design with online calculation of σ -surfaces (at BP-TZVP-MF QM level), selected molecules are reevaluated at a more accurate BP-TZVPD-FINE QM level in the refinement phase
Zhao et al. ⁴³²	Screening of large database of ionic liquids, calculation of thermodynamic properties through GC or COCMO-RS, gradual elimination of candidates based on thermodynamic and environmental and health properties
Ahmad et al. ⁴³³	Design of postcombustion CO ₂ capture solvents using GC-CAMD and subsequent estimation of heats of formation for selected solvents using QM models, considering two main reaction mechanisms to derive stoichiometry, DFT calculations

t0090 **Table 17** Overview of molecular and process design approaches considering reactive processes and/or QM models

References	Key features
Folic et al. ³⁹¹	CAM ^R PD using solvatochromic model for reaction rate constant prediction, regression of rate constant model, GC model for rate constant prediction, simple CSTR models, deterministic and stochastic CAM ^R PD evaluation, direct optimization, Menschutkin reaction
Mac Dowell et al., ³⁴³ Pereira et al. ^{152,438}	CAMPD for CO ₂ capture using chemical absorption where CO ₂ solvent solubility is predicted rigorously using SAFT-based models that simultaneously account for phase and chemical equilibrium, without the need to postulate reaction products or mechanisms
Salazar et al. ⁴³⁹	Molecular design, use of eNRTL and UNIFAC at process level, chemical CO ₂ capture solvents and processes, decomposed approach
Sioukrou et al. ³⁷⁷	Design of GXL blends, identification of optimum blend concentration and CO ₂ cosolvent from prespecified set of options, process flow sheet comprising CSTR, separator and compressor, Diels-Alder reaction, process cost and solvent inventory optimization, direct approach
Papadokostantakis et al. ³²⁵	Chemisorption solvent and process design for CO ₂ capture, thermodynamic, reactivity and sustainability criteria at solvent design stage, economic and sustainability criteria at process design stage, rigorous equilibrium column models, NLP process optimization considering different flow sheet structures and operating conditions, SAFT-VR thermodynamic model, decomposed approach
Zhou et al. ⁴⁴²	QM-CAM ^R PD using COSMO model, reaction rate constant prediction from linear model regressed based on σ -profiles, COSMO model calculated once for set of solvents used to derive reaction rate constant model, reactor, distillation column and heat exchanger flow sheet model, maximization of profit, Diels-Alder reaction, direct MINLP formulation
Liu et al. ⁴⁴³ Song et al. ⁴⁴⁴	Screening of ionic liquids with properties predicted through COSMO-RS, evaluation of their performance in process simulations Ionic liquid structure enumeration and extractive desulfurization process simulation, gradual reduction of ionic liquid options based on constraint satisfaction, use of UNIFAC and COSMO-RS for calculation of missing parameters and results validation
Scheffczyk et al. ⁴³¹	QM-CAMPD using COSMO-RS in the course of molecular and process design using a GA, design phase includes TZVP-MF implementation of COSMO-RS and pinch-based process separation process model with infinite number of stages, refinement stage includes TZPVD-FINE COSMO-RS with pinch-based process model for selected solvents first and then use of optimum results in ASPEN Plus [®] stage-by-stage rigorous simulations, decomposed approach

“Molecular and Process Design” section (Table 17) discusses contributions for molecular and process design. With respect to reactive systems and processes, this section includes contributions where the designed molecule may promote the chemical reaction by directly affecting kinetics (e.g., reaction rate) or participates in the chemical reaction as a reactant or as a product. Contributions where a molecule participates in reactive processes as a mass separating agent are not considered here.

s0310 Molecular Design

s0315 CAMD in reactive systems

p1905 In the area of reaction path synthesis, Buxton et al.^{407,408} propose a systematic approach for reaction path synthesis which enables the identification of potential alternative routes for the production of a desired product and the selection of the optimal route. Central to the proposed methodology is the use of CAMD for the generation of alternative candidate comaterials (raw materials and stoichiometric coproducts) based on the structure of the product as well as desired chemical and property constraints. After determining the comaterials, the methodology involves the identification of potential stoichiometries using a knowledge-based approach combined with an optimization selection procedure, the generation of corresponding reaction mechanisms, and the evaluation of the mechanism steps.

p1910 Gani et al.⁴⁰⁹ propose a systematic, knowledge-based approach for the screening of solvents as reaction media. The methodology considers and analyzes various different reactions and classes and it allows the incorporation of reaction rate models or data. It includes a scoring system to rank-order solvents, which is based on properties such as reactivity, phase-split formation, solubility toward desired solutes, heating/cooling, association/dissociation, selectivity, etc. The work is further extended to multistep^{410,411} and replacement reaction systems.⁴¹¹ The proposed steps and rules as well as the scoring system are implemented to all the reaction steps where solvents are identified and successively eliminated, resulting in few options worth of further analysis and experimentation.

p1915 Folic et al.³⁸⁹ propose a CAM^RD approach where parameters for the solvatochromic equation are regressed to predict the reaction rate constant. The equation is then used in the design of solvents, considering both deterministic and stochastic scenarios. Prior to the stochastic application, the authors perform global sensitivity analysis to evaluate the effect of uncertain coefficient ranges on the model and to identify representative combinations of the coefficients. The method is further used by Folic et al.³⁹¹ in CAM^RPD discussed in “Molecular and Process Design” section.

p1920 Papadopoulos et al.³⁹⁴ propose the design of solvents for CO₂ capture through chemical absorption. The CAMD stage includes calculation of CO₂ solubility in solvents using the Hansen solubility parameter⁴¹² and of solvent reactivity based on basicity constant, among other properties used as design criteria. Selected solvents are then evaluated rigorously with respect to CO₂ solubility based on vapor–liquid equilibrium predicted from a GC version of SAFT- γ -SW^{134, 135} EoS. This EoS accounts implicitly for both phase and chemical equilibrium without the need to postulate reaction products or mechanisms.

p1925 Chemmangattuvalappil and Eden²⁹⁷ propose a systematic approach to represent molecular structures as signature descriptors and use it in CAMD to track changes in structures due to reactions. The underlying idea is that the changes in the chemical structure can be correlated with the changes in the properties of the molecule. Therefore, the changes in the molecular structure due to reactions can be represented as a function of the property. Dev³⁶⁹ extends this approach for design of reactants and products irrespective of their numbers, while also using nonlinear models for structure–property predictions. Furthermore, Dev³⁶⁹ evaluates different machine learning algorithms with respect to modeling the reaction rate constant of a Diels-Alder reaction based on reactants.

p1930 Gerbaud et al.⁴¹³ address the design of molecules originating from renewable feedstocks and with an eco-friendly synthesis pathway. A computer-aided organic synthesis (CAOS) approach is used which takes as input a bio-sourced building block, a list of readily available coreactants and a list of eco-friendly chemical transformations suitable for large-scale production. The approach then enumerates all possible products, resulting in thousands of structures. These molecules exhibit an overall performance that is usually not satisfactory. Then, the CAMD approach of Heintz et al.¹⁵³ is used to enhance the properties by proposing derivatives thanks to additional substitutes. The work is part of a broader product design framework which also considers (a) an intelligence phase prior to molecular design including strategic, tactical, and operational decisions regarding the products that should be developed and (b) a choice phase where experimental work is performed after molecular design for selected candidate molecules.

s0320 **QM models in reactive systems**

p1935 Stanescu and Achenie⁴¹⁴ and Stanescu et al.⁴¹⁵ propose a QM-CAMRD approach for the identification of solvents that create a homogeneous reaction mixture, while at the same time promote the Kolbe–Schmitt reaction. DFT calculations are used to confirm the reaction mechanism, and to compute the energies and thermodynamic properties of the optimized structures and transition states along the reaction path. Solvents are designed using GC-CAMD and few of them are selected for calculation of the reaction rate constant using the transition-state theory. The QM calculations are therefore used as a final step, after the solvents have been designed.

p1940 In a different line of work, Hechinger et al.⁴¹⁶ propose the combination of reaction network flux analysis (RNFA)⁴¹⁷ with CAMD for the identification novel biofuels and chemical pathways of producing them. CAMD is used in a strategy where optimum molecules are identified based on properties which point toward useful fuels. Then, a reaction network toward these components is constructed in order to identify and compare suitable production routes. A different strategy is also considered where fuel-related properties are calculated for all molecules in a prespecified reaction network and are then employed as performance criteria in order to identify target molecules and the corresponding synthesis pathways using RNFA. In this case, QM models such as DFT are used for the calculation of properties. The latter include the standard enthalpies of formation and combustion, the normal boiling point, and the oxygen content. The work is extended by Dahmen and Marquardt⁴¹⁸ in a four-stage approach for the design of biofuel blends. In stage 1, palette compounds are identified using CAMD, followed by the development of a conversion pathway map in stage 2. Stage 3 involves the blend design problem formulation which includes process-related constraints such as production yield as well as pathway models. Stage 4 includes the design and analysis of the blends using MOO.

p1945 Struebing et al.¹⁵⁵ propose a QM-CAMRD approach that employs a surrogate model in the course of CAMD to reduce the expensive QM calculations, which are performed once in each iteration after the solution of the optimum CAMD problem. The methodology is illustrated in Fig. 11, obtained by Struebing et al.⁴¹⁹ who present a more extensive implementation of the approach. An initial set of solvents is used to derive a surrogate model, namely the solvatochromic equation, for the prediction of the reaction rate constant. CAMD is then performed using the surrogate model and the optimum solvent is evaluated with the QM model. The results are used to update the initial set of solvents and the surrogate model which provides increasingly reliable predictions as more optimum solvents are included in the regression set at each iteration. The iterations are repeated until no new optimum solvent is found. The method enables the evaluation of 1300 solvents for the Menschutkin reaction with only 10 calls of the QM models. The latter include DFT with SMD continuum solvation model.⁴²⁰ Struebing et al.⁴¹⁹ extend the approach to take into account the impact of solid reactant solubility and solvent density on solvent choice.

p1950 Sioumkrou¹⁵⁹ proposes a QM-CAMRD approach where the reaction rate constant of each designed solvent is calculated directly through QM models. To facilitate computations, fixed molecular geometry is considered when moving from the gas to the liquid phase. More details regarding the resulting MINLP formulation are reported in “Deterministic optimization approaches in molecular design” section. As noted in Sioumkrou,¹⁵⁹ this assumption introduces uncertainty in the design. To address this uncertainty, the author adopts the approach of Struebing et al.,¹⁵⁵ but the employed surrogate for reaction rate constant prediction is now a Kriging²³⁶ model. Kriging interpolation models have advantages compared to conventional linear regression. In the latter approach, predictions are provided based on a model (e.g., polynomial) that is postulated in advance. The model parameters are selected to optimize some criterion of best fit at the data points. Kriging produces an interpolation function based on a covariance model derived from the data rather than an a priori model of the interpolating function.⁴²¹ Kriging further accounts for uncertainty in modeling and provides a measure of confidence in the predictions. The use of this approach by Sioumkrou¹⁵⁹ results in a solvent which exhibits a 293% increase in the reaction rate constant compared to using the solvatochromic equation as the surrogate model for the Menschutkin reaction. The method is also applied in a Cope elimination reaction with a similarly high increase in the rate constant, compared to the best solvent of the initial set.

p1955 Zhou et al.³⁹² propose a QM-CAMRD approach investigating both deterministic and stochastic solvent design (CAM^{uRD}). The authors derive a surrogate model first and then use it in molecular design. The treatment of uncertainty is reported in “Molecular Design Under Uncertainty” section.

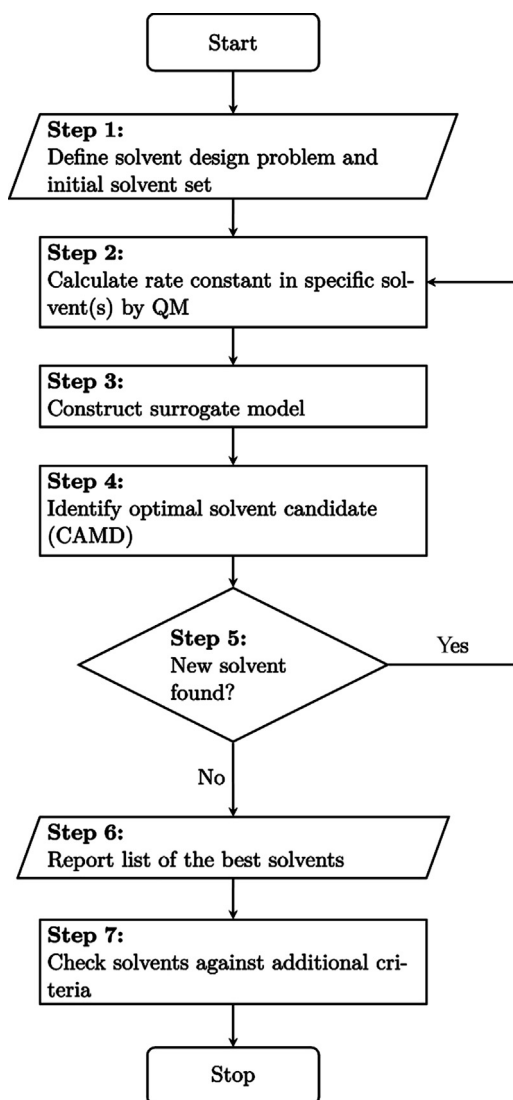


Fig. 11 QM-CAM^{bD} approach of Struebing et al.⁴¹⁹ Reproduced from Struebing, H.; Obermeier, S.; Sioumkrou, E.; Adjiman, C.S., *Chem. Eng. Sci.* **2017**, *159*, 69–83 by permission of Elsevier.

Austin et al.¹⁴⁹ propose a generic QM-CAM^{brD} approach exploiting models such as COSMO-RS and COSMO-SAC,⁴²² which is illustrated for both reactive and nonreactive systems. The approach is based on the CAM^{bD} framework of Austin et al.¹⁴⁹ which is described in “Mixture and Blend Design” section. In the QM-CAM^{brD} approach, the necessary component and mixture properties are calculated through σ -moments, σ -profiles, and cavity volumes of every species in solution from specifically developed GC models. The GC models are derived from data already available in a database, so there is no need for DFT or other QM calculations at any stage. The approach is illustrated for a liquid–liquid extraction solvent design case and for the design of solvents needed in the Menshutkin reaction. In the latter case, the authors address the design of both a single solvent and a mixed solvent. In both cases, the authors report significant performance improvements compared to solvents from the literature.

A non-GC, inverse engineering approach which applies optimization simultaneously with QM models for the calculation of molecular properties related to reactivity is proposed by De Vleeschouwer et al.⁴²³ The approach uses different substituents to generate 2-naphthol derivatives which are evaluated with respect to their acidity when excited from the singlet ground state to the first singlet excited state. The change in acidity constant exhibits a difference of seven orders of magnitude. For a set of 10 possible substituents on each of seven naphthol sites, the design space is in the size of 10^7 structures. DFT calculations are used to determine the acidity constant, while the optimization approach used is the Best First Search,⁴²⁴ a BB-type algorithm. Variations of this algorithm are also proposed in De Vleeschouwer et al.⁴²⁵ for inverse engineering of molecules using QM models.

QM models in nonreactive systems

Harper et al.¹⁸⁹ propose a systematic methodology (reviewed in “Generate and Test Approaches” section), where after three CAMD-based screening levels including improved property prediction models, the remaining molecules are introduced into QM software.

The latter is used to evaluate energy minima of different conformers and to predict desired properties. Applications are reported for nonreactive separation solvent design.

p1975 Lehmann and Maranas¹⁷² propose a QM-CAMD approach which uses a GA and combined GC-CAMD with ab initio QM models in a three-step approach. The first step includes the GC-CAMD of molecules using MILP or complete enumeration. The second step includes GA-based GC-CAMD in order to fine-tune the GA control parameters based on the results of the first step, so that it finds the maximal average number of globally best solutions. The fine-tuned GA is used in the third step which includes the actual QM-CAMD approach, where QM models are used in the course of CAMD. Molecules are generated by the GA and their properties are evaluated based on GC. Those that are predicted to perform close to a desired target performance are introduced into QM calculations. This reduces the number of expensive ab initio calculations, together with a database that registers previously calculated structures so that they are not re-calculated. QM models include DFT calculations and an internal geometry optimization step which finds a local energy optimum. The desired properties are then evaluated for the converged geometries. Continuum solvation QM models are also used in one case. Applications involve refrigerant and solvent design.

p1980 Sheldon et al.²³⁴ also use QM models in the course of CAMD. A continuum model of solvation is used, employing GC property values as input in order to calculate the free energy of solvation, which is the objective function in optimization. Details regarding the optimization approach are reported in "Deterministic optimization approaches in molecular design" section.

p1985 Satyanarayana et al.⁴²⁶ use the approach of Harper and Gani¹⁹⁰ to first design polymer repeat unit structures satisfying a subset of (macroscale) properties. Then, a polymer repeat unit is selected for further analysis with respect to properties at the microscale using molecular dynamic simulations.

p1990 Hada et al.²⁹⁹ propose the use of QM models as a first step toward the design of ionic liquids using a CAMD approach based on property clustering. DFT calculations are used to derive the necessary infrared spectra descriptors (based on a training set of ionic liquids) in order to develop latent variable property models the prediction of viscosity, melting point, and density. These models are used in order to develop property targets which are matched by ionic liquids generate using a property clustering technique.

p1995 Farahipour et al.⁴²⁷ propose a systematic screening approach for the identification of ionic liquids as CO₂ capture options. Initially, cation options are screened based on viscosity and the ones with low viscosity are further matched with an arbitrary anion and screened based on the absorption-desorption index (ADI) which accounts for Henry constant. The top 10 are selected and all combinations are evaluated in terms of ADI. Anions contributing to low ADI are selected and the ones contributing to low viscosity are further screened in their combinations with selected cations and are further evaluated in terms of melting point. The ones satisfying the latter constraint are selected based on their ADI values.

p2000 The GC-COSMO of Mu et al.¹⁴⁶ is extended by Peng et al.⁴²⁸ for the design of ionic liquids. The GC is derived through DFT calculations (optimization to the lowest energy in the ideal gas phase) for a dataset of 828 cations and 61 anions. The GC is introduced into COSMO-SAC for the mixture property calculations required in the applications of liquid-liquid extraction and postcombustion CO₂ capture. A similar approach is used by Zhang et al.⁴²⁹ for the design of ionic liquids, but the GC method is based on the work of Zhang et al.,⁴³⁰ who propose a new approach of predicting the surface charge density of ionic liquids. Furthermore, the CAMD optimization algorithm is a mixed SA-GA approach. Scheffczyk et al.²⁵⁹ propose a QM-CAMD approach based on COSMO-RS. The approach is detailed in "Molecular and Process Design" section as part of the QM-CAMPD framework proposed by Scheffczyk et al.⁴³¹

p2005 Zhao et al.⁴³² propose a systematic screening approach of a large database of ionic liquids. The first step involves the identification of the gas separation goals to help determine the screening objectives. The second step involves the refinement of Henry law constant using COSMO-RS and calculation of properties such as selectivity and absorption separation index. The resulting pool of ionic liquids is further evaluated in terms of toxicity, water/octanol partition coefficient, viscosity, and melting point, based on existing data, hence forming a new pool of candidates. In the fourth step, the viscosity, melting point, and ADI are calculated using QSPR, GC, and COSMO-RS in order to select the final candidates.

p2010 Ahmad et al.⁴³³ propose a QM-CAMD approach for the design of solvents for chemical CO₂ absorption. A GC-CAMD approach is used first to identify amine solvent candidates. The reaction stoichiometry is evaluated for each one of them considering the zwitterion mechanism for primary and secondary amines and the base-catalyzed hydration mechanism for tertiary amines.⁴³⁴ The heats of formation for reactants and products are then calculated using DFT QM models, aiming to calculate the enthalpy of reaction. The authors consider linear and branched amines.

s0330 **Molecular and Process Design**

p2015 Folic et al.³⁹¹ propose a CAM^RPD approach where the solvatochromic equation and a simple reactor process model are considered to identify solvents which increase the desired product formation. The reactor model is a continuous stirred tank reactor (CSTR) and enables the consideration of competing as well as consecutive reactions. In the former case, the aim is to maximize the formation of the desired product and to suppress the formation of the undesired product. In consecutive reactions, the reaction that yields the desired product is followed by another reaction that consumes it, hence the latter must be demoted. In both cases, the process variables are the outlet concentrations of the reactant, as well as of the desired and side-reaction products. The method includes a reaction model building step where the solvatochromic equation model is regressed from an initial set of solvents. The approach further includes results for both deterministic and stochastic solvent design. In the latter case, a global sensitivity analysis approach is employed to quantify the effect of uncertainty on model reliability and to determine the optimal solvent candidate, given this uncertainty. The method is applied to a representative of the S_N² reaction class, namely the Menshutkin reaction.

- p2020 Mac Dowell et al.³⁴³ and Pereira et al.^{152,438} employ the SAFT-VR¹⁰² EoS for CAMPD of CO₂ capture processes using chemical absorption. With this EoS, the phase and chemical equilibrium are considered implicitly, whereas process models are used in the course of solvent design. Mac Dowell et al.³⁴³ design a blend and Pereira et al.^{152,438} design solvents using a continuous representation of the molecular design space (see “Continuous Molecular Representations” section for details).
- p2025 Salazar et al.⁴³⁹ propose the design of CO₂ capture solvents for chemisorption systems in a two-step approach. First, CAMD is used to design solvents considering properties such as Gibbs energy of reaction, CO₂ solubility, and boiling point as performance criteria to short-list few options, which are then introduced into process design using ASPEN Plus[®]. UNIFAC and electrolyte NRTL⁴⁴⁰ parameters are fitted for the selected solvents in order to model the equilibrium at the process level. In this case, the reaction mechanisms considered for the calculation Gibbs energy of reaction or the activity coefficient model parameters are not disclosed.
- p2030 Papadokonstantakis et al.³²⁵ propose a CAMPD approach for the design CO₂ capture solvents and processes which builds on Papadopoulos et al.³⁹⁴ (see “CAMD in reactive systems” section). Selected solvents designed through CAMD, where equilibrium, reactivity, and sustainability properties as considered as performance criteria, are introduced into process design considering various different, unconventional absorption/desorption flow sheets. Gate-to-gate sustainability assessment is considered for the investigated solvent–process systems, together with their economic performance. SAFT-VR is used to account for phase and chemical equilibria at both the solvent selection and process design stages.
- p2035 The solvatochromic equation is used by Sioungkrou et al.³⁷⁷ in the simultaneous blend design of gas-expanded liquids (GXL) and the corresponding reactive separation process in which they are used. GXL comprise a liquid solvent and a compressible gas (usually CO₂); CO₂ enhances gas solubility and mass transfer while the liquid solvent increases the solubility of liquid and solid solutes. Considering CO₂ as the compressible gas, the aim is to identify the optimum cosolvent (out of few prespecified options) and concentration in the binary mixture. In the proposed approach, the solvatochromic equation is used to identify the reaction rate constant of the mixed solvent. Solid–vapor–liquid equilibrium is calculated using the group-contribution volume-translated PR⁴⁴¹ EoS. The process consists of a CSTR, a separator, and a compressor and is implemented in a Diels-Alder-type reaction. Process cost and solvent inventory are considered as objective functions.
- p2040 Zhou et al.⁴⁴² propose a QM-CAM^RPD approach for the design of solvents and processes for a Diels-Alder-type reaction. The COSMO¹³⁶ model is used in this case to derive σ -profiles which quantify the number of surface segments that can be found within a certain screening charge density σ -interval. The areas within each interval are used as solvent theoretical descriptors to quantify their effects on chemical reactions. A linear model is regressed using these descriptors to predict the reaction rate constant using a prespecified set of solvents. The COSMO model and the required QM methods are therefore implemented only once for each solvent in the set. A GC model is then developed for the prediction of the reaction rate constant using the derived descriptors. The model is used in QM-CAM^RPD of solvents for a process comprising a reactor, a distillation column, and a heat exchanger, with recycle. The aim is to maximize profit by determining the optimum solvent molecular structure, reaction conversion, and the product recovery using a MINLP formulation.
- p2045 Liu et al.⁴⁴³ use an enumeration-based selection approach for the identification of ionic liquids used in separations. Desired, solubility-based properties are calculated using COSMO-RS. Solvents are evaluated in terms of process performance through process simulations. An enumeration-based approach is used also by Song et al.⁴⁴⁴ for the design of ionic liquids for an extractive desulfurization process. Ionic liquid structures are enumerated and gradually eliminated until few of them are selected and introduced into process simulations. UNIFAC is used for ionic liquids for equilibrium property predictions and COSMO-RS is used to calculate missing UNIFAC parameters or to validate the results.
- p2050 Scheffczyk et al.⁴³¹ propose a QM-CAMPD approach where the COSMO-RS model is used iteratively in the course of CAMPD. A GA is used for the generation of molecular structures and the operating optimization of a process. COSMO-RS is called for the calculation of the properties required to evaluate each molecular structure in terms of process operation. To avoid increasing computational effort, the approach involves a first phase where QM-CAMPD is implemented using molecular and process models of lower fidelity and a second phase where selected molecules and process operating points are refined using molecular and process models of higher fidelity. In the first phase, COSMO-RS is used at the TZVP-MF level while a COSMO database of molecules is developed in order to avoid recalculating properties of already evaluated molecules. Furthermore, a pinch-based separation process model is used which assumes an infinite number of stages and sharp component splits in order to identify lower bounds on process variables, namely minimum solvent flowrate and energy demand. The separation process temperature is a degree of freedom in this phase. The second, refinement phase employs COSMO-RS at the TZPVD-FINE level. This is initially used with the pinch-based process models which are reoptimized for the selected molecules. The obtained molecules and process conditions are introduced into ASPEN Plus[®] process models to perform the simulations necessary for the complete determination of the process operating characteristics.

s0335 Applications

- p2055 This section presents an overview of applications proposed in published literature for the previously analyzed models and methods. The presented tables include a brief description of the application together with keywords based on the problem classification and on the algorithms used to address the design problem. Advanced molecular modeling and property prediction approaches are also cited.

s0340 **Solvents for Industrial Separations, Reactive Separations, and Promotion of Reactions**

p2060 Table 18 includes applications reported for separations, reactive separations, and promotion of chemical reactions. Solvents in separation and reactive separation systems are generally used as mass separation agents, whereas in the case of promotion of reactions they directly affect the reaction kinetics. The main processes included in Table 18 are the following:

- u0090 ● Liquid–liquid extraction
- u0095 ● Gas absorption
- u0100 ● Extractive distillation
- u0105 ● Extractive fermentation
- u0110 ● Blanket wash solvents for lithographic printing
- u0115 ● Metal degreasing
- u0120 ● Crystallization
- u0125 ● Reaction processes

p2105 Most applications involve solvent design for liquid–liquid extraction, gas absorption, and extractive distillation systems. Such systems require liquid–liquid or vapor–liquid equilibrium prediction, hence activity coefficient models or EoS are commonly utilized. For liquid–liquid extraction and gas absorption, selectivity of the solvent toward the solute and solvent–solute solubility is among the most important properties. For extractive distillation, the solvent–solute relative volatility is very important. In all these processes, azeotrope formation should be avoided. Furthermore, solvents are regenerated and recycled in a separate process, hence this separation should be easy. Solvents should therefore be designed to exhibit a high boiling point difference with the solute and a low energy of vaporization, among other requirements. To avoid solvent losses, the solvent–vapor pressure should be low, while in liquid–liquid extraction the solvent should ideally be immiscible toward the raffinate to avoid liquid-phase losses.^{161,321} Extractive fermentation processes represent a case of liquid-phase reactive extraction processes where the solvent is used to remove the desired product hence enhancing the production. The requirements in solvent design are similar with liquid–liquid extraction processes, with the addition that solvent and product should not react.³²⁶

p2110 Blanket wash solvents are needed in the printing industry in order to clean ink from lithographic printing presses.²²¹ The latter are used to transfer the printed image from a plate to a rubber or a plastic blanket and then to the paper. The quality of the printed images depends on the ability to maintain a clean blanket. Blanket washes, consisting of organic solvents, are used to remove ink, paper dust, etc. from the blanket cylinders.²⁹³ Solvents should be designed to dissolve ink, to enable minimum drying time,²⁹³ while their viscosity, surface tension, density, health, and safety impact characteristics should also be within desired limits.²²¹ Their vapor pressure should also be low to avoid vapor losses.

p2115 In the case of metal degreasing processes, the flow sheet requires a fresh solvent in the degreaser which is followed by a regeneration unit.^{290,447} Solvent vapor losses from the regenerator are recovered in an absorption unit, whereas an off-gas solvent stream from the degreaser is flared. The latter contains volatile organic content (VOC) which can be condensed and reused instead of flaring. The solvent needs to have specific sulfur content to avoid corrosion, density to facilitate hydrodynamics within the processes and vapor pressure to avoid losses. The range of these properties is different in the degreaser and the absorber.

p2120 Crystallization is a process widely used for the production of high-value chemicals such as pharmaceuticals.⁴⁴⁸ The latter are retrieved from their solutions in the form of solid crystals, hence solvents play an important role. Solvent design steps require solid–liquid equilibrium calculations. The solvent–solute solubility should be high, but depending on the type of crystallization this may vary. For example, Karunanithi et al.²²⁷ note that for cooling crystallization the solvent–solute solubility should be high at higher temperature, but lower at lower temperature. This is because the solubility determines the equipment size, but the temperature determines the yield. For other types of crystallization, such requirements may vary. The crystal morphology is also affected by the solvent, hence affecting downstream processing such as filtering, washing, drying, packaging, handling, and storage. Viscosity is also important as well as solvent toxicity and flammability.

p2125 In the case of designing solvents for promotion of chemical reactions, a key requirement is the maximization of the reaction rate constant or of the reaction rate which also requires process-related calculations as it is a function of concentration. Additional issues to consider include the effects of the solvent in the catalyst solubility and activity.¹⁵⁹ A complete work that illustrates how solvents with different reaction rate constants affect process design decisions is presented in Sioungkrou et al.,³⁷⁷ where solvents are eventually selected based on process economic criteria.

s0345 **Catalysts, Adsorbents, and Ionic Liquids**

p2130 The application of CAMD in the design of catalysts is reported in publications of Table 19. The properties considered in transition metal catalyst design include electronegativity, which directly affects the catalyst activity, density, and toxicity.²⁷⁷ The oxidation state is additionally considered in a follow-up work.²⁷⁸ The presented applications focus on homogeneous catalysts containing a molybdenum center with organic substituents. The values imposed on the target properties allow the consideration of the designed catalysts for epoxidation,^{277,278} and hydroformylation reactions.²⁷⁷

p2135 Benavides et al.²⁸⁸ propose the design of inorganic materials such as adsorbents using a GC approach. The latter is developed by Benavides et al.²⁸⁹ to predict activity coefficients in the bulk phase and the adsorbate solid solution using UNIFAC. The method is

10095 **Table 18** Applications in the design of solvents for industrial separations, reactive separations, and promotion of reactions

References	Application	Problem and model class	Algorithm	Sections
<i>Liquid–liquid extraction, gas absorption, extractive distillation, and other separations</i>				
Gani and Brignole ²	Liquid–liquid extraction	CAMD	Generate and test	Generate and Test Approaches
Brignole et al. ¹⁸¹	Separation	CAMD	Generate and test	Generate and Test Approaches
Joback ¹⁸⁴	Liquid–liquid extraction of acetic acid from water	CAMD	Generate and test	Generate and Test Approaches
Macchietto et al. ¹⁷³	Liquid–liquid extraction and gas absorption	CAMD	NLP	Deterministic optimization approaches in molecular design
Gani et al. ¹⁸⁷	Liquid–liquid extraction, azeotropic, or extractive distillation	CAMD	Generate and test	Generate and Test Approaches
Naser and Fournier ¹⁷⁴	Liquid–liquid extraction	CAMD	NLP, SQP	Deterministic optimization approaches in molecular design
Klein et al., ²¹² Gani and Fredenslund ²¹⁴	Solvents	Blend design	SRLP	Deterministic optimization approaches in molecular design, Mixture and Blend Design
Odele and Macchietto ¹⁶⁰	Liquid–liquid extraction and gas-absorption	CAMD	MINLP, OA/AP	Deterministic optimization approaches in molecular design
Pretel et al. ¹⁸²	Separation	CAMD	Generate and test	Generate and Test Approaches
Jakslund and Gani ³⁰³	Separation of methyl acetate and methanol	CAMPD, CAMPCD problem formulation (no case study reported)	Generate and test	Generate and Test Approaches, Molecular, Process and Control Design
Hamad and El-Halwagi ³⁰⁴	Butane recovery in a waste-to-oil plant	CAMPD	MINLP	Deterministic Optimization Approaches
Pistikopoulos and Stefanis ³⁰⁵	Solvents for reduction of acrylonitrile and dimethylformamide using gas absorption, environmental impact minimization, acetaldehyde recovery	CAMPD	MINLP	Deterministic Optimization Approaches
Ourique and Telles ²⁷¹	Solvent for butanol recovery	CAMD, molecular graphs	SA	Simulated annealing
Harper et al. ¹⁸⁹	Solvent for recovery of phenol, extractive distillation solvent	CAMD, QM	Generate and test	Generate and Test Approaches, QM models in reactive systems
Hostrup et al. ²³¹	Separation of acetone from chloroform and of acetic acid from water	CAMPD	Generate and test, MINLP for final solvent/process options	Deterministic optimization approaches in molecular design, Generate and Test Approaches
Buxton et al. ²⁰³	Solvents for reduction of acrylonitrile and dimethylformamide using gas absorption, environmental impact minimization	CAMPD, CAM ^b D	MINLP, GBD	Deterministic optimization approaches in molecular design, Deterministic Optimization Approaches, Mixture and Blend Design
Harper and Gani ¹⁹⁰	Solvent for recovery of phenol, replacement of benzene	CAMD	Generate and test	Generate and Test Approaches
Marcoulaki and Kokossis ^{269,270}	Liquid–liquid extraction, extractive distillation and gas absorption	CAMPD	SA	Simulated annealing, Stochastic Optimization Approaches
Van Dyk and Nieuwoudt ²⁵⁴	Extractive distillation	CAMPD, CAM ^b D	GA	Genetic algorithms, Mixture and Blend Design
Kim and Diwekar ¹⁶⁷	Extraction of acetic acid from water	CAMD, uncertainty	HSTA	Molecular Design Under Uncertainty
Kim and Diwekar ²⁷³	Extraction of acetic acid from water	CAMPD, MOO, uncertainty	HSTA and NLP	Stochastic Optimization Approaches, Multi-Objective Optimization, Molecular Design Under Uncertainty
Giovanoglu et al. ²³³	Batch separation	CAMPD	MIDO	Deterministic Optimization Approaches
Cismondi and Brignole ¹⁸³	Separation	CAMD	Generate and test	Generate and Test Approaches
Kim et al. ³²⁸	Extraction of acetonitrile from waste stream, batch separation	CAMPD, MOO	–	Stochastic Optimization Approaches

(Continued)

Table 18 (Continued)

<i>References</i>	<i>Application</i>	<i>Problem and model class</i>	<i>Algorithm</i>	<i>Sections</i>
Lehmann and Maranas ¹⁷²	Liquid–liquid extraction of benzene from cyclohexane	CAMD, QM	GA	Genetic algorithms, QM models in non-reactive systems
Karunanithi et al. ²²⁶	Liquid–liquid extraction, process constraints	CAMPD, CAM ^b D	Framework for MINLP formulations	Deterministic optimization approaches in molecular design, Deterministic Optimization Approaches, Mixture and Blend Design
Xu and Diwekar ²⁵⁵	Liquid–liquid extraction of acetic acid from water	CAMD, uncertainty	Efficient and Hammersley GA	Genetic algorithms, Molecular Design Under Uncertainty
Xu and Diwekar ²⁵⁶	Heterogeneous azeotropic distillation of acetic acid from water	CAMPD, MOO, uncertainty	MOEGA	Stochastic Optimization Approaches, Multi-Objective Optimization, Molecular Design Under Uncertainty
Ulas and Diwekar ³⁴⁴	Batch separation of acetonitrile from water	MOO, uncertainty, CAMPCD	HSTA and NLP	Molecular, Process and Control Design
Papadopoulos and Linke ¹⁶¹	Liquid–liquid extraction and gas absorption	CAMD, MOO	SA	Simulated annealing, Multi-Objective Optimization
Papadopoulos and Linke ³²¹	Liquid–liquid extraction, gas absorption and extractive distillation	CAMPD, MOO	SA	Simulated annealing, Multi-Objective Optimization
Yang and Song ¹⁹⁴	Liquid–liquid extraction	CAMD	Generate and test	Generate and Test Approaches
Kazantzi et al. ³⁴⁷	Acid gas purification	CAMPD	Property clustering	Property Clustering and/or Reverse Formulations
Eljack and Eden ²⁹³	Liquid–liquid extraction of aniline, blanket wash solvents	CAMD	Property clustering	Property Clustering Approaches, Property Clustering and/or Reverse Formulations
Song and Song ²⁷⁴	Liquid–liquid extraction and extractive distillation	CAMD	SA	Simulated annealing
Papadopoulos and Seferlis ¹⁶⁹ , Papadopoulos et al. ¹⁶²	Extractive distillation	MOO, CAMPCD	SA, NLP	Molecular, Process and Control Design
Papadopoulos and Linke ³²⁶ , Papadopoulos et al. ¹⁵⁰	Liquid–liquid extraction	CAMPD, MOO, Grid/parallel computing	SA	Exploitation of Advanced Computing Infrastructures
Bommareddy et al. ^{349,350}	Acid gas purification	CAMPD	Property clustering	Property Clustering and/or Reverse Formulations
Diky et al. ³⁹³	Solvent design in ThermoData Engine of NIST	Uncertainty	Knowledge-based approach	Molecular Design Under Uncertainty
Kheireddine et al. ³⁵²	Recycling of lubricating oils	Blend and process design	Property clustering	Property Clustering and/or Reverse Formulations, Mixture and Blend Design
Gebreslassie and Diwekar ^{282,285}	Extraction of acetic acid from waste stream	Parallel computing	Efficient ACO	Ant colony optimization, Exploitation of Advanced Computing Infrastructures
Herring and Eden ²⁵⁷	De novo molecular design, design of molecules with desired boiling point	CAMD, spatial fragment descriptors	GA	Genetic algorithms
Zhang et al. ²³²	Extraction	CAMD	General MILP/ MINLP formulation	Deterministic optimization approaches in molecular design
Scheffczyk et al. ²⁵⁹	Liquid–liquid extraction of phenol and hydroxymethylfurfural from water	CAMD, QM	GA	Genetic algorithms, QM models in non-reactive systems
Khor et al. ³⁶¹	Oil extraction from palm pressed fiber, safety and health indices	CAMD, MOO	Disjunctive programming, fuzzy representation	Multi-Objective Optimization
Ooi et al. ^{362,363}	Solvents for oil extraction from palm pressed fiber, safety and health indices	CAMD, MOO	MILP	Multi-Objective Optimization
Ten et al. ³⁹⁸	Extraction of carotenoids, safety and occupational health indices	CAMD, MOO, uncertainty	MILP, fuzzy representation	Molecular Design Under Uncertainty
Liu et al. ²⁵⁸	Extractive distillation	CAMD	Hybrid gene and SA algorithm	Genetic algorithms

Table 18 (Continued)

References	Application	Problem and model class	Algorithm	Sections
Zhou et al. ³³³	Absorption–desorption process, separation of acetone from air	CAMPD	GA, NLP, CONOPT	Stochastic Optimization Approaches
Scheffczyk et al. ⁴³¹	Hybrid extraction–distillation process for the bio-based platform chemical γ -valerolactone	CAMPD, QM	GA	Molecular and Process Design
<i>Reactive extraction</i>				
Wang and Achenie ²²⁰	Extractive fermentation	CAMPD	MINLP, AP/OA	Deterministic optimization approaches in molecular design
Linke and Kokossis ³²⁰	Extractive fermentation	CAMPD	SA	Stochastic Optimization Approaches
Papadopoulos and Linke ^{163,322}	Extractive fermentation, liquid–liquid extraction, and extractive distillation	CAMPD, MOO	SA	Stochastic Optimization Approaches
Cheng and Wang ^{244,306,307}	Extractive fermentation	CAMPD	MHDE and MISQP	Deterministic Optimization Approaches, Deterministic optimization approaches in molecular design
Zhou et al. ²⁵²	Extractive reaction processes, liquid-phase reactions	CAMD	GA	Genetic algorithms
<i>Blanket wash solvents</i>				
Ostrovsky et al. ²²²	Blanket wash solvents for lithographic printing	CAMD	IA global optimization	Deterministic optimization approaches in molecular design
Achenie and Sinha ²²⁴	Blanket wash solvents for lithographic printing, process constraints	Blend and process design	IA global optimization	Deterministic optimization approaches in molecular design, Mixture and Blend Design
Chemngattavalappil et al. ²⁹⁵	Blanket wash solvents	CAMD	Property clustering	
Heintz et al. ¹⁵³	Blanket wash mixture substitution, substitution of chlorinated paraffins, extraction solvents	CAMP ^b D	GA	Genetic algorithms, Mixture and Blend Design
Sinha et al. ²²⁵	Blanket wash solvents for lithographic printing	CAMP ^b D	IA global optimization	Deterministic optimization approaches in molecular design, Mixture and Blend Design
<i>Metal degreasing</i>				
Eden et al. ²⁹²	VOC recovery from metal degreasing process	CAMPD	Property clustering	Property Clustering and/or Reverse Formulations
Eljack et al. ^{348,447}	Metal degreasing	CAMPD	Property clustering	Property Clustering and/or Reverse Formulations
Chemngattavalappil et al. ⁴⁴⁵	Metal degreasing	CAMD, topological connectivity indices	Property clustering	Solvents for Industrial Separations, Reactive Separations and Promotion of Reactions
Samudra and Sahinidis ¹⁵⁸	Metal degreasing and crystallization solvents	CAMD, molecular graphs	MILP, BARON	Deterministic optimization approaches in molecular design
<i>Solvents for pharmaceutical industry with emphasis on crystallization</i>				
Karunanithi et al. ²²⁷	Cooling and drowning out crystallization of ibuprofen, process constraints	CAMPD, CAMP ^b D	Framework for MINLP formulations	Deterministic optimization approaches in molecular design, Deterministic Optimization Approaches, Mixture and Blend Design
Sheldon et al. ²³⁴	Solvents for pharmaceutical and agrochemical industry	CAMD, QM	MINLP, OA/ER	Deterministic optimization approaches in molecular design, QM models in non-reactive systems
Karunanithi et al. ²²⁸	Crystallization of carboxylic acids, process constraints	CAMPD, CAMP ^b D	Framework for MINLP formulations	Deterministic optimization approaches in molecular design
Weis and Visco ⁴⁴⁶	Solvents for pharmaceuticals industry	Solvent selection, signature molecular descriptor	Complete enumeration	Solvents for Industrial Separations, Reactive Separations and Promotion of Reactions

(Continued)

Table 18 (Continued)

References	Application	Problem and model class	Algorithm	Sections
Jonuzaj et al. ²³⁸ , Jonuzaj and Adjiman ²⁴⁰	Liquid–liquid extraction, solubility of ibuprofen	Blend design	MINLP, GDP, HR	Deterministic optimization approaches in molecular design, Mixture and Blend Design
Austin et al. ¹⁴⁹	Purification of ibuprofen via cooling crystallization	CAMPD	DFO, MILP, BARON	Deterministic optimization approaches in molecular design, Mixture and Blend Design
Wang and Lakerveld ³⁴⁰	Continuous antisolvent crystallization of paracetamol, PC-SAFT	CAMPD	NLP, continuous representation	Continuous Molecular Representations
Jonuzaj et al. ²³⁹	Liquid–liquid extraction, solubility of ibuprofen	CAMPD	MINLP, GDP, BM	Mixture and Blend Design
<i>Solvents for promotion of reactions or reactive processes</i>				
Gani et al. ^{409,411}	Selection of solvents promoting organic reactions	Incorporation of reactivity properties in solvent screening	Knowledge- based approach	CAMPD in reactive systems
Stanescu and Achenie, ⁴¹⁴ Stanescu et al. ⁴¹⁵	Kolbe–Schmitt reaction	CAMPD, QM, reaction	ProCAMPD, ICAS	QM models in reactive systems
Folic et al. ³⁸⁹	Menschutkin reaction	CAMPD, uncertainty, reaction	MILP and MINLP formulations	CAMPD in reactive systems
Folic et al. ³⁹¹	Menschutkin reaction	CAMPD, uncertainty, reaction	MILP and MINLP formulations	CAMPD in reactive systems, QM models in reactive systems
Folic et al. ⁴¹⁰	Selection of solvents for replacement and multistep reactions	Incorporation of reactivity properties in solvent screening	Knowledge- based approach	CAMPD in reactive systems
De Vleeschouwer et al. ^{423,425}	Photoacidic compounds for fast proton transfer in absence of water	QM, reaction	Best First Search, BB-type algorithm	QM models in reactive systems
Sioukrou ³⁷⁷	Menschutkin and Cope elimination reaction	CAMPD, QM, reaction	MINLP, BARON, CONOPT	QM models in reactive systems
Struebing et al. ⁴¹⁹	Menschutkin reaction solvents	QM, reaction	MILP, CPLEX	QM models in reactive systems
Sioukrou et al. ³⁷⁷	GXL solvents, Diels–Alder reaction	Blend and process design, reaction	MINLP, SMIN- α BB	Deterministic optimization approaches in molecular design, Mixture and Blend Design, Molecular and Process Design
Zhou et al. ⁴⁴²	Diels–Alder reaction	CAMPD, QM, reaction	MINLP	Molecular and Process Design
Zhou et al. ³⁹²	Diels–Alder reaction	CAMPD, uncertainty, QM, reaction	MINLP	Molecular Design Under Uncertainty, QM models in reactive systems
Struebing et al. ⁴¹⁹	Menschutkin reaction	CAMPD, QM, reaction	MINLP, BARON	QM models in reactive systems
Austin et al. ¹⁴⁹	Menschutkin reaction, liquid–liquid extraction	CAMPD, QM, reaction	DFO, BARON	Mixture and Blend Design, QM models in reactive systems
Dev ³⁶⁹	Design of reactant esters and alcohols that generate corresponding ester and alcohol product optimizing a property index	CAMPD, MOO, tracking of changes in structures due to reactions, signature molecular descriptors	MILP or MINLP	Multi-Objective Optimization, QM models in reactive systems

developed for the removal of radioactive radium and barium from water used in hydraulic fracking, while it is further used by Doshi et al.²⁸⁷ for the removal of arsenic from water.

Screening of adsorbents for different separations is also performed using optimization by Gounaris et al.^{449,450} Although it is not a GC approach, it is an interesting application of optimization-based design of materials. The approach is based on the work of First et al.,^{451–453} who propose a systematic method for the characterization of the 3D pore structures of zeolites and metal-organic frameworks (MOFs) using optimization, geometry, and graph algorithms. The method starts with the crystallographic coordinates of a structure and automatically identifies the portals, channels, and cages of a zeolite and MOF, describing their geometry and connectivity. It can be used to calculate pore size distribution, accessible volume, accessible surface area, pore limiting diameter, and largest cavity diameter. This work is based on the original models of Gounaris et al.,^{449,450} who mention that they may be used for both shape-selective separations and catalysis.

10100 **Table 19** Applications in the design of catalysts, adsorbents, and ionic liquids

References	Application	Problem and model class	Algorithm	Sections
<i>Catalysts</i>				
Chavali et al. ²⁷⁷	Molybdenum catalyst for epoxidation and hydroformylation reaction	CAMD, topological connectivity indices	TS	Tabu search, Catalysts, Adsorbents and Ionic Liquids
Lin et al. ²⁷⁸	Molybdenum catalyst for an epoxidation reaction	CAMD, topological connectivity indices	TS	Tabu search, Catalysts, Adsorbents and Ionic Liquids
<i>Adsorbents</i>				
Gounaris et al. ^{449,454}	Zeolites for hydrogen recovery from inorganic gases and light hydrocarbon streams, dehydration of gas streams, separation of light olefin/paraffin, separation of normal paraffins from branched isomers, separation of fructose from glucose, separation of fatty acids	Optimization-based characterization of zeolites	NLP	Catalysts, Adsorbents and Ionic Liquids
First et al. ^{451–453}	Zeolites and MOFs for separation of CO ₂ /N ₂ , CO ₂ /CH ₄ , CO ₂ /H ₂ , O ₂ /N ₂ , propane/propylene, ethane/ethylene, styrene/ethylbenzene, xylene isomers	Optimization-based characterization of zeolites and MOFs	MILP or NLP	Catalysts, Adsorbents and Ionic Liquids
Behavides et al. ^{288,289}	Adsorbents for NORM of natural gas fracking waste	CAMD, fitting of UNIFAC parameters	ACO	Ant colony optimization, Catalysts, Adsorbents and Ionic Liquids
Doshi et al. ²⁸⁷	Adsorbents for arsenic removal from water	CAMD, fitting of UNIFAC parameters	Efficient ACO	Ant colony optimization, Catalysts, Adsorbents and Ionic Liquids
<i>Ionic liquids</i>				
Matsuda et al. ⁴⁵⁴	Ionic liquids for favorable ionic conductivity and viscosity	Exhaustive search, fitting of GC-QSPR model for property prediction	Generate and test	Catalysts, Adsorbents and Ionic Liquids
McLeese et al. ²⁸⁰	Ionic liquids for use within a hydrofluorocarbon (refrigerant) gas separation system	CAMD, topological connectivity indices	TS	Tabu search
Chavez-Islas et al. ⁴⁵⁵	Ionic liquids as solvents for bio-ethanol recuperation	CAMPD	MINLP, BARON, SBB, CONOPT	Catalysts, Adsorbents and Ionic Liquids
Roughton et al. ³¹⁰	Ionic liquids as azeotropic distillation solvents	CAMPD	MILP	Deterministic Optimization Approaches
Karunanithi and Mehrkesh ⁴⁵⁶	Ionic liquids for high electrical conductivity and for toluene–heptane separation	CAMD	Enumeration and gradual reduction of initial set based on constraints, GA used to solve the problem and compare results	Catalysts, Adsorbents and Ionic Liquids
Hada et al. ²⁹⁹	Ionic liquid design as EBS	CAMD, QM, latent property GC models	Property clustering, enumeration of all candidates	QM models in non-reactive systems
Peng et al. ⁴²⁸	Ionic liquids for separation of cyclohexane from benzene	CAMD, QM	MINLP	QM models in non-reactive systems
Zhang et al. ⁴²⁹	Ionic liquid design for separation of benzene from cyclohexane and deep desulfurization from gasoline	CAMD, QM	Mixed SA-GA algorithm	QM models in non-reactive systems
Zhao et al. ⁴³²	Ionic liquids, separation of C ₂ H ₂ /C ₂ H ₄	Screening of ionic liquids database, QM	Gradual reduction of initial set of candidates based on constraints	QM models in non-reactive systems
Song et al. ⁴⁴⁴	Ionic liquids for extractive desulfurization of fuel oils	Enumeration of structures and process simulation, QM	Generate and test	Molecular and Process Design

10105 **Table 20** Contributions in materials and process design for CO₂ capture

References	Application	Problem and model class	Algorithm	Sections
<i>Solvents for physical separation</i>				
Gani et al. ¹⁸⁷ Bardow et al. ¹⁷⁵	Physical absorption Physical absorption	CAMD CAMPD, PC-SAFT	Generate and test NLP, continuous molecular representation	Generate and Test Approaches Continuous Molecular Representations
Stavrou et al. ³³⁴	Physical absorption	CAMPD, PC-SAFT	NLP, continuous molecular representation	Continuous Molecular Representations
Qadir et al. ³²⁹	Pre-combustion capture, physical absorption	CAMPD, pure solvent and blend design, PPC-SAFT	GA, internal SQP for continuous parameters	Stochastic Optimization Approaches
Burger et al. ²³⁷	Physical absorption	CAMPD, MOO, GC SAFT- γ -Mie	MINLP, OA/ER/AP, sandwiching algorithm	Deterministic optimization approaches in molecular design, Deterministic Optimization Approaches, Multi-Objective Optimization
Lampe et al. ³³⁶	Physical absorption	CAMPD, GPC-SAFT	NLP, continuous molecular representation	Continuous Molecular Representations
Chong et al. ⁴⁵⁷	Ionic liquids, physical absorption	CAMD	MINLP	Materials for CO ₂ Capture
Farahipour et al. ⁴²⁷	Ionic liquids, physical absorption	Screening of database, QM	Gradual reduction of initial set of candidates based on constraints	QM models in non-reactive systems
Liu et al. ⁴⁴³	Ionic liquids, physical absorption	Simulation of different ionic liquids in single and multistage flash flow sheets, QM	Enumeration-based solvent selection	Molecular and Process Design
Gopinath et al. ²⁴¹ , Gopinath ¹⁷⁹	Physical absorption	CAMPD, GC SAFT- γ -Mie	MINLP, modified OA	Deterministic optimization approaches in molecular design, Deterministic Optimization Approaches
Chong et al. ^{378,379}	Ionic liquids and mixtures, physical absorption	CAMP ^b D	Property clustering	Mixture and Blend Design
Chong et al. ⁴⁵⁸	Ionic liquids, physical absorption in bioenergy carbon capture and storage system	CAMD	MINLP, disjunctive programming	Materials for CO ₂ Capture
Zhao et al. ⁴³²	Ionic liquids, separation of CO ₂ /CH ₄	Screening of ionic liquids database, QM	Gradual reduction of initial set of candidates based on constraints	QM models in non-reactive systems
Peng et al. ⁴²⁸	Ionic liquids, physical absorption	CAMD, QM	MINLP, BB	QM models in non-reactive systems
Valencia-Marquez et al. ³⁷⁶	Ionic liquids, physical absorption	CAMPD, SOO, and MOO	MINLP, SBB	Multi-Objective Optimization
<i>Solvents for chemical separation</i>				
Mac Dowell et al. ³⁴³	Chemical absorption, chemical and phase equilibrium	Blend and process design, SAFT-VR	NLP, continuous molecular representation	Continuous Molecular Representations
Pereira et al. ^{149,438}	Chemical absorption, chemical and phase equilibrium	CAMPD, CAM ^b D, SAFT-VR	NLP, continuous molecular representation	Continuous Molecular Representations
Salazar et al. ⁴³⁹	Chemical absorption, phase equilibrium	CAMPD, UNIFAC, and eNRTL	N/A	Molecular and Process Design
Chemmangattuvalpil and Eden ²⁹⁷	Chemical absorption, nonreactive pure component properties as design criteria	CAMD within process constraints, signature molecular descriptors	Property clustering, MILP	Materials for CO ₂ Capture

Table 20 (Continued)

References	Application	Problem and model class	Algorithm	Sections
Ng et al. ³⁵⁹	Chemical absorption, nonreactive pure component properties as design criteria	CAMD, MOO, signature molecular descriptors	MILP, fuzzy representation	Multi-Objective Optimization
Papadokonstantakis et al. ³²⁵	Chemical absorption, chemical and phase equilibrium	CAMPD, MOO, SAFT-VR, cradle-to-gate and gate-to-gate sustainability assessment	SA, NLP	Molecular and Process Design
Papadopoulos et al. ³⁹⁴	Chemical absorption, chemical and phase equilibrium	CAMD, MOO, uncertainty, GC SAFT- γ -SW, cradle-to-gate sustainability assessment	SA	Molecular Design Under Uncertainty
Ten et al. ^{364,366}	Chemical absorption, nonreactive pure component properties as design criteria, health and safety indices	CAMD, MOO	MILP, fuzzy representation, disjunctive programming	Multi-Objective Optimization
Zarogiannis et al. ¹⁶⁵ , Papadopoulos et al. ¹⁶⁴	Chemical absorption, no reactivity properties	MOO, blend design, uncertainty	Complete enumeration	Mixture and Blend Design
Limleamthong et al. ³⁶⁰	Chemical absorption, no reactivity properties	MOO	Linear Programming, DEA	Multi-Objective Optimization
Ahmad et al. ⁴³³	Chemical absorption	CAMD, QM, reaction mechanisms	ProCAMD, ICAS	QM models in non-reactive systems
<i>Adsorbents for PSA/VSA</i> Hasan et al. ⁴⁵⁹	Zeolites for PSA/VSA	Screening of zeolites based on geometry, grand canonical Monte Carlo for determination of equilibrium and kinetics in selected zeolites, optimization-based process design using surrogate model	NLP	Materials for CO ₂ Capture
First et al. ⁴⁶⁰	Zeolites for PSA	Screening of zeolites based on geometry, grand canonical Monte Carlo for determination of equilibrium and kinetics in selected zeolites, optimization-based process design using surrogate model	MILP for minimum energy pathway through zeolite, NLP for process optimization	Materials for CO ₂ Capture

p2145 Ionic liquids are organic salts usually consisted of a bulky organic cation and an organic or inorganic anion and with melting points near room temperature. They are characterized by extremely low vapor pressures, wide liquid ranges, nonflammability, thermal stability, tunable polarity, good electrolytic and solvation properties, and high viscosity. Such properties are often considered as performance indices in the design of ionic liquids, together with solubility.

s0350 **Materials for CO₂ Capture**

p2150 The design of CO₂ capture materials using CAMD approaches involves mainly research into the identification of solvents that are able to reduce capital and operating costs. Table 20 summarizes the main developments. Zeolites have also been considered through optimization-based material screening and process design approaches. In the area of solvent-based separations, CO₂ capture processes which are based on physical absorption often include a pressurized absorption column which is followed by a flash process for solvent regeneration. Being an absorption process, the key requirement is to identify solvents that exhibit high CO₂ solubility. Notable works in this area include those using different versions of SAFT to model the phase equilibrium, as well as the use of GC or continuous molecular representation approaches for solvent and process design (see “Methodologies for Solution of CAMPD Problems” section). Ionic liquids have also been proposed as potential CO₂ capture solvents in this case. These are considered as EBSs, mainly due to extremely low vapor pressure. However, they often exhibit high viscosity hence it is important to use it in the minimization criteria during solvent design.

p2155 Chemical absorption processes include mainly amine-based solvents which react with CO₂ in packed-bed absorption columns and the solvent is then regenerated thermally in desorption columns. The process is energy intensive, because the regeneration takes place at approximately 120°C, while the amine solvents have significant health, safety, and environmental impacts. The

identification through CAMD approaches of solvents which address these shortcomings is difficult, because both the solvent CO₂ solubility and reaction kinetics need to be accounted for during solvent design. To date, there are different approaches addressing the design of CO₂ capture solvents through chemical absorption. One approach includes contributions which employ pure component properties as design criteria focusing on properties such as solvent heat of vaporization, molar volume, density, and viscosity (e.g., Ten et al.³⁶⁴). While these are important properties, solvent CO₂ solubility and reaction kinetics should also be considered. An important feature of such works is the consideration of health-related properties such as soil-sorption coefficient and toxic limit concentration Chemmangattuvalappil and Eden²⁹⁷ or of a more inclusive evaluation of multiple health and safety indices Ten et al.³⁶⁴

p2160 Different approaches consider equilibrium calculations to determine the CO₂ solubility in solvent–water mixtures. Such approaches include works by Salazar et al.,⁴³⁹ Mac Dowell et al.,³⁴³ Pereira et al.,^{152,438} Papadokonstantakis et al.,³²⁵ and Papadopoulos et al.³⁹⁴ reviewed in “Molecular and Process Design” section. An important feature of Papadokonstantakis et al.³²⁵ and Papadopoulos et al.³⁹⁴ is that CAMD is first implemented considering numerous pure component properties such as solvent CO₂ solubility, solvent basicity (which is a measure of reactivity), heat capacity, density and viscosity, and surface tension, together with a complete cradle-to-gate life cycle assessment and environmental, health, and safety hazards analysis. A set of solvents that exhibit desirable performance in these properties is then evaluated using GC SAFT- γ -SW EoS which allows the calculation of the equilibrium CO₂ solubility. Mac Dowell et al.³⁴³ and Pereira et al.^{152,438} also use SAFT-based implementations, while they also consider process together with solvent design. Papadokonstantakis et al.³²⁵ consider process design and gate-to-gate sustainability process assessment for selected solvents.

p2165 In a different line of work, Hasan et al.⁴⁵⁹ and First et al.⁴⁶⁰ propose a hierarchical screening of zeolites as adsorbents for CO₂ capture using pressure swing adsorption or vacuum swing adsorption (PSA/VSA). The method is not based on GC but falls within the CAMPD philosophy of evaluating materials, considering optimization-based process design. Potential zeolites based on their pore sizes are first identified. Next, shape, size, and adsorption selectivity are calculated to shortlist few options. Grand canonical Monte Carlo is used for the derivation of equilibrium and kinetic relationships. A detailed simulation and sampling-based process optimization technique is used to optimize a four-step PSA/VSA process, supported by the derivation of a Kriging-based model.

s0355 **Heat Exchange Fluids**

p2170 The design of heat exchange fluids involves applications in refrigeration cycles, heat pumps, and ORC (Table 21). All these processes involve pressure change operations with the fluid operating at two pressure levels, for basic cycle configurations. It is therefore important to identify fluids that operate at low pressures, but vacuum should be avoided if possible as it generally increases the capital and operating costs. At the desired pressure levels, the fluid should be able to extract as much heat as possible from the heat source. In the case of ORC, the fluid should enhance power generation through a turbine, where the formation of liquid should be avoided. The use of mixtures instead of pure fluids facilitates heat extraction as they provide a better match of the temperature profile of the heat source. This is because mixtures exhibit variable phase-change temperature, as opposed to the constant phase-change temperature of pure fluids, hence pinches are avoided. Matching of the heat source temperature profile can be achieved through combined selection of appropriate fluids and design of multipressure systems or tightly integrated heat exchange networks.³¹⁸ With respect to fluid properties, heat of vaporization, heat capacity, thermal conductivity, and viscosity are important properties. Flammability, toxicity, global warming, and ozone depletion potentials have also been considered in published research. Noteworthy contributions further include the use of SAFT-based models. A review of working fluid, process design control, and integration of ORC is available in Linke et al.¹⁷⁷

s0360 **Polymers**

p2175 The design of polymers includes the identification of the repeat unit structure of a polymer that satisfies a set of desired macroscopic properties.^{251,464} This is a typical CAMD problem which may use GC methods such as the ones proposed by Van Krevelen and Te Nijenhuis.¹⁷ There are currently numerous property classes that can be predicted by GC methods including volumetric, calorimetric, cohesive properties and solubility, transition temperatures, interfacial energy properties, optical, electrical, magnetic, mechanical, acoustic properties, chemical stability, and various transport properties, to name but a few. Table 22 shows the main features of the reported applications.

s0365 **Bio-Based Fuels and Chemicals**

p2180 The contributions included in Table 23 involve the design of bio-based fuels and other chemicals as well as design of potential production pathways. In the area of biofuels, there are several challenges that may be overcome through appropriate design of fuel mixtures and blends. As noted in Hada et al.,²⁹⁸ the oxidative stability and low-temperature operability are key issues, as they have an inverse relationship; structural modifications that improve oxidative stability adversely affect low-temperature operability and vice versa. Fuel additives are used to address such issues, but the proposed formulations need to cover biofuels originating from different feedstocks in order to meet quality regulations. To this end, properties which are important for the design of fuel additives

10110 **Table 21** Contributions in CAMD for heat exchange fluids in refrigeration, heat pump, and ORC processes

References	Application	Problem and model class	Algorithm	Sections
Joback ¹⁸⁴	Refrigerants	CAMD	Generate and test	Generate and Test Approaches
Gani et al. ¹⁸⁷	Refrigerants	CAMD	Generate and test	Generate and Test Approaches
Duvedi and Achenie ²¹⁷	Refrigeration	CAMPD	MINLP, OA/AP	Deterministic Optimization Approaches
Churi and Achenie ²¹⁹	Refrigeration	Blend design, process constraints	MINLP, OA/AP	Deterministic optimization approaches in molecular design
Duvedi and Achenie ²¹⁸	Refrigeration	Blend design, process constraints	MINLP, OA/AP	Deterministic Optimization Approaches, Mixture and Blend Design
Marcoulaki and Kokossis ²⁶⁸	Refrigerants	CAMD	SA	Simulated annealing, Stochastic Optimization Approaches
Lee et al. ³⁸⁰	Refrigeration	Blend and process design	Composite curves and NLP	Mixture and Blend Design
Sahinidis et al. ²⁴²	Refrigerants	CAMD	Branch-and-reduce global optimization	Deterministic optimization approaches in molecular design
Lehmann and Maranas ¹⁷²	Refrigerants	QM	GA	Genetic algorithms, QM models in non-reactive systems
Solvason et al. ³⁸¹	Refrigerants	CAM ^b D	Property clustering, complete enumeration	Mixture and Blend Design
Samudra and Sahinidis ¹⁵⁸	Refrigerants	CAMD, molecular graphs	MILP, BARON, XPRESS	Deterministic optimization approaches in molecular design
Cignitti et al. ³¹⁷	Refrigeration	CAMPD, CAM ^b D	MINLP, SBB	Deterministic Optimization Approaches
Ourique and Telles ²⁷¹	Heat pump	CAMD, molecular graphs	SA	Simulated annealing
Papadopoulos et al. ³²³	ORC	CAMPD, MOO	SA	Heat Exchange Fluids
Papadopoulos et al. ³²⁴	ORC	CAMPD, MOO, CAM ^b D	SA	Stochastic Optimization Approaches, Multi-Objective Optimization, Mixture and Blend Design, Molecular Design Under Uncertainty
Lampe et al. ³³⁵	ORC	CAMPD, PC-SAFT	NLP, continuous molecular representation	Continuous Molecular Representations
Roskosch and Atakan ^{176,341}	ORC and heat pump	CAMPD	NLP, continuous molecular representation	Continuous Molecular Representations
Lampe et al. ³³⁵	ORC	CAMPD, GPC-SAFT	NLP, continuous molecular representation	Continuous Molecular Representations
Molina-Thierry and Flores-Tlacuahuac ³¹²	ORC	Blend and process design	NLP	Deterministic Optimization Approaches, Mixture and Blend Design
Palma-Flores et al. ³¹¹	ORC	CAMPD	NLP	Deterministic Optimization Approaches
Mavrou et al. ³⁸²	ORC	Blend and process design, MOO	Complete enumeration	Mixture and Blend Design
Mavrou et al. ³⁸³	ORC	Blend and process design, MOO, uncertainty	Complete enumeration	Mixture and Blend Design
Frutiger et al. ^{399,402} , Frutiger ³⁴²	ORC and heat pump	CAMPD, uncertainty	See "Molecular Design Under Uncertainty" section	Continuous Molecular Representations, Molecular Design Under Uncertainty
Thierry et al. ³¹⁹	ORC	Blend and process design	MINLP, GDP	Deterministic Optimization Approaches
Schilling et al. ^{338,339}	ORC	CAMPD, MOO, GPC-SAFT	MINLP, OA/ER, BARON, DICOPT, SNOPT	Continuous Molecular Representations, Multi-Objective Optimization
Santos-Rodriguez et al. ³¹³	ORC	Blend and process design, uncertainty	Deterministic stochastic programming, MINLP, CONOPT	Deterministic Optimization Approaches, Mixture and Blend Design, Molecular Design Under Uncertainty

(Continued)

Table 21 (Continued)

References	Application	Problem and model class	Algorithm	Sections
Cignitti et al. ³¹³	ORC	CAMPD	Simultaneous approach: MINLP, LINDOGlobal ⁴⁶⁴ . Decomposed approach: MILP with BARON and NLP with CONOPT	Deterministic Optimization Approaches
White et al. ³¹⁴	ORC	CAMPD, GC SAFT- γ -Mie	MINLP, OA/ER/AP	Deterministic Optimization Approaches
Schwobel et al. ⁴⁶¹	ORC	Screening working fluids and ORC simulations, QM, COSMO-RS	Complete enumeration of 75 million fluids	Heat Exchange Fluids
Andres-Martinez et al. ⁴⁰⁴	ORC	CAMD, uncertainty	MINLP	Molecular Design Under Uncertainty
Stijepovic et al. ^{318,462}	ORC	Selection of molecules and process design	MINLP, BARON	Deterministic Optimization Approaches
Karunanithi and Mehrkesh ⁴⁵⁶	Ionic liquids for thermal energy storage (high thermal conductivity)	CAMD	Enumeration and gradual reduction of initial set based on constraints, GA used to solve the problem and compare results	Heat Exchange Fluids
Mehrkesh and Karunanithi ⁴⁶³	Ionic liquids for thermal energy storage (high thermal storage density)	CAMD	GA	Heat Exchange Fluids

include vapor pressure, octane rating and heating value, cetane number, viscosity and density, flash point, lethal concentration, Gibbs energy of mixing, oxygen content, etc. Such properties are associated with the ability of the fuel to be burned, engine efficiency, efficient fuel flow, safety and health, stability, and environmental impacts.³⁸⁷ Additives for lubricants are also important, with associated properties discussed in Yunus et al.³⁸⁷ Similar properties are used in approaches where biofuels or other chemicals are designed using CAMD and then the feasibility of their production is evaluated. The latter is based on the use of platform chemicals and either the evaluation of synthesis of the designed products through reaction network analysis or the evaluation of their production through process design approaches.

s0370 **Formulated Products and Other Chemicals**

p2185 This section investigates the design of formulated products⁴⁶⁹ as well as other chemicals which may be the active ingredients in formulated products and pertain to pharmaceuticals, agricultural, and cosmetics industry. Table 24 summarizes the main features of the reported approaches and applications. Formulated products are mainly complex blends, comprising several different components which result in desirable functionality and performance of the final product. For example, Conte et al.¹⁹¹ classify liquid product components into active ingredients which determine the function of the product (e.g., the ingredient that repels insects in the corresponding lotion), additives used to enhance the end-use properties, and the solvent which is used to dissolve the active ingredient and additives to enable easy and dependable use. Martin and Martinez³⁰⁸ provide an indicative example of the different components which may be found in detergents, including surfactants, builders, enzymes, polymers, bleach, softeners, stabilizers, preservatives, fragrances, and colorants. The need of molecular design approaches is quite clear, but these are just part of the product design problem. Fung et al.³⁰⁹ provide an overview of all the complex decision-making required from the selection of ingredients all the way to product commercialization. Other important applications of molecular design approaches include the design of active ingredients for human medication, of chemicals for agricultural applications, food enhancers, fragrances, etc.

s0375 **Future Outlook and Further Reading**

p2190 The presented work highlights all major classes of molecular design technologies and approaches, showing how the original concept evolved toward various different directions. The integration of process decisions and equipment models into molecular design is a mature, but still very active research field. The incorporation of rigorous QM-based models in molecular design is an

10115 **Table 22** Design of polymers

References	Application	Problem and model class	Algorithm	Sections
Derringer and Markham ¹⁸⁶	Polymer repeat units	CAMD	Generate and test	Generate and Test Approaches
Venkatasubramanian et al., ^{253,264} Patkar and Venkatasubramanian ²⁵¹	Various polymers	CAMD	GA	Genetic algorithms
Vaidyanathan and El-Halwagi ²²⁹	Design of polycarbonate of bisphenol A	CAM ^b D	MINLP, IA global optimization	Deterministic optimization approaches in molecular design, Mixture and Blend Design
Constantinou et al. ¹⁸⁸	Polymer design and foam blowing and cleaning agents for textiles	CAMD	Generate and test	Generate and Test Approaches
Maranas ²¹⁵	Polymer repeat units	CAMD	MINLP to MILP, GAMS/OSL	Deterministic optimization approaches in molecular design
Maranas ¹⁶⁸	Polymer repeat units	CAMD, uncertainty	MINLP, AP/OA, DICOPT	Deterministic optimization approaches in molecular design, Molecular Design Under Uncertainty
Vaidyanathan and El-Halwagi ⁴⁶⁵	Fiber re-enforced polymer composites, design of polymer forming the matrix of composite	CAMD	MINLP, GINO	Deterministic optimization approaches in molecular design, Mixture and Blend Design
Camarda and Maranas ²¹⁶	Polymer repeat units	CAMD, topological indices	MINLP, AP/OA, DICOPT	Deterministic optimization approaches in molecular design
Satyanarayana et al. ^{405,406}	Polymers as coatings and synthetic fibers	CAMD, Grid/parallel computing, connectivity indices	Harper and Gani ¹⁹⁰ approach	Exploitation of Advanced Computing Infrastructures
Eslick et al. ²⁷⁹	Cross-linked polymer networks	CAMD	TS	Tabu search
Solvason et al. ³⁸⁶	Polymer blend of spun yarn for marine application	Blend design, topological connectivity indices	Property clustering, complete enumeration	Mixture and Blend Design
Satyanarayana et al. ⁴²⁶	Polymers as hermetic stoppers	CAMD, molecular dynamics, connectivity indices	Harper and Gani ¹⁹⁰ approach	QM models in non-reactive systems
Pavurala and Achenie ⁴⁶⁶	Polymers for oral drug delivery	CAMD	MINLP, OA	Polymers
Zhang et al. ²³²	Polymer repeat units	CAMD	General MILP/ MINLP formulation	Deterministic optimization approaches in molecular design
Abedin et al. ²⁸¹	Water compatible visible light photosensitizers for dental adhesive	CAMD, topological connectivity indices	TS	Tabu search
Mukherjee et al. ²⁸⁶	Polymeric adsorbents for metal ion removal from water	CAMD, fitting of UNIFAC parameters	EACO	Ant Colony Optimization
Hada et al. ³⁰⁰	Development of thermoplastic from the mixing of starches, lactic acids and additives	Blend design, latent variable models	Property clustering	Property Clustering Approaches, Mixture and Blend Design

emerging area of research, with very intense interest in applications including reactive systems. The use of rigorous molecular models involves activities observed in several applications which have not yet been addressed using CAMD-based methodologies. For example, the combination of optimization algorithms with rigorous molecular models is also reported in the active domain of CADD.²⁶³ However, such activities lack the systematic approaches described here, hence there is a great scope to transfer and further develop CAMD methodologies in domains such as CADD. Furthermore, the use of EoS other than cubic is also emerging rapidly in CAMD. Such EoS enable molecular design for challenging activities, including applications that involve multiphase equilibria, high pressures, and systems, that exhibit electrolytic behavior or include reactions. In all these cases, the development of advanced optimization approaches is needed in order to efficiently address the mathematical challenges that result from very nonideal behaviors and high combinatorial complexity. The development and utilization of surrogate models also appears as an emerging approach to address computational challenges. The Kriging model has been reported as a surrogate model type in few applications,^{159,474} while latent variable methods such as principal component analysis have been used (e.g., Hada et al.^{299,300}). Recent developments also include technologies such as machine learning for high-throughput screening of chemicals⁴⁷⁵ and molecular design.⁴⁷² Additionally, efforts to exploit advanced computational infrastructures and information technologies are needed. Finally, the design of complex products is also an emerging area. New and efficient

10120 **Table 23** Applications in the design of bio-based fuels and chemicals

References	Application	Problem and model class	Algorithms	Sections
<i>Design of bio-based fuels and chemicals</i>				
Camarda and Sunderesan ⁴⁶⁷	Fuel additives and detergents, textile surfactants as soybean derivatives	CAMD, topological molecular connectivity indices	MILP	Bio-Based Fuels and Chemicals
Hada et al. ²⁹⁸	Biofuel additives	CAMP ^D , latent property GC models	Property clustering	Mixture and Blend Design
Yunus et al. ³⁸⁷	Additives for fuels and lubricants	Blend design	NLP	Mixture and Blend Design
Perdomo et al. ⁴⁶⁸	Optimum composition of fatty acid methyl esters (FAMEs)	Blend design, SAFT- γ	GA	Bio-Based Fuels and Chemicals
Woo et al. ³⁰¹	Design of biofuel additives	Blend design, statistical model development	Mixed-integer problem formulation	Property Clustering Approaches
<i>Design of bio-based fuels and chemicals with production pathways</i>				
Hechinger et al., ⁴¹⁶ Dahmen and Marquardt ⁴¹⁸	Biorenewable fuel and blend candidates and identification of production pathways	CAMD, RNFA	ProCAMD, ICAS	QM models in reactive systems
Ng et al. ³⁵³	Design of biochemicals and biorefineries to produce them	CAMPD, signature molecular descriptors	MILP	Property Clustering and/or Reverse Formulations
Ng et al. ³⁵⁴	Design of fuel additives from palm-based biomass and biorefineries to produce them	CAMPD, MOO, CAMP ^D , signature molecular descriptors	MILP, fuzzy representation	Property Clustering and/or Reverse Formulations, Multi-Objective Optimization, Mixture and Blend Design
Gerbaud et al. ⁴¹³	Identification of molecules from biorenewable feedstocks and of production pathways	CAMD, CAOS approach	GA (approach of Heintz et al. ¹⁵³)	CAMD in reactive systems

methods are needed to address the need for considering mixtures of multiple and diverse components and for accounting of a large number of activities prior to commercialization. This is an area where different molecular design approaches and technologies can converge into software systems which can cover the entire product design chain, from product identification to commercialization.

p2195 The iCAMD special issue⁴⁷⁶ includes a large collection of publications with contemporary advances in molecular design. Martin et al.⁴⁷⁷ published a collective volume with several different contributions on chemical product design. The volume includes methodological advances together with important applications. Adjiman and Galindo⁴⁷⁸ also published a collective volume on molecular systems engineering which includes contributions regarding the development of property prediction models as well as molecular design. Austin et al.⁴ present a structured review of molecular design approaches, with analysis of property modeling, mathematical optimization, and applications. In the area of product design, Gani and Ng⁴⁷⁹ present an inclusive work that identifies challenges in product design, classifies product design by market sector and product type, and proposes a systematic product design approach. In the area of pharmaceuticals, Papadakis et al.⁴⁸⁰ present a perspective paper on pharmaceutical process development where product design approaches are also discussed as part of the entire process. The work is included in a collective volume by Singh and Yan⁴⁸¹ which includes additional contributions regarding product and process design in pharmaceutical industries. An inclusive framework on pharmaceutical product design is also included in the work of Gernaey and Gani.⁴⁸²

p2200 The computer-aided design of materials for CO₂ capture processes is covered in an extensive volume by Papadopoulos and Seferlis.⁴⁸³ The work includes contributions that cover the entire area of computer-aided CO₂ capture systems design, from material modeling and design, to process modeling, design, control, and integration. Applications are reported for several CO₂ capture technologies, including solvent-based separations, PSA/VSA materials and systems, membrane-based separations, as well as oxycombustion and calcium looping systems. Finally, in the area of ORC, the review of Linke et al.¹⁷⁷ covers developments from working fluid design, to process design, control, and integration.

10125 **Table 24** Applications in the design of formulated and other chemicals

References	Application	Problem and model class	Algorithm	Sections
<i>Formulated products</i>				
Conte et al. ^{191,192}	Formulated products, paint formulation, insect repellent lotion, solvents for hair spray	CAMPD, CAM ^b D	Generate and test, ICAS	Generate and Test Approaches, Mixture and Blend Design
Martin and Martinez ³⁰⁸	Detergents	Blend and process design	MINLP, BARON	Deterministic Optimization Approaches, Mixture and Blend Design
Mattei et al. ^{193,388} Ng et al. ³⁹⁷	Surfactants for emulsions Design of alkyl substituent for fungicide	CAM ^b D CAMD, MOO, uncertainty, signature molecular descriptors	Generate and test MILP, fuzzy representation	Mixture and Blend Design Molecular Design Under Uncertainty
Zhang et al. ²³²	Surfactant design for UV sunscreen in emulsion	CAMD	General MILP/MINLP formulation	Deterministic optimization approaches in molecular design
Martin and Martinez ³⁹⁶	Detergents	Blend design, uncertainty	MINLP, BARON	Molecular Design Under Uncertainty
Fung et al. ³⁰⁹	Die attach adhesive, hand lotion	CAMPD	MINLP	Deterministic Optimization Approaches
<i>Other chemicals</i>				
Joback ¹⁸⁴	General methodology for drug design applied to asthma drugs	CAMD	Generate and test	Generate and Test Approaches
Raman and Maranas ¹⁰	Substitution of dialkyldithiolanylidenemalonates with optimum fungicidal and insecticidal properties	CAMD, uncertainty, topological connectivity indices	MILP, CPLEX, MINLP for deterministic problem, DICOPT	Deterministic optimization approaches in molecular design, Molecular Design Under Uncertainty
Buxton et al. ^{407,408}	Derivation of alternative reaction pathways for pesticide components	CAMD, reaction pathway	Optimization-based CAMD with knowledge-based approach	CAMD in reactive systems
Friedler et al. ²²³	Polyhalogenated biphenyls, flavor constituents in perfumes	CAMD	MINLP, BB	Deterministic optimization approaches in molecular design
Siddhaye et al. ⁴⁷⁰	Design of alcohol related to bloodstream uptake of orally delivered drugs	CAMD, topological connectivity indices	MILP, GAMS/OSL	Formulated Products and Other Chemicals
Siddhaye et al. ⁴⁷¹	Novel penicillin derivative	CAMD, topological connectivity indices	MILP, GAMS/OSL	Formulated Products and Other Chemicals
Chemngattuvallappil et al. ²⁹⁶	Design of alkyl substituent for fungicide	CAMD, signature molecular descriptors	Property clustering	Property Clustering Approaches
Zhang et al. ⁴⁷²	Design of fragrance molecules as shampoo additives and insect repellent sprays	CAMD, machine learning, GC	MINLP, GAMS, OptCAMD, ProCAMD ⁴⁷³	Formulated Products and Other Chemicals

References

1. Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. *AIChE J.* **1975**, *21*, 1086–1099.
2. Gani, R.; Brignole, E. A. *Fluid Phase Equilib.* **1983**, *13*, 331–340.
3. Adjiman, C. S.; Galindo, A.; Jackson, G. *Comput.-Aided Chem. Eng.* **2014**, *34*, 55–64.
4. Austin, N. D.; Sahinidis, N. V.; Trahan, D. W. *Chem. Eng. Res. Design* **2016**, *116*, 2–26.
5. Poling, B. E.; Prausnitz, J. M.; O'Connell, J. P. *The Properties of Gases and Liquids*, 5th edn; McGraw-Hill: New York, 2017.
6. Kolska, Z.; Zabransky, M.; Randova, A. Group contribution methods for estimation of selected physico-chemical properties of organic compounds. In *Thermodynamics-Fundamentals and Its Application in Science*; Morales-Rodriguez, R., Ed.; InTech: London, 2012;; pp 135–162. Chapter 6.
7. Langmuir, I. The Distribution and Orientation of Molecules. In *Third Colloid Symposium Monograph*, 73; Chemical Catalog Company: New York, 1925.
8. Redlich, O.; Derr, E. L.; Pierotti, G. J. *J. Am. Chem. Soc.* **1959**, *81* (10), 2283–2285.
9. Marrero, J.; Gani, R. *Fluid Phase Equilib.* **2001**, *183*, 183–208.
10. Raman, V. S.; Maranas, C. D. *Comput. Chem. Eng.* **1998**, *22*, 747.
11. Randic, M. *J. Am. Chem. Soc.* **1975**, *97* (23), 6609–6615.

12. Churchwell, C. J.; Rintoul, M. D.; Martin, S.; Visco, D. P., Jr.; Kotu, A.; Larson, R. S.; Sillerud, L. O.; Brown, D. C.; Faulon, J. L. *J. Mol. Graph. Model.* **2004**, *22* (4), 263–273.
13. Constantinou, L.; Gani, R. *AIChE J.* **1994**, *40* (10), 1697–1710.
14. Hukkerikar, A. S.; Sarup, B.; Ten Kate, A.; Abildskov, J.; Sin, G.; Gani, R. *Fluid Phase Equilib.* **2012**, *321*, 25–43.
15. Joback, K. G.; Reid, R. C. *Chem. Eng. Commun.* **1987**, *57* (1–6), 233–243.
16. Hukkerikar, A. S.; Kalakul, S.; Sarup, B.; Young, D. M.; Sin, G.; Gani, R. *J. Chem. Inf. Model.* **2012**, *52* (11), 2823–2839.
17. Van Krevelen, D. W.; Te Nijenhuis, K. *Properties of Polymers: Their Correlation With Chemical Structure and Their Numerical Estimation and Prediction From Additive Group Contributions*. Elsevier: Amsterdam, 2009.
18. Coutinho, J. A.; Carvalho, P. J.; Oliveira, N. M. *RSC Adv.* **2012**, *2* (19), 7322–7346.
19. Gmehling, J.; Constantinescu, D.; Schmid, B. *Annu. Rev. Chem. Biomol. Eng.* **2015**, *6*, 267–292.
20. Derr, E. L.; Deal, C. H. *Inst. Chem. Eng. Symp. Ser.* **1969**, *32* (3), 4051.
21. Kojima, K.; Tochigi, K. *Prediction of Vapor-Liquid Equilibria by the ASOG Method*. Elsevier: Amsterdam, 1979.
22. Weidlich, U. Ph.D. Thesis, University of Dortmund, Dortmund, DE, 1985.
23. Weidlich, U.; Gmehling, J. *Ind. Eng. Chem. Res.* **1987**, *26*, 1372–1381.
24. Larsen, B.L. Ph.D. Thesis, Technical University of Denmark, DK, 1986.
25. Larsen, B. L.; Rasmussen, P.; Fredenslund, A. *Ind. Eng. Chem. Res.* **1987**, *26*, 2274–2286.
26. Hector, T.; Gmehling, J. *Fluid Phase Equilib.* **2014**, *371*, 82–92.
27. Nebig, S.; Gmehling, J. *Fluid Phase Equilib.* **2011**, *302*, 220–225.
28. van der Waals, J. H. In *Over de Continuïteit van den Gas en Vloeistofoestand (On the Continuity of the Gaseous and Liquid State)*, Ph.D. Dissertation, Leiden University, 1873.
29. Lopez-Echeverry, J. S.; Reif-Acherman, S.; Araujo-Lopez, E. *Fluid Phase Equilib.* **2017**, *447*, 39–71.
30. Valderrama, J. O. *Ind. Eng. Chem. Res.* **2003**, *42*, 1603–1618.
31. Kontogeorgis, G. K.; Folas, G. K. *Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories*, 1st ed.; John Wiley & Sons: West Sussex, 2010.
32. Economou, I. G. Cubic and Generalized van der Waals Equations of State. In *Applied Thermodynamics of Fluids*; Goodwin, A. R. H., Sengers, J. V., Peters, C. J., Eds.; RSC Publishing: Cambridge, 2010; pp 53–83. Chapter 4.
33. de Hemptinne, J.-C.; Ledanois, J.-M.; Mouglin, P.; Barreau, A. *Select Thermodynamic Models for Process Simulation. A Practical Guide Using a Three Steps Methodology*, 1st edn; Editions Technip: Paris, 2012.
34. Michelsen, M. L.; Mollerup, J. M. *Thermodynamic Models: Fundamentals & Computational Aspects*, 2nd ed.; Tie-Line Publications: Holte, 2007.
35. Tassios, D. P. *Applied Chemical Engineering Thermodynamics*, 1st ed.; Springer-Verlag: Berlin Heidelberg, 1993.
36. Wei, Y. S.; Sadus, R. J. *AIChE J.* **2000**, *46*, 169–196.
37. Redlich, O.; Kwong, J. N. S. *Chem. Rev.* **1949**, *44*, 233–244.
38. Soave, G. *Chem. Eng. Sci.* **1972**, *27*, 1197–1203.
39. Peng, D.; Robinson, D. B. *Ind. Eng. Chem. Fundam.* **1976**, *15*, 59–64.
40. Chen, C.-C.; Mathias, P. M. *AIChE J.* **2002**, *48*, 194–200.
41. Ambrose, D., Correlation and Estimation of Vapour-Liquid Critical Properties. I: Critical Temperatures of Organic Compounds, NPL Technical Report Chem. 92, (National Physical Laboratory, Teddington, UK, 1978).
42. Ambrose, D., Correlation and Estimation of Vapour-Liquid Critical Properties. II: Critical Pressure and Critical Volume, NPL Technical Report Chem. 92, (National Physical Laboratory, Teddington, UK, 1979).
43. Constantinou, L.; Gani, R.; O'Connell, J. P. *Fluid Phase Equilib.* **1995**, *103*, 11–22.
44. Kontogeorgis, G. K.; Tassios, D. P. *Chem. Eng. J.* **1997**, *66*, 35–49.
45. Tsvintzelis, I.; Kontogeorgis, G. K. *Ind. Eng. Chem. Res.* **2012**, *51*, 13496–13517.
46. Tsvintzelis, I.; Ali, S.; Kontogeorgis, G. K. *Fluid Phase Equilib.* **2016**, *430*, 75–92.
47. Hudson, G. H.; McCoubrey, J. C. *Trans. Faraday Soc.* **1960**, *56*, 761–766.
48. Coutinho, J. A. P.; Kontogeorgis, G. M.; Stenby, E. H. *Fluid Phase Equilib.* **1994**, *102*, 31–60.
49. Coutinho, J. A. P.; Vlamos, P. M.; Kontogeorgis, G. M. *Ind. Eng. Chem. Res.* **2000**, *39*, 3076–3082.
50. Gao, G.; Daridon, J. L.; Saint-Guirons, H.; Xans, P.; Montel, F. *Fluid Phase Equilib.* **1992**, *74*, 85–93.
51. Kordas, A.; Magoulas, K.; Stamataki, S.; Tassios, D. *Fluid Phase Equilib.* **1995**, *112*, 33–44.
52. Avlonitis, G.; Mourikas, G.; Stamataki, S.; Tassios, D. *Fluid Phase Equilib.* **1994**, *101*, 53–68.
53. Kordas, A.; Tsoutsouras, K.; Stamataki, S.; Tassios, D. *Fluid Phase Equilib.* **1994**, *93*, 141–166.
54. Tsvintzelis, I.; Ali, S.; Kontogeorgis, G. M. *Fluid Phase Equilib.* **2015**, *397*, 1–17.
55. Jaubert, J.-N.; Mutelet, F. *Fluid Phase Equilib.* **2004**, *224*, 285–304.
56. Jaubert, J.-N.; Vitu, S.; Mutelet, F.; Corriou, J.-P. *Fluid Phase Equilib.* **2005**, *237*, 193–211.
57. Jaubert, J.-N.; Privat, R.; Mutelet, F. *AIChE J.* **2010**, *56*, 3225–3235.
58. Peneloux, A.; Abdoul, W.; Rauzy, E. *Fluid Phase Equilib.* **1989**, *47*, 115–132.
59. Abdoul, W.; Rauzy, E.; Peneloux, A. *Fluid Phase Equilib.* **1991**, *68*, 47–102.
60. Jaubert, J.-N.; Privat, R. *Fluid Phase Equilib.* **2010**, *295*, 26–37.
61. Heidemann, R. A. *Fluid Phase Equilib.* **1996**, *116* (1–2), 454–464.
62. Kontogeorgis, G. M.; Coutsikos, P. *Ind. Eng. Chem. Res.* **2012**, *51*, 4119–4142.
63. Huron, M. J.; Vidal, J. *Fluid Phase Equilib.* **1979**, *3* (40), 255–271.
64. Mollerup, J. *Fluid Phase Equilib.* **1986**, *25*, 323–327.
65. Michelsen, M. L. *Fluid Phase Equilib.* **1990**, *60*, 47–58.
66. Michelsen, M. L. *Fluid Phase Equilib.* **1990**, *60*, 213–219.
67. Dahl, S.; Michelsen, M. L. *AIChE J.* **1990**, *36* (12), 1829–1836.
68. Wong, D. S. H.; Sandler, S. I. *AIChE J.* **1992**, *38* (5), 671–680.
69. Lermite, C.; Vidal, J. *Fluid Phase Equilib.* **1992**, *72*, 111–130.
70. Soave, G. *Fluid Phase Equilib.* **1993**, *87*, 23–35.
71. Soave, G. S.; Sama, S.; Olivera, M. I. *Fluid Phase Equilib.* **1999**, *156*, 35–50.
72. Soave, G.; Gamba, S.; Pellegrini, L. A. *Fluid Phase Equilib.* **2010**, *299*, 285–293.
73. Dahl, S.; Fredenslund, A.; Rasmussen, P. *Ind. Eng. Chem. Res.* **1991**, *30*, 1936–1942.
74. Holderbaum, T.; Gmehling, J. *Fluid Phase Equilib.* **1991**, *70*, 251–265.
75. Fischer, K.; Gmehling, J. *Fluid Phase Equilib.* **1996**, *121*, 185–206.
76. Gmehling, J.; Li, J.; Fischer, K. *Fluid Phase Equilib.* **1997**, *141*, 113–127.
77. Li, J.; Fischer, K.; Gmehling, J. *Fluid Phase Equilib.* **1998**, *143*, 71–82.
78. Schmid, B.; Gmehling, J. *Fluid Phase Equilib.* **2011**, *302*, 213–219.

79. Schmid, B.; Schedemann, A.; Gmehling, J. *Ind. Eng. Chem. Res.* **2014**, *53*, 3393–3405.
80. Orbey, H.; Sandler, S. I.; Wong, D. S. H. *Fluid Phase Equilib.* **1993**, *85*, 41–54.
81. Boukouvalas, C.; Spiliotis, N.; Coutsikos, P.; Tzouvaras, N.; Tassios, D. *Fluid Phase Equilib.* **1994**, *92*, 75–106.
82. Voutsas, E.; Magoulas, K.; Tassios, D. *Ind. Eng. Chem. Res.* **2004**, *43*, 6238–6246.
83. Voutsas, E.; Louli, V.; Boukouvalas, C.; Magoulas, K.; Tassios, D. *Fluid Phase Equilib.* **2006**, *241*, 216–228.
84. Pappa, G. D.; Louli, V.; Dedousi, K.; Voutsas, E. C. *J. Supercrit. Fluids* **2011**, *58*, 321–329.
85. Wertheim, M. S. *J. Stat. Phys.* **1984**, *35*, 19–34.
86. Wertheim, M. S. *J. Stat. Phys.* **1984**, *35*, 35–47.
87. Wertheim, M. S. *J. Stat. Phys.* **1986**, *42*, 459–476.
88. Wertheim, M. S. *J. Stat. Phys.* **1986**, *42*, 477–492.
89. Jackson, G.; Chapman, W. G.; Gubbins, K. E. *Mol. Phys.* **1988**, *65* (1), 1–31.
90. Chapman, W. G.; Jackson, G.; Gubbins, K. E. *Mol. Phys.* **1988**, *65* (5), 1057–1079.
91. Huang, H. S.; Radosz, M. *Ind. Eng. Chem. Res.* **1990**, *29*, 2284–2294.
92. Huang, H. S.; Radosz, M. *Ind. Eng. Chem. Res.* **1991**, *30*, 1994–2005.
93. McCabe, C.; Galindo, A. SAFT Associating Fluids and Fluid Mixtures. In *Applied Thermodynamics of Fluids*; Goodwin, A. R. H., Sengers, J. V., Peters, C. J., Eds.; RSC Publishing: Cambridge, 2010; pp 215–279. Chapter 4.
94. Economou, I. *Ind. Eng. Chem. Res.* **2002**, *41*, 953–962.
95. Muller, E. A.; Gubbins, K. E. *Ind. Eng. Chem. Res.* **2001**, *40*, 2193–2211.
96. Tan, S. P.; Adidharma, H.; Radosz, M. *Ind. Eng. Chem. Res.* **2008**, *47*, 8063–8082.
97. Polishuk, I.; Mulero, A. *Rev. Chem. Eng.* **2011**, *27*, 241–251.
98. Fu, Y.-H.; Sandler, S. I. *Ind. Eng. Chem. Res.* **1995**, *34*, 1897–1909.
99. Kraska, T.; Gubbins, K. E. *Ind. Eng. Chem. Res.* **1996**, *35*, 4727–4737.
100. Kraska, T.; Gubbins, K. E. *Ind. Eng. Chem. Res.* **1996**, *35*, 4738–4746.
101. Blas, F. J.; Vega, L. F. *Mol. Phys.* **1997**, *92*, 135–150.
102. Gil-Vilegas, A.; Galindo, A.; Whitehead, P. J.; Mills, S. J.; Jackson, G.; Burgess, A. N. *J. Chem. Phys.* **1997**, *106*, 4168–4186.
103. Galindo, A.; Davies, L. A.; Gil-Vilegas, A.; Jackson, G. *Mol. Phys.* **1998**, *93*, 241–252.
104. Lafitte, T.; Bessieres, D.; Pineiro, M. M.; Daridon, J. L. *J. Chem. Phys.* **2006**, *124*, 024509, 1-16.
105. Gross, J.; Sadowski, G. *Ind. Eng. Chem. Res.* **2001**, *40*, 1244–1260.
106. Gross, J.; Sadowski, G. *Ind. Eng. Chem. Res.* **2002**, *41*, 5510–5515.
107. von Solms, N.; Michelsen, M. L.; Kontogeorgis, G. M. *Ind. Eng. Chem. Res.* **2003**, *42*, 1098–1105.
108. Karakatsani, E. K.; Kontogeorgis, G. M.; Economou, I. G. *Ind. Eng. Chem. Res.* **2006**, *45*, 6063–6074.
109. Polishuk, I. *Ind. Eng. Chem. Res.* **2011**, *50*, 4183–4198.
110. Liang, X.; Maribo-Mogensen, B.; Thomsen, K.; Yan, W.; Kontogeorgis, G. K. *Ind. Eng. Chem. Res.* **2012**, *51*, 14903–14914.
111. Breil, M. P.; Kontogeorgis, G. K. *Ind. Eng. Chem. Res.* **2009**, *48*, 5472–5480.
112. Liang, X.; Maribo-Mogensen, B.; Tsvintzelis, I.; Kontogeorgis, G. K. *Fluid Phase Equilib.* **2016**, *407*, 2–6.
113. Albers, K.; Sadowski, G. *Fluid Phase Equilib.* **2012**, *326*, 21–30.
114. Wolbach, J. P.; Sandler, S. I. *Ind. Eng. Chem. Res.* **1997**, *36*, 4041–4049.
115. Wolbach, J. P.; Sandler, S. I. *Int. J. Thermophys.* **1997**, *18*, 1001–1016.
116. Wolbach, J. P.; Sandler, S. I. *Ind. Eng. Chem. Res.* **1998**, *37*, 2917–2928.
117. Singh, M.; Leonhard, K.; Lucas, K. *Fluid Phase Equilib.* **2007**, *258*, 16–28.
118. Leonhard, K.; van Nhu, N.; Lucas, K. *Fluid Phase Equilib.* **2007**, *258*, 41–50.
119. Leonhard, K.; van Nhu, N.; Lucas, K. *J. Phys. Chem. C* **2007**, *111*, 15533–15543.
120. van Nhu, N.; Singh, M.; Leonhard, K. *J. Phys. Chem. B* **2008**, *112*, 5693–5701.
121. von Muller, A.; Leonhard, K. *Fluid Phase Equilib.* **2013**, *356*, 96–101.
122. Lucia, A.; Octavio, L. M.; Visco, D. P., Jr. *Comput. Chem. Eng.* **2009**, *33*, 531–533.
123. Padiuszynski, K.; Domanska, U. *Ind. Eng. Chem. Res.* **2012**, *51*, 12967–12983.
124. Papaioannou, V.; Adjiman, C. S.; Jackson, G.; Galindo, A. Group Contribution Methodologies for the Prediction of Thermodynamic Properties and Phase Behavior in Mixtures. In *Process Systems Engineering: Molecular Systems Engineering*; Adjiman, C. S., Galindo, A., Eds.; Wiley-VCH: Weinheim, 2011; vol. 6; pp 135–172. Chapter 4.
125. Papaioannou, V.; Lafitte, T.; Avendaño, C.; Adjiman, C. S.; Jackson, G.; Müller, E. A.; Galindo, A. *J. Chem. Phys.* **2014**, *140*, 054107–054116.
126. Lora, M.; Rindfleisch, F.; McHugh, M. A. *J. Appl. Polym. Sci.* **1999**, *73*, 1979–1991.
127. Tamouza, S.; Passarello, J.-P.; Tobaly, P.; de Hemptinne, J.-C. *Fluid Phase Equilib.* **2004**, *222–223*, 67–76.
128. Tamouza, S.; Passarello, J. P.; Tobaly, P.; de Hemptinne, J.-C. *Fluid Phase Equilib.* **2005**, *228–229*, 409–419.
129. NguyenHuynh, D. *Fluid Phase Equilib.* **2016**, *430*, 33–46.
130. Vijande, J.; Piñeiro, M. M.; Legido, J. L. *Ind. Eng. Chem. Res.* **2010**, *49*, 9394–9406.
131. Tihic, A.; Kontogeorgis, G. M.; Von Solms, N.; Michelsen, M. L.; Constantinou, L. *Ind. Eng. Chem. Res.* **2008**, *47*, 5092–5101.
132. Tihic, A.; von Solms, N.; Michelsen, M. L.; Kontogeorgis, G. M. *Fluid Phase Equilib.* **2009**, *281*, 60–69.
133. Peng, Y.; Goff, K. D.; dos Ramos, M. C.; McCabe, C. *Fluid Phase Equilib.* **2009**, *277*, 131–144.
134. Lymeriadis, A.; Adjiman, C. S.; Galindo, A.; Jackson, G. A. *J. Chem. Phys.* **2007**, *127*, 234903.
135. Lymeriadis, A.; Adjiman, C. S.; Jackson, G.; Galindo, A. A. *Fluid Phase Equilib.* **2008**, *274*, 85–104.
136. Klamt, A.; Schuurmann, G. J. G. *J. Perkin Trans.* **1993**, *2*, 799–805.
137. Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224–2235.
138. Klamt, A.; Jonas, V.; Burger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.
139. Klamt, A. *COSMO-RS, From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, 1st ed.; Elsevier: Amsterdam, 2004.
140. Lin, S.-T.; Sandler, S. I. *Ind. Eng. Chem. Res.* **2001**, *41*, 899–913.
141. Xue, Z.; Mu, T.; Gmehling, J. *Ind. Eng. Chem. Res.* **2012**, *51*, 11809–11817.
142. Panayiotou, C.; Tsvintzelis, I.; Aslanidou, D.; Hatzimanikatis, V. *J. Chem. Thermodyn.* **2015**, *90*, 294–309.
143. Costa, C. T. O. G.; Tavares, F. W.; Secchi, A. R. *Fluid Phase Equilib.* **2016**, *409*, 472–481.
144. COSMObase, *COSMObase Ver. C30_1401*. COSMOlogic GmbH & Co. K.G.: Leverkusen, 2014.
145. Mullins, E.; Liu, Y. A.; Ghaderi, A.; Fast, S. *Ind. Eng. Chem. Res.* **2008**, *47*, 1707–1725.
146. Mu, T.; Rarey, J.; Gmehling, J. *AIChE J.* **2009**, *55*, 3298–3300.
147. Mu, T.; Rarey, J.; Gmehling, J. *AIChE J.* **2007**, *53*, 3231–3240.
148. Grensemann, H.; Gmehling, J. *Ind. Eng. Chem. Res.* **2005**, *44*, 1610–1624.
149. Austin, N. D.; Sahinidis, N. V.; Trahan, D. W. *Chem. Eng. Sci.* **2017**, *159*, 93–105.

150. Papadopoulos, A. I.; Linke, P.; Seferlis, P. Integrated Multiobjective Molecular and Process Design: Operational and Computational Frontiers. In *Tools for Chemical Product Design: From Consumer Products to Biomedicine*; Martin, M., Eden, M., Chemmangattuvalappil, N., Eds.; Computer Aided Chemical Engineering Elsevier: Amsterdam, 2016; vol. 39.; pp 269–313. Chapter 11.
151. Gani, R. *Comput. Chem. Eng.* **2004**, *28* (12), 2441–2457.
152. Pereira, F. E.; Keskes, E.; Galindo, A.; Jackson, G.; Adjiman, C. S. *Comput. Chem. Eng.* **2011**, *35* (3), 474–491.
153. Heintz, J.; Belaud, J. P.; Pandya, N.; Dos Santos, M. T.; Gerbaud, V. *Comput. Chem. Eng.* **2014**, *71*, 362–376.
154. Austin, N. D. Ph.D. Thesis, Carnegie Mellon University, USA, 2017.
155. Struebing, H.; Ganase, Z.; Karamertzanis, P. G.; Sioukrou, E.; Haycock, P.; Piccione, P. M.; Armstrong, A.; Galindo, A.; Adjiman, C. S. *Nat. Chem.* **2013**, *5* (11), 952.
156. Marcoulaki, E.C., Ph.D. Thesis, Institute of Technology, University of Manchester, UK, 1998.
157. Adjiman, C. S. Optimal Solvent Design Approaches. In *Encyclopedia of Optimization*; Floudas, C. A., Pardalos, P. M., Eds.; Springer: Boston, MA, 2008.; pp 2750–2757.
158. Samudra, A. P.; Sahinidis, N. V. *AIChE J.* **2013**, *59* (10), 3686–3701.
159. Sioukrou, E. Ph.D. Thesis, Imperial College London, UK, 2014.
160. Odele, O.; Macchietto, S. *Fluid Phase Equilib.* **1993**, *82*, 47–54.
161. Papadopoulos, A. I.; Linke, P. *AIChE J.* **2006**, *52* (3), 1057–1070.
162. Papadopoulos, A. I.; Seferlis, P.; Linke, P. *Chem. Eng. Sci.* **2017**, *159*, 154–176.
163. Papadopoulos, A. I.; Linke, P. *Chem. Eng. Res. Design* **2005**, *83* (6), 674–678.
164. Papadopoulos, A. I.; Zargiannis, T.; Seferlis, P. Computer-Aided Molecular Design of Pure and Mixed CO₂ Capture Solvents. In *Process Systems and Materials for CO₂ Capture: Modelling, Design, Control and Integration*; Papadopoulos, A. I., Seferlis, P., Eds.; John Wiley & Sons: Hoboken, NJ, 2017; pp 173–202. Chapter 7.
165. Zargiannis, T.; Papadopoulos, A. I.; Seferlis, P. *J. Clean. Prod.* **2016**, *136*, 159–175.
166. Zargiannis, T.; Papadopoulos, A. I.; Seferlis, P.; Linke, P. *Comput.-Aided Chem. Eng.* **2017**, *40*, 2443–2448.
167. Kim, K.; Diwekar, U. *Ind. Eng. Chem. Res.* **2002**, *41*, 1285–1296.
168. Maranas, C. D. *AIChE J.* **1997**, *43* (5), 1250–1264.
169. Papadopoulos, A. I.; Seferlis, P. *Comput.-Aided Chem. Eng.* **2009**, *26*, 177–181.
170. Adjiman, C. S.; Clarke, A. J.; Cooper, G.; Taylor, P. C. *Chem. Commun.* **2008**, *24*, 2806–2808.
171. Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11* (6), 700–733.
172. Lehmann, A.; Maranas, C. D. *Ind. Eng. Chem. Res.* **2004**, *43* (13), 3419–3432.
173. Macchietto, S.; Odele, O.; Omatsone, O. *Chem. Eng. Res. Design* **1990**, *68*, 429–433.
174. Naser, S. F.; Fournier, R. L. *Comput. Chem. Eng.* **1991**, *15* (6), 397–414.
175. Bardow, A.; Steur, K.; Gross, J. *Ind. Eng. Chem. Res.* **2010**, *49* (6), 2834–2840.
176. Roskosch, D.; Atakan, B. *Energy* **2015**, *81*, 202–212.
177. Linke, P.; Papadopoulos, A. I.; Seferlis, P. *Energies* **2015**, *8* (6), 4755–4801.
178. Linke, P.; Kokossis, A. *Comput. Chem. Eng.* **2003**, *27* (5), 733–758.
179. Gopinath, S. Ph.D. Thesis, Imperial College London, UK, 2017.
180. Brignole, E. A.; Cismondi, M. Molecular Design–Generation & Test Methods. In *Computer Aided Chemical Engineering*; Achenie, L. E. K., Gani, R., Venkatasubramanian, V., Eds.; Elsevier: Amsterdam, 2003; vol. 12; pp 23–41. Chapter 2.
181. Brignole, E. A.; Bottini, S. B.; Gani, R. *Fluid Phase Equilib.* **1986**, *29*, 125–132.
182. Pretel, E. J.; Lopez, P. A.; Bottini, S. B.; Brignole, E. A. *AIChE J.* **1994**, *40*, 1349–1360.
183. Cismondi, M.; Brignole, E. A. *Ind. Eng. Chem. Res.* **2004**, *43* (3), 784–790.
184. Joback, K.G. Ph.D. Thesis, Massachusetts Institute of Technology, MA, US, 1989.
185. Joback, K. G.; Stephanopoulos, G. In *Designing Molecules Possessing Desired Physical Property Values*, Proceedings of FOCAPD '89, Snowmass, CO, 1989; p 363.
186. Derringer, G. C.; Markham, R. L. *J. Appl. Polym. Sci.* **1985**, *30* (12), 4609–4617.
187. Gani, R.; Nielsen, B.; Fredenslund, A. *AIChE J.* **1991**, *37* (9), 1318–1332.
188. Constantinou, L.; Bagherpour, K.; Gani, R.; Klein, J. A.; Wu, D. T. *Comput. Chem. Eng.* **1996**, *20* (6–7), 685–702.
189. Harper, P. M.; Gani, R.; Kolar, P.; Ishikawa, T. *Fluid Phase Equilib.* **1999**, *158*, 337–347.
190. Harper, P. M.; Gani, R. *Comput. Chem. Eng.* **2000**, *24* (2–7), 677–683.
191. Conte, E.; Gani, R.; Ng, K. M. *AIChE J.* **2011**, *57* (9), 2431–2449.
192. Conte, E.; Gani, R.; Malik, T. I. *Fluid Phase Equilib.* **2011**, *302* (1–2), 294–304.
193. Mattei, M.; Kontogeorgis, G. M.; Gani, R. *Fluid Phase Equilib.* **2014**, *362*, 288–299.
194. Yang, X. G.; Song, H. H. *Chem. Eng. Technol.* **2006**, *29* (1), 33–43.
195. Adjiman, C. S.; Schweiger, C. A.; Floudas, C. A. Mixed-Integer Nonlinear Optimization in Process Synthesis. In *Handbook of Combinatorial Optimization*; Pardalos, P. M., Du, D.-Z., Graham, R., Eds.; Springer: Boston, MA, 1998; pp 1–76.
196. Rios, L. M.; Sahinidis, N. V. *J. Glob. Optim.* **2013**, *56* (3), 1247–1293.
197. Floudas, C. A.; Gounaris, C. E. *J. Glob. Optim.* **2009**, *45* (1), 3–38.
198. Boukouvala, F.; Misener, R.; Floudas, C. A. *Eur. J. Oper. Res.* **2016**, *252* (3), 701–727.
199. Little, J. D.; Murty, K. G.; Sweeney, D. W.; Karel, C. *Oper. Res.* **1963**, *11* (6), 972–989.
200. Adjiman, C. S.; Androulakis, J. P.; Floudas, C. A. *AIChE J.* **2000**, *46* (9), 1769–1797.
201. Ryoo, H. S.; Sahinidis, N. V. *Comput. Chem. Eng.* **1995**, *19* (5), 551–566.
202. Geoffrion, A. M. *J. Optim. Theory Appl.* **1972**, *10* (4), 237–260.
203. Buxton, A.; Livingston, A. G.; Pistikopoulos, E. N. *AIChE J.* **1999**, *45*, 817–843.
204. Viswanathan, J.; Grossmann, I. E. *Comput. Chem. Eng.* **1990**, *14* (7), 769–782.
205. Kocis, G. R.; Grossmann, I. E. *Ind. Eng. Chem. Res.* **1987**, *26* (9), 1869–1880.
206. Raman, R.; Grossmann, I. E. *Comput. Chem. Eng.* **1994**, *18* (7), 563–578.
207. Jonuzaj, S. Ph.D. Thesis, Imperial College London, UK, 2017.
208. Nemhauser, G. L.; Wolsey, L. A. Integer and Combinatorial Optimization. In *Interscience Series in Discrete Mathematics and Optimization*, John Wiley & Sons: Hoboken, NJ, 1986.
209. Grossmann, I. E.; Lee, S. *Comput. Optim. Appl.* **2003**, *26* (1), 83–100.
210. Vaidyanathan, R.; El-Halwagi, M. *Comput. Chem. Eng.* **1994**, *18* (10), 889–897.
211. Schittkowski, K. *Ann. Oper. Res.* **1986**, *5*, 485–500.
212. Klein, J. A.; Wu, D. T.; Gani, R. *Comput. Chem. Eng.* **1992**, *16*, S229–S236.
213. Lasdon, L. S.; Waren, A. D.; Sarkar, S.; Palacios, F. *ACM Sigmap Bulletin* **1979**, (27), 9–15.
214. Gani, R.; Fredenslund, A. *Fluid Phase Equilib.* **1993**, *82*, 39–46.
215. Maranas, C. D. *Ind. Eng. Chem. Res.* **1996**, *35* (10), 3403–3414.
216. Camarda, K. V.; Maranas, C. D. *Ind. Eng. Chem. Res.* **1999**, *38* (5), 1884–1892.

217. Duvedi, A. P.; Achenie, L. E. K. *Chem. Eng. Sci.* **1996**, *51*, 3727–3739.
218. Duvedi, A.; Achenie, L. E. K. *Comput. Chem. Eng.* **1997**, *21*, 915–923.
219. Churi, N.; Achenie, L. E. K. *Comput. Chem. Eng.* **1997**, *21*, S349–S354.
220. Wang, Y.; Achenie, L. E. *Fluid Phase Equilib.* **2002**, *201* (1), 1–18.
221. Sinha, M.; Achenie, L. E. K.; Ostrovsky, G. M. *Comput. Chem. Eng.* **1999**, *23*, 1381–1394.
222. Ostrovsky, G. M.; Achenie, L. E.; Sinha, M. *Comput. Chem.* **2002**, *26* (6), 645–660.
223. Friedler, F.; Fan, L. T.; Kalotai, L.; Dallos, A. *Comput. Chem. Eng.* **1998**, *22* (6), 809–817.
224. Achenie, L. E. K.; Sinha, M. *Reliab. Comput.* **2003**, *9*, 317–338.
225. Sinha, M.; Achenie, L. E. K.; Gani, R. *Ind. Eng. Chem. Res.* **2003**, *42*, 516–527.
226. Karunanithi, A. T.; Achenie, L. E. K.; Gani, R. *Ind. Eng. Chem. Res.* **2005**, *44*, 4785–4797.
227. Karunanithi, A. T.; Achenie, L. E. K.; Gani, R. *Chem. Eng. Sci.* **2006**, *61*, 1247–1260.
228. Karunanithi, A. T.; Acquah, C.; Achenie, L. E. K.; Sithambaram, S.; Suib, S. L. *Comput. Chem. Eng.* **2009**, *33*, 1014–1021.
229. Vaidyanathan, R.; El-Halwagi, M. *Ind. Eng. Chem. Res.* **1996**, *35* (2), 627–634.
230. Liebman, J.; Lasdon, L.; Schrage, L.; Waren, A. *Modeling and Optimization With GINO*. The Scientific Press: South San Francisco, CA, 1986.
231. Hostrup, M.; Harper, P. M.; Gani, R. *Comput. Chem. Eng.* **1999**, *23* (10), 1395–1414.
232. Zhang, L.; Cignitti, S.; Gani, R. *Comput. Chem. Eng.* **2015**, *78*, 79–84.
233. Giovanoglou, A.; Barlatier, J.; Adjiman, C. S.; Pistikopoulos, E. N.; Cordier, J. L. *AIChE J.* **2003**, *49*, 3095–3109.
234. Sheldon, T. J.; Folic, M.; Adjiman, C. S. *Ind. Eng. Chem. Res.* **2006**, *45*, 1128–1140.
235. Gill, P. E.; Murray, W.; Saunders, M. A. User's Guide for Snopt Version 6, A Fortran Package for Large-Scale Nonlinear Programming; <http://www.sbsi-sol-optimize.com>, Stanford Business Software Inc.: Palo Alto, CA, 2002.
236. Krige, D. G. *J. South. Afr. Inst. Min. Metall.* **1951**, *52* (6), 119–139.
237. Burger, J.; Papaioannou, V.; Gopinath, S.; Jackson, G.; Galindo, A.; Adjiman, C. S. *AIChE J.* **2015**, *61* (10), 3249–3269.
238. Jonuzaj, S.; Akula, P. T.; Kleniati, P. M.; Adjiman, C. S. *AIChE J.* **2016**, *62* (5), 1616–1633.
239. Jonuzaj, S.; Gupta, A.; Adjiman, C. S. *Comput. Chem. Eng.* **2018**, <https://doi.org/10.1016/j.compchemeng.2018.1.16>.
240. Jonuzaj, S.; Adjiman, C. S. *Chem. Eng. Sci.* **2017**, *159*, 106–130.
241. Gopinath, S.; Jackson, G.; Galindo, A.; Adjiman, C. S. *AIChE J.* **2016**, *62*, 3484–3504.
242. Sahinidis, N. V.; Tawarmalani, M.; Yu, M. *AIChE J.* **2003**, *49*, 1761–1775.
243. Samudra, A.; Sahinidis, N. V. *Ind. Eng. Chem. Res.* **2013**, *52* (25), 8518–8526.
244. Cheng, H. C.; Wang, F. S. *Chem. Eng. J.* **2010**, *162* (2), 809–820.
245. Liao, C. T.; Tzeng, W. J.; Wang, F. S. *J. Chin. Inst. Chem. Eng.* **2001**, *32* (6), 491–502.
246. Exler, O.; Schittkowski, K. *Opt. Lett.* **2007**, *1* (3), 269–280.
247. Blum, C.; Roli, A. *ACM Comput. Surv.* **2003**, *35* (3), 268–308.
248. Pardalos, P. M.; Romeijn, H. E.; Tuy, H. *J. Comput. Appl. Math.* **2000**, *124*, 209–228.
249. Papadopoulos, A. I.; Linke, P. *Comput. Chem. Eng.* **2004**, *28* (11), 2391–2406.
250. Holland, J. H. *SIAM J. Comput.* **1973**, *2* (2), 88–105.
251. Patkar, P. R.; Venkatasubramanian, V. Genetic Algorithms Based CAMD. In *Computer-Aided Chemical Engineering*; Achenie, L. E. K., Gani, R., Venkatasubramanian, V., Eds.; Vol. 12; Elsevier: Amsterdam, 2003; pp 95–128. Chapter 5.
252. Zhou, T.; Wang, J.; McBride, K.; Sundmacher, K. *AIChE J.* **2016**, *62* (9), 3238–3249.
253. Venkatasubramanian, V.; Chan, K.; Caruthers, J. M. *Comput. Chem. Eng.* **1994**, *18* (9), 833–844.
254. Van Dyk, B.; Nieuwoudt, I. *Ind. Eng. Chem. Res.* **2000**, *39* (5), 1423–1429.
255. Xu, W.; Diwekar, U. M. *Ind. Eng. Chem. Res.* **2005**, *44* (18), 7138–7146.
256. Xu, W.; Diwekar, U. M. *Int. J. Environ. Pollut.* **2007**, *29* (1–3), 70–89.
257. Herring, R. H.; Eden, M. R. *Comput. Chem. Eng.* **2015**, *83*, 267–277.
258. Liu, X.; Huang, Y.; Zhao, Y.; Gani, R.; Zhang, X.; Zhang, S. *Ind. Eng. Chem. Res.* **2016**, *55* (20), 5931–5944.
259. Scheffczyk, J.; Fleitmann, L.; Schwarz, A.; Lampe, M.; Bardow, A.; Leonhard, K. *Chem. Eng. Sci.* **2017**, *159*, 84–92.
260. Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31–36.
261. Douguet, D.; Thoreau, E.; Grassy, G. *J. Comput. Aided Mol. Des.* **2000**, *14* (5), 449–466.
262. Douguet, D.; Munier-Lehmann, H.; Labesse, G.; Pochet, S. *J. Med. Chem.* **2005**, *48* (7), 2457–2468.
263. Devi, R. V.; Sathya, S. S.; Coumar, M. S. *Appl. Soft Comput.* **2015**, *27*, 543–552.
264. Venkatasubramanian, V.; Chan, K.; Caruthers, J. M. *J. Chem. Inf. Comput. Sci.* **1995**, *35* (2), 188–195.
265. Kirkpatrick, S. C. D.; Gelatt, J.; Vecchi, M. P. *Science* **1983**, *220*, 671.
266. Metropolis, N. A.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
267. Marcoulaki, E. C.; Kokossis, A. C. *Comput. Chem. Eng.* **1998**, *22*, S11–S18.
268. Marcoulaki, E. C.; Kokossis, A. C. *Chem. Eng. Sci.* **2000**, *55*, 2529–2546.
269. Marcoulaki, E. C.; Kokossis, A. C. *Chem. Eng. Sci.* **2000**, *55* (13), 2547–2561.
270. Marcoulaki, E. C.; Kokossis, A. C.; Batzias, F. A. *Comput. Chem. Eng.* **2000**, *24* (2–7), 705–710.
271. Ourique, J. E.; Telles, A. S. *Comput. Chem. Eng.* **1998**, *22*, S615–S618.
272. Kim, K.; Diwekar, U. *Ind. Eng. Chem. Res.* **2002**, *41*, 1276–1284.
273. Kim, K. J.; Diwekar, U. M. *Ind. Eng. Chem. Res.* **2002**, *41* (18), 4479.
274. Song, J.; Song, H. H. *Chem. Eng. Technol.* **2008**, *31* (2), 177–187.
275. Song, J.; Qi, T.; Qu, J. K.; Zhang, X. G.; Shen, W. F. *Adv. Mat. Res.* **2011**, *233–235*, 2938–2944.
276. Ashley, V.; Linke, P. *Chem. Eng. Res. Design* **2004**, *82* (8), 952–960.
277. Chavali, S.; Lin, B.; Miller, D. C.; Camarda, K. V. *Comput. Chem. Eng.* **2004**, *28* (5), 605–611.
278. Lin, B.; Chavali, S.; Camarda, K.; Miller, D. C. *Comput. Chem. Eng.* **2005**, *29* (2), 337–347.
279. Eslick, J. C.; Ye, Q.; Park, J.; Topp, E. M.; Spencer, P.; Camarda, K. V. *Comput. Chem. Eng.* **2009**, *33* (5), 954–963.
280. McLeese, S. E.; Eslick, J. C.; Hoffmann, N. J.; Scurto, A. M.; Camarda, K. V. *Comput. Chem. Eng.* **2010**, *34* (9), 1476–1480.
281. Abedin, F.; Roughton, B.; Ye, Q.; Spencer, P.; Camarda, K. *Chem. Eng. Sci.* **2017**, *159*, 131–139.
282. Gebreslassie, B. H.; Diwekar, U. M. *Comput. Chem. Eng.* **2015**, *78*, 1–9.
283. Diwekar, U. M.; Gebreslassie, B. H. *Int. J. Swarm Intel. Evol. Comput.* **2016**, *5*, 131.
284. Hammersley, J. M. *Ann. N. Y. Acad. Sci.* **1960**, *86* (1), 844–874.
285. Gebreslassie, B. H.; Diwekar, U. M. *Chem. Eng. Sci.* **2017**, *159*, 194–206.
286. Mukherjee, R.; Gebreslassie, B.; Diwekar, U. M. *Clean Techn. Environ. Policy* **2017**, *19* (2), 483–499.
287. Doshi, R. K.; Mukherjee, R.; Diwekar, U. M. *ACS Sustain. Chem. Eng.* **2018**, *6* (2), 2603–2611.

288. Benavides, P. T.; Gebreslassie, B. H.; Diwekar, U. M. *Chem. Eng. Sci.* **2015**, *137*, 977–985.
289. Benavides, P. T.; Diwekar, U. *Chem. Eng. Sci.* **2015**, *137*, 964–976.
290. Shelley, M. D.; El-Halwagi, M. M. *Comput. Chem. Eng.* **2000**, *24* (9–10), 2081–2091.
291. El-Halwagi, M. M.; Glasgow, I. M.; Qin, X.; Eden, M. R. *AIChE J.* **2004**, *50* (8), 1854–1869.
292. Eden, M. R.; Jorgensen, S. B.; Gani, R.; El-Halwagi, M. M. *Chem. Eng. Process.* **2004**, *43*, 595.
293. Eljack, F. T.; Eden, M. R. *Comput. Chem. Eng.* **2008**, *32* (12), 3002–3010.
294. Eljack, F.T. Ph.D. Thesis, Auburn University, Alabama, US, 2007.
295. Chemmangattualappil, N. G.; Eljack, F. T.; Solvason, C. C.; Eden, M. R. *Comput. Chem. Eng.* **2009**, *33* (3), 636–643.
296. Chemmangattualappil, N. G.; Solvason, C. C.; Bommareddy, S.; Eden, M. R. *Comput. Chem. Eng.* **2010**, *34* (12), 2062–2071.
297. Chemmangattualappil, N. G.; Eden, M. R. *Ind. Eng. Chem. Res.* **2013**, *52* (22), 7090–7103.
298. Hada, S.; Solvason, C. C.; Eden, M. R. *Front. Energy Res.* **2014**, *2*, 20.
299. Hada, S.; Herring, R. H.; Davis, S. E.; Eden, M. R. *Comput. Chem. Eng.* **2015**, *81*, 310–322.
300. Hada, S.; Herring, R. H., III; Eden, M. R. *Comput. Chem. Eng.* **2017**, *107*, 26–36.
301. Woo, W. Q.; Ng, L. Y.; Sivaneswaran, U.; Chemmangattualappil, N. G. *J. Phys. Sci.* **2017**, *28*, 1.
302. Adjiman, C. S.; Dallwig, S.; Floudas, C. A.; Neumaier, A. *Comput. Chem. Eng.* **1998**, *22* (9), 1137–1158.
303. Jaksland, C.; Gani, R. *Comput. Chem. Eng.* **1996**, *20*, S151–S156.
304. Hamad, A. A.; El-Halwagi, M. M. *Trans. IChemE* **1998**, *76*, 376.
305. Pistikopoulos, E. N.; Stefanis, S. K. *Comput. Chem. Eng.* **1998**, *22* (6), 717–733.
306. Cheng, H. C.; Wang, F. S. *Chem. Eng. Sci.* **2007**, *62* (16), 4316.
307. Cheng, H. C.; Wang, F. S. *Biochem. Eng. J.* **2008**, *41* (3), 258.
308. Martin, M.; Martinez, A. *Chem. Eng. Res. Design* **2013**, *91* (5), 795–809.
309. Fung, K. Y.; Ng, K. M.; Zhang, L.; Gani, R. *Comput. Chem. Eng.* **2016**, *91*, 15–27.
310. Roughton, B. C.; Christian, B.; White, J.; Camarda, K. V.; Gani, R. *Comput. Chem. Eng.* **2012**, *42*, 248–262.
311. Palma-Flores, O.; Flores-Tlacuahuac, A.; Canseco-Melchor, G. *Comput. Chem. Eng.* **2015**, *72*, 334–349.
312. Molina-Thierry, D. P.; Flores-Tlacuahuac, A. *Ind. Eng. Chem. Res.* **2015**, *54*, 3367–3383.
313. Santos-Rodríguez, M. M.; Flores-Tlacuahuac, A.; Zavala, V. M. *Appl. Energy* **2017**, *198*, 145–159.
314. White, M. T.; Oyewunmi, O. A.; Haslam, A. J.; Markides, C. N. *Energy Convers. Manag.* **2017**, *150*, 851–869.
315. Cignitti, S.; Andreasen, J. G.; Haglind, F.; Woodley, J. M.; Abildskov, J. *Appl. Energy* **2017**, *203*, 442–453.
316. Drud, A. S. *ORSA J. Comput.* **1994**, *6* (2), 207–216.
317. Cignitti, S.; Mansouri, S. S.; Woodley, J. M.; Abildskov, J. *Ind. Eng. Chem. Res.* **2018**, *57* (2), 677–688.
318. Stijepovic, M. Z.; Papadopoulos, A. I.; Linke, P.; Stijepovic, V.; Grujic, A. S.; Kijevanin, M.; Seferlis, P. *J. Clean. Prod.* **2017**, *142*, 1950–1970.
319. Thierry, D. M.; Flores-Tlacuahuac, A.; Grossmann, I. E. *Comput. Chem. Eng.* **2016**, *89*, 106–126.
320. Linke, P.; Kokossis, A. C. In *Simultaneous Synthesis and Design of Novel Chemicals and Process Flowsheets*, Proceedings of European Symposium on Computer Aided Process Engineering, Delft, The Netherlands, 2002; Grievink, J., van Schijnde, K., Eds.; Elsevier: Amsterdam, 2012.
321. Papadopoulos, A. I.; Linke, P. *Chem. Eng. Sci.* **2006**, *61* (19), 6316–6336.
322. Papadopoulos, A. I.; Linke, P. *Chem. Eng. Process.* **2009**, *48* (5), 1047–1060.
323. Papadopoulos, A. I.; Stijepovic, M.; Linke, P. *Appl. Therm. Eng.* **2010**, *30* (6–7), 760–769.
324. Papadopoulos, A. I.; Stijepovic, M.; Linke, P.; Seferlis, P.; Voutetakis, S. *Ind. Eng. Chem. Res.* **2013**, *52* (34), 12116–12133.
325. Papadokostas, S.; Badr, S.; Hugerbuhler, K.; Papadopoulos, A. I.; Damartzis, T.; Seferlis, P.; Forte, E.; Chremos, A.; Galindo, A.; Adjiman, C. S.; Jackson, G. Towards Sustainable Solvent-Based Post Combustion CO₂ Capture: From Molecules to Conceptual Flowsheet Design. In *Sustainability of Products, Processes and Supply Chains: Theory and Applications, Computer Aided Chemical Engineering*; You, F., Ed.; Vol. 36; Elsevier: Amsterdam, 2015; pp 279–305. Chapter 11.
326. Papadopoulos, A. I.; Linke, P. *Comput. Chem. Eng.* **2009**, *33* (1), 72–87.
327. HellasGrid. HellasGrid webpage, 2018, www.hellasgrid.gr.
328. Kim, K. J.; Diwekar, U. M.; Tomazi, K. G. *Chem. Eng. Commun.* **2004**, *191* (12), 1606–1633.
329. Qadir, A.; Chiesa, M.; Abbas, A. *Int. J. Greenhouse Gas Control* **2014**, *30*, 179–187.
330. Jog, P. K.; Sauer, S. G.; Blaesing, J.; Chapman, W. G. *Ind. Eng. Chem. Res.* **2001**, *40* (21), 4641–4648.
331. MATLAB. MATLAB webpage, 2018, www.mathworks.com/products/matlab.html.
332. ASPEN Plus webpage, 2018, www.aspentech.com.
333. Zhou, T.; Zhou, Y.; Sundmacher, K. *Chem. Eng. Sci.* **2017**, *159*, 207–216.
334. Stavrou, M.; Lampe, M.; Bardow, A.; Gross, J. *Ind. Eng. Chem. Res.* **2014**, *53* (46), 18029–18041.
335. Lampe, M.; Stavrou, M.; Bucker, H. M.; Gross, J.; Bardow, A. *Ind. Eng. Chem. Res.* **2014**, *53* (21), 8821–8830.
336. Lampe, M.; Stavrou, M.; Schilling, J.; Sauer, E.; Gross, J.; Bardow, A. *Comput. Chem. Eng.* **2015**, *81*, 278–287.
337. Sauer, E.; Stavrou, M.; Gross, J. *Ind. Eng. Chem. Res.* **2014**, *53* (38), 14854–14864.
338. Schilling, J.; Lampe, M.; Gross, J.; Bardow, A. *Chem. Eng. Sci.* **2017**, *159*, 217–230.
339. Schilling, J.; Tillmanns, D.; Lampe, M.; Hopp, M.; Gross, J.; Bardow, A. *Mol. Syst. Design Eng.* **2017**, *2* (3), 301–320.
340. Wang, J.; Lakerveld, R. *AIChE J.* **2017**, <https://doi.org/10.1002/aic.15998>.
341. Roskosch, D.; Atakan, B. In *Reverse Engineering of Fluid Selection for ORCs Using Cubic Equations of State*, Proceedings of 3rd International Seminar on ORC Power Systems, October 12–14, Brussels, Belgium, 2015; Lemort, V., Quoilin, S., De Paepe, M., van den Broek, M., Eds.; 2015.
342. Frutiger, J. Ph.D. Thesis, Technical University of Denmark, DK, 2017.
343. Mac Dowell, N.; Galindo, A.; Jackson, G.; Adjiman, C. S. *Comput.-Aided Chem. Eng.* **2010**, *28*, 1231–1236.
344. Ulas, S.; Diwekar, U. M. *Chem. Eng. Sci.* **2006**, *61* (6), 2001–2009.
345. Mac Dowell, N., Ph.D. Thesis, Imperial College London, UK, 2010.
346. Mac Dowell, N.; Samsatli, N. J.; Shah, N. *Int. J. Greenhouse Gas Control* **2013**, *12*, 247–258.
347. Kazantzi, V.; Qin, X.; El-Halwagi, M.; Eljack, F. T.; Eden, M. *Ind. Eng. Chem. Res.* **2007**, *46* (10), 3400.
348. Eljack, F. T.; Eden, M. R.; Kazantzi, V.; Qin, X.; El-Halwagi, M. M. *AIChE Journal* **2007**, *53* (5), 1232–1239.
349. Bommareddy, S.; Chemmangattualappil, N. G.; Solvason, C. C.; Eden, M. R. *Comput. Chem. Eng.* **2010**, *34* (9), 1481–1486.
350. Bommareddy, S.; Chemmangattualappil, N. G.; Solvason, C. C.; Eden, M. R. *Br. J. Chem. Eng.* **2010**, *27* (3), 441–450.
351. Eljack, F. Computer-Aided Molecular Design and Property Predictions. In *Tools for Chemical Product Design: From Consumer Products to Biomedicine*; Martin, M., Eden, M., Chemmangattualappil, N., Eds.; Computer Aided Chemical Engineering vol. 39; Elsevier: Amsterdam, 2016 pp 153–175. Chapter 6.
352. Kheireddine, H. A.; El-Halwagi, M. M.; Elbasher, N. O. *Clean Techn. Environ. Policy* **2013**, *15* (1), 35–44.
353. Ng, L. Y.; Andiappan, V.; Chemmangattualappil, N. G.; Ng, D. K. *Ind. Eng. Chem. Res.* **2015**, *54* (21), 5722–5735.
354. Ng, L. Y.; Andiappan, V.; Chemmangattualappil, N. G.; Ng, D. K. *Comput. Chem. Eng.* **2015**, *81*, 288–309.
355. Murata, T.; Ishibuchi, H.; Tanaka, H. *Comput. Ind. Eng.* **1996**, *30*, 957–968.

356. Diwekar, U. *Introduction to Applied Optimization*. Kluwer Academic: Boston, MA, 2003.
357. Fu, Y.; Diwekar, U. M. *Ann. Oper. Res.* **2004**, *132* (1–4), 109–134.
358. Fu, Y.; Diwekar, U. M.; Young, D.; Cabezas, H. *Clean Prod. Process.* **2000**, *2* (2), 92–107.
359. Ng, L. Y.; Chemmangattuvalappil, N. G.; Ng, D. K. *Ind. Eng. Chem. Res.* **2014**, *53* (44), 17429–17444.
360. Limleamthong, P.; Gonzalez-Miquel, M.; Papadokostantakis, S.; Papadopoulos, A. I.; Seferlis, P.; Guillén-Gosálbez, G. *Green Chem.* **2016**, *18* (24), 6468–6481.
361. Khor, S. Y.; Liam, K. Y.; Loh, W. X.; Tan, C. Y.; Ng, L. Y.; Hassim, M. H.; Ng, D. K.; Chemmangattuvalappil, N. G. *Process Saf. Environ. Prot.* **2017**, *106*, 211–223.
362. Ooi, J.; Promentilla, M. A. B.; Tan, R. R.; Ng, D. K.; Chemmangattuvalappil, N. G. *Process Saf. Environ. Prot.* **2017**, *111*, 663–677.
363. Ooi, J.; Promentilla, M. A. B.; Tan, R. R.; Ng, D. K.; Chemmangattuvalappil, N. G. *Comput. Chem. Eng.* **2018**, *109*, 191–202.
364. Ten, J. Y.; Hassim, M. H.; Chemmangattuvalappil, N.; Ng, D. K. *J. Loss Prev. Process Ind.* **2016**, *40*, 67–80.
365. Aminbakhsh, S.; Gunduz, M.; Sonmez, R. *J. Safety Res.* **2013**, *46*, 99–105.
366. Ten, J. Y.; Hassim, M. H.; Ng, D. K. S.; Chemmangattuvalappil, N. G. *Chem. Eng. Sci.* **2017**, *159*, 140–153.
367. Bortz, M.; Burger, J.; Aspiron, N.; Blagov, S.; Bottcher, R.; Nowak, U.; Scheithauer, A.; Welke, R.; Kuffer, K. H.; Hasse, H. *Comput. Chem. Eng.* **2014**, *60*, 354–363.
368. Dev, V. A.; Chemmangattuvalappil, N. G.; Eden, M. R. *Comput.-Aided Chem. Eng.* **2016**, *38*, 2055–2060.
369. Dev, V.A. Ph.D. Thesis, Auburn University, Alabama, US, 2017.
370. Mavrotas, G. *Appl. Math Comput.* **2009**, *213* (2), 455–465.
371. Mavrotas, G.; Florios, K. *Appl. Math Comput.* **2013**, *219* (18), 9652–9669.
372. Kocis, G. R.; Grossmann, I. E. *Comput. Chem. Eng.* **1989**, *13* (3), 307–315.
373. Grossmann, I.E.; Viswanathan, J.; Vecchiotti, A.; Raman, R.; Kalvelagen, E., 2018, DICOPT webpage, www.gams.com/248/docs/solvers/dicopt/index.html.
374. GAMS. GAMS webpage, 2018, www.gams.com.
375. Messac, A.; Ismail-Yahaya, A.; Mattson, C. A. *Struct. Multidiscipl. Optim.* **2003**, *25* (2), 86–98.
376. Valencia-Marquez, D.; Flores-Tlacuahuac, A.; Vasquez-Medrano, R. *J. Clean. Prod.* **2017**, *168*, 1652–1667.
377. Sioukrou, E.; Galindo, A.; Adjiman, C. S. *Chem. Eng. Sci.* **2014**, *115*, 19–30.
378. Chong, F. K.; Eljack, F. T.; Atilhan, M.; Foo, D. C.; Chemmangattuvalappil, N. G. *Comput. Chem. Eng.* **2016**, *91*, 219–232.
379. Chong, F. K.; Chemmangattuvalappil, N. G.; Eljack, F. T.; Atilhan, M.; Foo, D. C. *Clean Technol. Environ. Policy* **2016**, *18* (4), 1177–1188.
380. Lee, G. C.; Smith, R.; Zhu, X. X. *Ind. Eng. Chem. Res.* **2002**, *41* (20), 5016–5028.
381. Solvason, C. C.; Chemmangattuvalappil, N. G.; Eden, M. R. *Comput. Chem. Eng.* **2009**, *33* (5), 977–991.
382. Mavrou, P.; Papadopoulos, A. I.; Stjepovic, M. Z.; Seferlis, P.; Linke, P.; Voutetakis, S. *Appl. Therm. Eng.* **2015**, *75*, 384–396.
383. Mavrou, P.; Papadopoulos, A. I.; Seferlis, P.; Linke, P.; Voutetakis, S. *Appl. Therm. Eng.* **2015**, *89*, 1054–1067.
384. Vecchiotti, A.; Grossmann, I. E. *Comput. Chem. Eng.* **1999**, *23* (4–5), 555–565.
385. SBB. SBB webpage, 2018, www.gams.com/latest/docs/S_SBB.html.
386. Solvason, C. C.; Chemmangattuvalappil, N. G.; Eljack, F. T.; Eden, M. R. *Ind. Eng. Chem. Res.* **2009**, *48* (4), 2245–2256.
387. Yunus, N. A.; Gernaey, K. V.; Woodley, J. M.; Gani, R. *Comput. Chem. Eng.* **2014**, *66*, 201–213.
388. Mattei, M.; Kontogeorgis, G. M.; Gani, R. *Comput.-Aided Chem. Eng.* **2012**, *31*, 220–224.
389. Folic, M.; Adjiman, C. S.; Pistikopoulos, E. N. *AIChE J.* **2007**, *53* (5), 1240–1256.
390. Sobol', I. Y. M. *Comput. Math. Math. Phys.* **1967**, *7* (4), 86–112.
391. Folic, M.; Adjiman, C. S.; Pistikopoulos, E. N. *Ind. Eng. Chem. Res.* **2008**, *47*, 5190–5202.
392. Zhou, T.; Lyu, Z.; Qi, Z.; Sundmacher, K. *Chem. Eng. Sci.* **2015**, *137*, 613–625.
393. Diky, V.; Chirico, R. D.; Muzny, C. D.; Kazakov, A. F.; Kroenlein, K.; Magee, J. W.; Abdulgatov, I.; Kang, J. W.; Gani, R.; Frenkel, M. *J. Chem. Inf. Model.* **2012**, *53* (1), 249–266.
394. Papadopoulos, A. I.; Badr, S.; Chremos, A.; Forte, E.; Zargiannis, T.; Seferlis, P.; Papadokostantakis, S.; Galindo, A.; Jackson, G.; Adjiman, C. S. *Mol. Syst. Design Eng.* **2016**, *1* (3), 313–334.
395. Papadopoulos, A. I.; Badr, S.; Chremos, A.; Forte, E.; Zargiannis, T.; Seferlis, P.; Papadokostantakis, S.; Adjiman, C. S.; Galindo, A.; Jackson, G. *Chem. Eng. Trans.* **2014**, *39*, 211–216.
396. Martin, M.; Martinez, A. *Ind. Eng. Chem. Res.* **2015**, *54* (22), 5990–6001.
397. Ng, L. Y.; Chemmangattuvalappil, N. G.; Ng, D. K. *Comput. Chem. Eng.* **2015**, *83*, 186–202.
398. Ten, J. Y.; Ng, L. Y.; Hassim, M. H.; Ng, D. K.; Chemmangattuvalappil, N. G. *Ind. Eng. Chem. Res.* **2017**, *56* (37), 10413–10427.
399. Frutiger, J.; Andreassen, J.; Liu, W.; Spliethoff, H.; Haglind, F.; Abildskov, J.; Sin, G. *Energy* **2016**, *109*, 987–997.
400. McKay, M. D.; Beckman, R. J.; Conover, W. J. *Dent. Tech.* **1979**, *21* (2), 239–245.
401. Kucherenko, S.; Rodriguez-Fernandez, M.; Pantelides, C.; Shah, N. *Reliab. Eng. Syst. Saf.* **2009**, *94* (7), 1135–1148.
402. Frutiger, J.; Cignitti, S.; Abildskov, J.; Woodley, J. M.; Sin, G. *Comput.-Aided Chem. Eng.* **2017**, *40*, 973–978.
403. Frutiger, J.; Marcarie, C.; Abildskov, J.; Sin, G. *J. Chem. Eng. Data* **2016**, *61* (1), 602–613.
404. Andres-Martinez, O.; Flores-Tlacuahuac, A. *Ind. Eng. Chem. Res.* **2018**, *57*, 5058–5069.
405. Satyanarayana, C. K.; Gani, R.; Abildskov, J. *Fluid Phase Equilib.* **2007**, *261* (1–2), 58–63.
406. Satyanarayana, K. C.; Abildskov, J.; Gani, R. *Comput. Chem. Eng.* **2009**, *33* (5), 1004–1013.
407. Buxton, A.; Hugo, A.; Livingston, A. G.; Pistikopoulos, E. N. Identification of Multistep Reaction Stoichiometries: CAMD Problem Formulation. In *Computer Aided Chemical Engineering*; Achenie, L. E. K., Gani, R., Venkatasubramanian, V., Eds.; Vol. 12; Elsevier: Amsterdam, 2003; pp 167–209. Chapter 7.
408. Buxton, A.; Livingston, A. G.; Pistikopoulos, E. N. *Comput. Chem. Eng.* **1997**, *21*, S959–S964.
409. Gani, R.; Jimenez-Gonzalez, C.; Constable, D. J. *Comput. Chem. Eng.* **2005**, *29* (7), 1661–1676.
410. Folic, M.; Gani, R.; Jimenez-Gonzalez, C.; Constable, D. J. *Chin. J. Chem. Eng.* **2008**, *16* (3), 376–383.
411. Gani, R.; Gomez, P. A.; Folic, M.; Jimenez-Gonzalez, C.; Constable, D. J. *Comput. Chem. Eng.* **2008**, *32* (10), 2420–2444.
412. Hansen, C. M. *Prog. Org. Coat.* **2004**, *51* (1), 77–84.
413. Gerbaud, V.; Dos Santos, M. T.; Pandya, N.; Aubry, J. M. *Chem. Eng. Sci.* **2017**, *159*, 177–193.
414. Stanescu, I.; Achenie, L. E. *Chem. Eng. Sci.* **2006**, *61* (18), 6199–6212.
415. Stanescu, I.; Gupta, R. R.; Achenie, L. E. K. *Molecular Simulation* **2006**, *32* (3–4), 279–290.
416. Hechinger, M.; Voll, A.; Marquardt, W. *Comput. Chem. Eng.* **2010**, *34* (12), 1909–1918.
417. Voll, A.; Marquardt, W. *AIChE J.* **2012**, *58* (6), 1788–1801.
418. Dahmen, M.; Marquardt, W. *Energy Fuel* **2017**, *31* (4), 4096–4121.
419. Strubeing, H.; Obermeier, S.; Sioukrou, E.; Adjiman, C. S. *Chem. Eng. Sci.* **2017**, *159*, 69–83.
420. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B.* **2009**, *113* (18), 6378–6396.
421. Chilès, J.-P.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*. John Wiley and Sons: Hoboken, NJ, 2013.
422. Lin, S. T.; Sandler, S. I. *Ind. Eng. Chem. Res.* **2002**, *41* (5), 899–913.
423. De Vleeschouwer, F.; Geerlings, P.; De Proft, F. *ChemPhysChem* **2016**, *17* (10), 1414–1424.
424. Pearl, J.; Korf, R. E. *Annu. Rev. Comput. Sci.* **1987**, *2*, 451–467.

425. De Vleeschouwer, F.; Yang, W.; Beratan, D. N.; Geerlings, P.; De Profijt, F. *Phys. Chem. Chem. Phys.* **2012**, *14* (46), 16002–16013.
426. Satyanarayana, K. C.; Abildskov, J.; Gani, R.; Tsolou, G.; Mavrantzas, V. G. *Br. J. Chem. Eng.* **2010**, *27* (3), 369–380.
427. Farahipour, R.; Mehrkesh, A.; Karunanithi, A. T. *Chem. Eng. Sci.* **2016**, *145*, 126–132.
428. Peng, D.; Zhang, J.; Cheng, H.; Chen, L.; Qi, Z. *Chem. Eng. Sci.* **2017**, *159*, 58–68.
429. Zhang, J.; Qin, L.; Peng, D.; Zhou, T.; Cheng, H.; Chen, L.; Qi, Z. *Chem. Eng. Sci.* **2017**, *162*, 364–374.
430. Zhang, J.; Peng, D.; Song, Z.; Zhou, T.; Cheng, H.; Chen, L.; Qi, Z. *Chem. Eng. Sci.* **2017**, *162*, 355–363.
431. Scheffczyk, J.; Schafer, P.; Fleitmann, L.; Thien, J.; Redepenning, C.; Leonhard, K.; Marquardt, W.; Bardow, A. *Mol. Syst. Design Eng.* **2018**. <https://doi.org/10.1039/C7ME00125H>.
432. Zhao, Y.; Gani, R.; Afzal, R. M.; Zhang, X.; Zhang, S. *AIChE J.* **2017**, *63* (4), 1353–1367.
433. Ahmad, M. Z.; Hashim, H.; Mustafa, A. A.; Maarof, H.; Yunus, N. A. J. *Clean. Prod.* **2018**, *176*, 704–715.
434. Vaidya, P. D.; Kenig, E. Y. *Chem. Eng. Technol.* **2007**, *30* (11), 1467–1474.
435. Dev, V. A.; Chemmangattavalappil, N. G.; Eden, M. R. *Comput.-Aided Chem. Eng.* **2014**, *33*, 151–156.
436. Dev, V. A.; Chemmangattavalappil, N. G.; Eden, M. R. *Comput.-Aided Chem. Eng.* **2015**, *37*, 1445–1450.
437. Moity, L.; Molinier, V.; Benazzouz, A.; Barone, P.; Marion, P.; Aubry, J.-M. *Green Chem.* **2014**, *16*, 146–160.
438. Pereira, F. E.; Keskes, E.; Galindo, A.; Jackson, G.; Adjiman, C. S. Integrated Design of CO₂ Capture Processes From Natural Gas. In *Process Systems Engineering: Energy Systems Engineering*; Georgiadis, M. C., Kikkinides, E. S., Pistikopoulos, E. N., Eds.; Vol. 5; John Wiley & Sons: Hoboken, NJ, 2008; pp 231–248. Chapter 8.
439. Salazar, J.; Diwekar, U.; Joback, K.; Berger, A. H.; Bhow, A. S. *Energy Procedia* **2013**, *37*, 257–264.
440. Chen, C. C.; Britt, H. I.; Boston, J. F.; Evans, L. B. *AIChE J.* **1982**, *28* (4), 588–596.
441. Ahlers, J.; Yamaguchi, T.; Gmehling, J. *Ind. Eng. Chem. Res.* **2004**, *43* (20), 6569–6576.
442. Zhou, T.; McBride, K.; Zhang, X.; Qi, Z.; Sundmacher, K. *AIChE J.* **2015**, *61* (1), 147–158.
443. Liu, B.; Wen, Y.; Zhang, X. *Can. J. Chem. Eng.* **2017**, *95* (4), 767–774.
444. Song, Z.; Zhang, C.; Qi, Z.; Zhou, T.; Sundmacher, K. *AIChE J.* **2018**, *64* (3), 1013–1025.
445. Chemmangattavalappil, N. G.; Solvason, C. C.; Bommarreddy, S.; Eden, M. R. *Comput. Chem. Eng.* **2010**, *34* (5), 582–591.
446. Weis, D. C.; Visco, D. P. *Comput. Chem. Eng.* **2010**, *34* (7), 1018–1029.
447. Eljack, F. T.; Solvason, C. C.; Chemmangattavalappil, N.; Eden, M. R. *Chin. J. Chem. Eng.* **2008**, *16* (3), 424–434.
448. Chen, J.; Sarma, B.; Evans, J. M.; Myerson, A. S. *Cryst. Growth Des.* **2011**, *11* (4), 887–895.
449. Gounaris, C. E.; Floudas, C. A.; Wei, J. *Chem. Eng. Sci.* **2006**, *61* (24), 7933–7948.
450. Gounaris, C. E.; Wei, J.; Floudas, C. A. *Chem. Eng. Sci.* **2006**, *61* (24), 7949–7962.
451. First, E. L.; Gounaris, C. E.; Wei, J.; Floudas, C. A. *Phys. Chem. Chem. Phys.* **2011**, *13* (38), 17339–17358.
452. First, E. L.; Floudas, C. A. *Microporous Mesoporous Mater.* **2013**, *165*, 32–39.
453. First, E. L.; Gounaris, C. E.; Floudas, C. A. *Langmuir* **2013**, *29* (18), 5599–5608.
454. Matsuda, H.; Yamamoto, H.; Kurihara, K.; Tochigi, K. *Fluid Phase Equilib.* **2007**, *261* (1–2), 434–443.
455. Chavez-Islas, L. M.; Vasquez-Medrano, R.; Flores-Tlacuahuac, A. *Ind. Eng. Chem. Res.* **2011**, *50* (9), 5153–5168.
456. Karunanithi, A. T.; Mehrkesh, A. *AIChE J.* **2013**, *59* (12), 4627–4640.
457. Chong, F. K.; Foo, D. C.; Eljack, F. T.; Atilhan, M.; Chemmangattavalappil, N. G. *Clean Technol. Environ. Policy* **2015**, *17* (5), 1301–1312.
458. Chong, F. K.; Andiappan, V.; Eljack, F. T.; Foo, D. C. Y.; Chemmangattavalappil, N. G.; Ng, D. K. S. Ionic Liquid Design for Biomass-based Tri-Generation System With Carbon Capture. In *Process Systems and Materials for CO₂ Capture: Modelling, Design, Control and Integration*; Papadopoulos, A. I., Seferlis, P., Eds.; John Wiley & Sons: Hoboken, NJ, 2017; pp 203–227. Chapter 8.
459. Hasan, M. M. F.; First, E. L.; Floudas, C. A. *Phys. Chem. Chem. Phys.* **2013**, *15* (40), 17601–17618.
460. First, E. L.; Hasan, M. M.; Floudas, C. A. *AIChE J.* **2014**, *60* (5), 1767–1785.
461. Schwobel, J. A.; Preifinger, M.; Bruggemann, D.; Klamt, A. *Ind. Eng. Chem. Res.* **2017**, *56* (3), 788–798.
462. Stijepovic, M. Z.; Papadopoulos, A. I.; Linke, P.; Grujic, A. S.; Seferlis, P. *Energy* **2014**, *69*, 285–298.
463. Mehrkesh, A.; Karunanithi, A. T. *Comput. Chem. Eng.* **2016**, *93*, 402–412.
464. Lin, Y.; Schrage, L. *Optim. Methods Softw.* **2009**, *24* (4–5), 657–668.
465. Vaidyanathan, R.; Gowayed, Y.; El-Halwagi, M. *Comput. Chem. Eng.* **1998**, *22* (6), 801–808.
466. Pavurala, N.; Achenie, L. E. K. *Comput. Chem. Eng.* **2014**, *71*, 734–744.
467. Camarda, K. V.; Sunderesan, P. *Ind. Eng. Chem. Res.* **2005**, *44* (12), 4361–4367.
468. Perdomo, F. A.; Perdomo, L.; Millan, B. M.; Aragon, J. L. *Chem. Eng. Res. Design* **2014**, *92* (8), 1482–1494.
469. Zhang, L.; Babi, D. K.; Gani, R. *Annu. Rev. Chem. Biomol. Eng.* **2016**, *7*, 557–582.
470. Siddhaye, S.; Camarda, K. V.; Topp, E.; Southard, M. *Comput. Chem. Eng.* **2000**, *24* (2–7), 701–704.
471. Siddhaye, S.; Camarda, K.; Southard, M.; Topp, E. *Comput. Chem. Eng.* **2004**, *28* (3), 425–434.
472. Zhang, L.; Mao, H.; Liu, L.; Du, J.; Gani, R. *Comput. Chem. Eng.* **2018**, *115*, 295–308.
473. ICAS webpage, 2018, ICAS Software package, www.capec.kt.dtu.dk.
474. Hasan, M. F. Multi-Scale Process Systems Engineering for Carbon Capture, Utilization, and Storage: A Review. In *Process Systems and Materials for CO₂ Capture: Modelling, Design, Control and Integration*; Papadopoulos, A. I., Seferlis, P., Eds.; John Wiley & Sons: Hoboken, NJ, 2017; pp 231–247. Chapter 9.
475. Chen, J. J. F.; Visco, D. P., Jr. *Chem. Eng. Sci.* **2017**, *159*, 31–42.
476. Adjiman, C. S.; Bardow, A. *Chem. Eng. Sci.* **2017**, *159*, 1–230.
477. Martin, M. M.; Eden, M. R.; Chemmangattavalappil, N. G. *Tools for Chemical Product Design: From Consumer Products to Biomedicine, Computer-Aided Chemical Engineering (Vol. 39)*. Elsevier: Amsterdam, 2016.
478. Adjiman, C. S.; Galindo, A. In *Molecular Systems Engineering. Process Systems Engineering*; Pistikopoulos, E. N., Georgiadis, M. C., Dua, V., Eds.; Vol. 6; John Wiley & Sons: Hoboken, NJ, 2010.
479. Gani, R.; Ng, K. M. *Comput. Chem. Eng.* **2015**, *81*, 70–79.
480. Papadakis, E.; Woodley, J. M.; Gani, R. *Comput.-Aided Chem. Eng.* **2018**, *41*, 597–656.
481. Singh, R.; Yuan, Z. *Process Systems Engineering for Pharmaceutical Manufacturing. Computer-Aided Chemical Engineering 41*. Elsevier: Amsterdam, 2018.
482. Gernaey, K. V.; Gani, R. *Chem. Eng. Sci.* **2010**, *65* (21), 5757–5769.
483. Papadopoulos, A. I.; Seferlis, P. *Process Systems and Materials for CO₂ Capture: Modelling, Design, Control and Integration*. John Wiley & Sons: Hoboken, NJ, 2017.

Further Reading

- Thi, C. L.; Tamouza, S.; Passarello, J.-P.; Tobaly, P.; de Hemptinne, J.-C. *Ind. Eng. Chem. Res.* **2006**, *45* (20), 6803–6810.
- Haslam, A. J.; Galindo, A.; Jackson, G. *Fluid Phase Equilib.* **2008**, *266*, 105–128.
- Glover, F. *ORSA J. Comput.* **1989**, *1* (3), 190–206.