Global Flowsheet Optimization for Reductive Dimethoxymethane Production Using Data-Driven Thermodynamic Models

Jannik Burre,[†] Christoph Kabatnik,[‡] Mohamed Al-Khatib,[†] Dominik Bongartz,[†] Andreas Jupke,[‡] and Alexander Mitsos[*,¶,†,§]

[†]*Process Systems Engineering (AVT.SVT), RWTH Aachen University, Forckenbeckstr. 51, 52074 Aachen, Germany*

[‡]*Fluid Process Engineering (AVT.FVT), RWTH Aachen University, Forckenbeckstr. 51, 52074 Aachen, Germany*

[¶]*JARA-ENERGY, 52056 Aachen, Germany*

[§]*Energy Systems Engineering (IEK-10), Forschungszentrum Jülich, 52425 Jülich, Germany*

E-mail: amitsos@alum.mit.edu

Phone: +49 241 80 94808. Fax: +49 241 80 92326

Abstract

The absence of accurate thermodynamic models for reductive dimethoxymethane (DMM) synthesis has limited the design of corresponding processes to approximate calculations only. To enable a more reliable process design, we measure liquid equilibrium densities and fit parameters for the PCP-SAFT equation of state (EOS). This EOS is highly accurate for systems at high pressures and therefore suitable for the high pressure reactor and the flash unit for gas recycling. As the resulting flowsheet optimization problem is nonconvex, we use our deterministic global solver MAiNGO to solve the problem. To improve computational tractability, we approximate process models that require the PCP-SAFT EOS with artificial neural networks and Gaussian processes. Finally, the so-called *reduced-space* problem formulation and a hybrid of the McCormick and the auxiliary variable method enable convergence within 5.8 CPUh. At the optimal operating conditions, an exergy efficiency of 91.9 % is achieved for a reactor pressure of 120 bar.

Keywords: hybrid modeling, global optimization, process design, dimethoxymethane, PCP-SAFT

1 Introduction

As the mobility sector accounts for more than 20 % of global greenhouse gas emissions,[1] it plays a crucial role in climate change mitigation. Whereas battery electric vehicles become increasingly relevant for new vehicles, the existing fleet of passenger cars and long-haul transportation most likely will continue to depend on liquid fuels with a high energy-density for the next decades. In this regard, dimethoxymethane (DMM) and its derivatives polyoxymethylene ethers ($OME_{3\text{-}5}$) are promising synthetic fuels, which can be produced from renewable hydrogen (H_2) and carbon dioxide (CO_2)[2–6] or from biomass.[7–11] In addition to their potentially clean production from renewable resources, their clean combustion[12,13] and similar fuel properties to diesel[14,15] make them an attractive (partial) replacement for diesel. However, their economic and environmentally benign production still remains challenging.

By only replacing fossil-based raw materials with renewable ones, major process inefficiencies remain present within the established process concepts via DMM and trioxane.[3,4] Therefore, new processes need to be developed. One alternative to the established process for $OME_{3\text{-}5}$ production is its direct synthesis from methanol and aqueous formaldehyde (FA),[16,17] thus circumventing the energy intensive process step for trioxane production. However, the presence of water in such a process route inhibits $OME_{3\text{-}5}$ formation and complicates product purification.[18] Promising process concepts that can effectively remove water from the system are currently under development.[4,19,20] Another process alternative is the anhydrous synthesis of $OME_{3\text{-}5}$ from dimethyl ether (DME) and trioxane.[21,22] However, this still requires an energy-intensive trioxane production step. In contrast, a process route via DMM and gaseous FA[23,24] can be highly resource efficient—which is key for power-to-X processes—if gaseous FA is produced via methanol dehydrogenation[25] and DMM via the direct reduction of CO_2.[26]

Whereas there has been extensive work on FA synthesis via methanol dehydrogenation within the last decades, DMM synthesis via direct CO_2 reduction has just recently been discovered. The existing publications on process development, including our own,[27] rely on intermediate-fidelity models incorporating activity coefficient thermodynamic models. Activity coefficient models are

however usually not accurate at high pressures[28] and for systems containing significant amounts of quadrupolar components (e.g., CO_2). For such systems, to which the reductive synthesis of DMM belongs to, equations of state (EOS) should be preferred. In this regard, the perturbed-chain polar statistical associating fluid theory (PCP-SAFT) EOS[29,30] has been successfully applied to various systems.[31] The necessity of a complicated thermodynamic model in combination with typically nonlinear process unit models for such chemical processes results in a nonconvex optimization problem, which makes deterministic process optimization for the direct CO_2 reduction challenging. Especially deterministic global optimization, which is required to guarantee optimal solutions for nonconvex problems, is often computationally not tractable.

In this work, we measure liquid equilibrium densities and use data from the open literature to fit missing binary parameters of the PCP-SAFT EOS for the underlying system of components. We then use the PCP-SAFT EOS to develop data-driven models to make global optimization for reductive DMM production tractable while keeping model accuracy high. More specifically, we use a data-driven model for the reactor to predict the solubility of H_2 and CO_2 in the liquid reaction mixture at high pressure and their conversion with methanol to DMM. Additionally, a data-driven model is developed for the flash downstream the reactor unit that recycles unreacted gaseous components. To also account for DMM purification and identify the least energy-intensive separation process, we consider two alternative distillation column sequences formulated as a simple superstructure.

Section 2 introduces the process concept that is used for process optimization and the underlying chemical reaction. In Section 3, we fit binary PCP-SAFT parameters and develop data-driven models for the reactor and the flash unit, as well as the superstructure model for distillation column sequencing. On the basis of these models, Section 4 describes how the optimization problem is solved using our open-source deterministic global solver MAiNGO,[32] before the results are discussed. In Section 5, we conclude our findings.

2 DMM production via direct CO_2 reduction

The production of DMM via the direct reduction of CO_2 is based on the reaction

$$2CH_3OH + CO_2 + 2H_2 \rightleftharpoons DMM + 2H_2O, \tag{R1}$$

which is catalyzed by a ruthenium-based catalyst.[26,33] In this reaction, CO_2 and H_2 are dissolved under high pressure in methanol, where the formation of the intermediate product methyl formate (MF) takes place, before it is finally converted to DMM.

In our previous publication,[27] we developed a hierarchical process development and comparison methodology that we applied to several reaction pathways for DMM production. Using intermediate-fidelity models, the reaction pathway according to Reaction (R1) was found to be the most suitable one for sustainable DMM production at its current state of development. As the goal of this study is to refine the process and optimize its design and operating conditions using a more accurate thermodynamic model, some process modifications are required (Fig. 1). First, the unreacted gases dissolved in the liquid reactor effluent are separated by a flash and a low-temperature distillation column. Second, as the amount of MF at reaction equilibrium is negligible (based on our own calculations considering the two-step reaction[26] with MF as an intermediate) and its recycling together with H_2 and CO_2 is simple, we do not consider MF formation in this study. Third, we do not consider a fixed distillation column sequence for the purification of DMM as it is dependent on the upstream reaction performance. Instead, we use a superstructure model to incorporate the choice for the optimal sequence into the optimization resulting in a mixed-integer nonlinear program (MINLP). As the mixture contains an azeotrope between methanol and DMM, a pressure swing distillation (column D1 and D2) is considered (Fig. 1). Finally, we only consider DMM synthesis from methanol, CO_2, and H_2. Thus, we do not optimize the upstream process for methanol production via CO_2 hydrogenation.

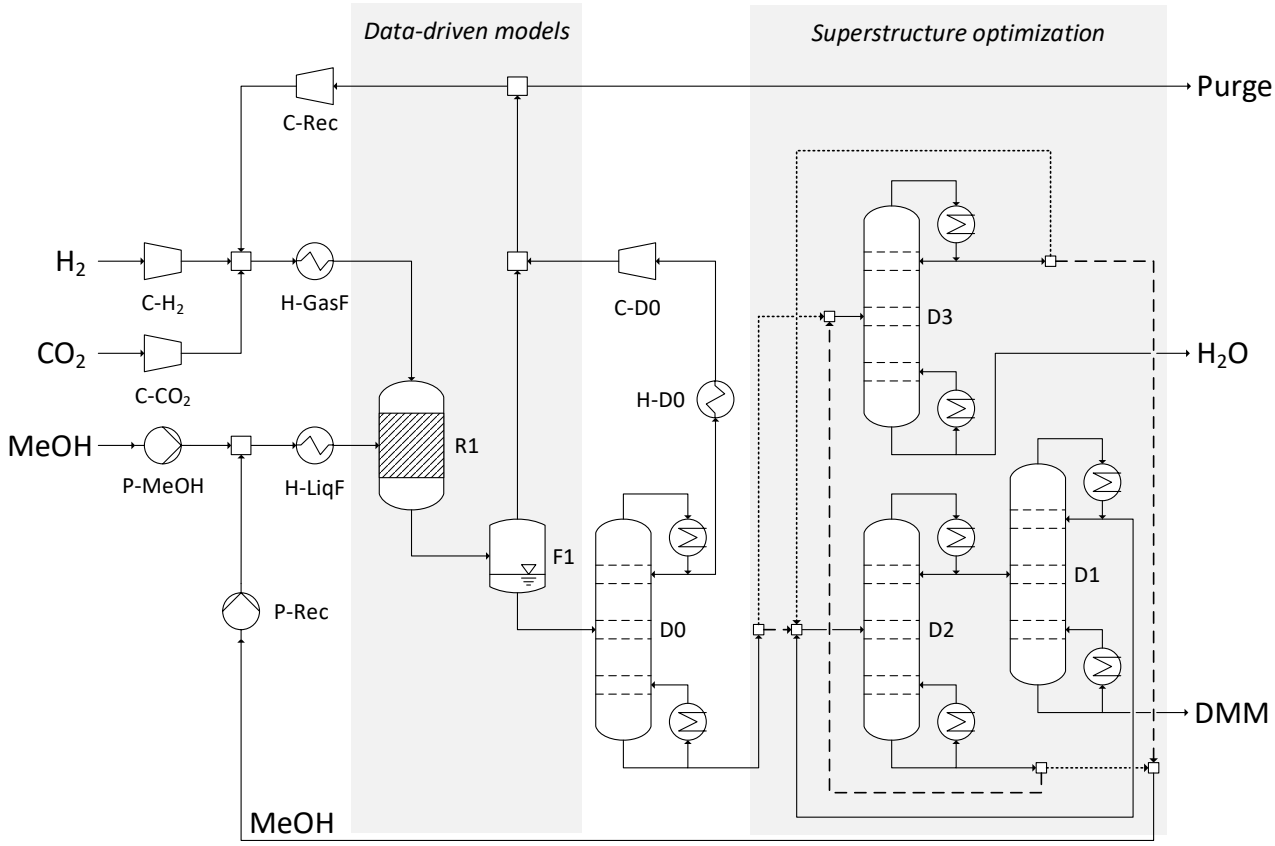

Figure 1 Process flowsheet for the reductive production of DMM. The corresponding process model includes a data-driven model for the reactor, where the vapor-liquid reaction (R1) takes place at high pressure, and a data-driven model for a flash unit, where unreacted H_2 and CO_2 is separated and recycled. Downstream the distillation column D0, which separates remaining CO_2, a superstructure model for the purification of DMM using pressure swing distillation (column D1 and D2) is used for column sequencing. The different line types correspond to the different column configurations.

For the exothermic vapor-liquid reaction (R1), a temperature of 80 °C has been found experimentally to be ideal for the ruthenium-based catalyst at its current state due to its required minimum activation energy.[26] In contrast, the ideal reactor pressure still remains unknown. Experiments have shown that a high pressure enhances the solubility of gases into methanol and may increase conversion.[26,33,34] However, a high reactor pressure results in high compression cost. The use of co-solvents with an enhanced solubility for H_2 and CO_2 has high potential to counteract corresponding mass transfer limitations within the reactor. However, these co-solvents also have an

influence on the catalytic reaction,[35] which has not been investigated sufficiently so far. Therefore, we do not consider the use of co-solvents within our study. The combination of a high pressure and the presence of components with a quadrupole moment (CO_2) that might influence fluid properties significantly makes the consideration of an accurate thermodynamic model inevitable.

3 Modeling

3.1 Thermodynamic model

Since the reactor is the central unit of the process and has a significant influence on the downstream process units, we place a firm focus on the accuracy of the reactor model. As the reaction according to Equation (R1) requires high pressures (based on our own experiments and those from literature,[26,33,34] a reactor pressure of 50 bar is sufficient to ensure a reasonable solubility of H_2 and CO_2 in the liquid), the reactor is modeled with the PCP-SAFT EOS. As the subsequent flash unit recycles unreacted gaseous educts potentially at high pressures (to avoid recycling of gaseous DMM), it has a direct influence on the reaction and is therefore also modeled with the PCP-SAFT EOS. For the downstream process at moderate pressures, the system is approximated as an ideal system (with the azeotrope between DMM and methanol being considered as a pseudo-component, cf. Section 3.2.1) to maintain optimization tractability while still obtaining estimates on the exergy demand for DMM purification. As the exergy demand for DMM purification makes up only a smaller part of the total energy demand,[27] the resulting inaccuracies are expected to affect overall process performance only moderately.

3.1.1 PCP-SAFT equation of state

PCP-SAFT EOS is based on PC-SAFT EOS[36] with additional polar terms for dipole-dipole[30] and quadrupole-quadrupole[29] interactions. For further information on the used model, we refer to Aigner et al.[37]

Phase equilibrium calculations using only pure component parameters (ESI Tab. S1) often deliver results with significant deviations to experimental data.[28] To gain reliable phase equilibria for multicomponent systems, conventional combining rules usually require the adjustment of interaction parameters to the binary subsystems. We consider two cases: For non associating systems, the binary interaction parameter k_{ij} is used for the correction of dispersive interactions.[38] In case of cross-associating systems, polar interactions occur involving a molecule that is either a hydrogen

bond acceptor or a hydrogen bond donor. This effect is modeled by the binary association strenght ε^{A_iBj} and binary association volume κ^{A_iBj}.[37] These corrections lead to a significant improvement of phase equilibrium calculations.

In our case, the solubility of the gaseous educts H_2 and CO_2 is of utter importance because of its significant influence on the chemical equilibrium and thus on the total process performance. Therefore, binary parameters are used for all binary subsystems including H_2 or CO_2. Parameters are either taken directly from literature or adjusted to vapor-liquid equilibria from literature. Since there is no experimental VLE-data for H_2 or CO_2 with DMM available in the open literature, we measured liquid equilibrium densities (ESI Section 2), which we used to calculate the binary interaction parameters $k_{CO_2,DMM}$ and $k_{H_2,DMM}$ (Tab. 1). For parameter fitting, the deviation between the experimental results and those generated by the PCP-SAFT EOS are evaluated by the weighted root-mean-square deviation RMSDw

$$\mathrm{RMSDw}(\rho) = \sqrt{\frac{1}{N}\sum_{n}^{N}\left(\frac{\rho_{\mathrm{eq},n}^{\mathrm{calc}} - \rho_{\mathrm{eq},n}^{\mathrm{exp}}}{u_{\mathrm{c}}(\rho)_{i,n}}\right)^2} \tag{1}$$

taking the uncertainties $u_{\mathrm{c}}(\rho)_{i,n}$ of the experimental liquid equilibrium densities $\rho_{\mathrm{eq}}^{\mathrm{exp}}$ into account. It is noticeable that all gas solubilities were significantly underestimated without the use of binary parameters.

Table 1 Binary parameters for the PCP-SAFT EOS.

Parameter source	Component i	Component j	k_{ij} / –	ε^{A_iBj} / K	κ^{A_iBj} / –	Data Ref.
Literature	CO_2	Water	-	2882.3	5.72967×10^{-4}	Aigner et al.[37]
Regression	CO_2	Methanol	-	3127.43	6.06313×10^{-4}	Leu et al.[39]
	H_2	Water	-0.4622	-	-	Gillespie and Wilson[40], Kling and Maurer[41], DeVaney et al.[42]
	H_2	Methanol	-0.5132	-	-	Brunner et al.[43]
Experiments	CO_2	DMM	-0.0875	-	-	this work
	H_2	DMM	-0.127	-	-	this work

3.1.2 Data-driven thermodynamic model

Given the complicated form of the PCP-SAFT EOS, it is not straightforward to consider this thermodynamic model directly in deterministic global process optimization. It is not available in most commercial process simulation tools and complicated to rigorously implement in process models. The multitude of terms required to describe the complex interactions between different types of molecules and phases introduces many variables and makes the solution of the corresponding system of equations challenging. Only few standalone model implementations are available in the open literature,[44] which can however not be integrated into commercial process simulation tools. To still benefit from the high accuracy of the PCP-SAFT EOS for process optimization, we develop data-driven models for the reactor and the flash unit for gas recycling, which can efficiently be used for deterministic process optimization.[45,46] The presence of both data-driven and mechanistic models results in a hybrid process model. With this hybrid modeling approach, we enable the integration of models that are too complicated for deterministic global optimization while exploiting the large validity range of mechanistic models for the remaining process units.

Flash model

In accordance to the reported suitability of artificial neural networks[47] (ANNs) for approximating phase equilibrium calculations for systems described by the PC-SAFT EOS,[48] we use an ANN to model the vapor-liquid equilibrium (VLE) within the flash unit F1 considering the PCP-SAFT EOS. The input variables of this model are the operating conditions of the flash (T_{F1} and P_{F1}) and the component mole fractions of the liquid reactor effluent ($x_{\mathrm{Prod},i}$). The bounds for $x_{\mathrm{Prod},i}$ correspond to the attainable region of the reaction effluent and are summarized in Tab. 2. In order to yield a linear process model for the flash unit and keep the output dimension as small as possible, the output variables are chosen to be the split factors ξ_i of each component i. As the consideration of 6 input variables requires a large set of samples, we generate 10,000 data points using a mechanistic flash model implemented in Matlab incorporating the PCP-SAFT EOS.

For deterministic global process optimization, it is crucial to keep the problem size as small as

Table 2 Input variables and their bounds for the training of the ANN flash model.

Input variable	Description	Bounds
T_{F1} / °C	Temperature of flash F1	[25;100]
P_{F1} / bar	Pressure of flash F1	[4;40]
x_{Prod,H_2} / -	Liquid mole fraction of component H_2 in reactor effluent	[0.01;0.03]
x_{Prod,CO_2} / -	Liquid mole fraction of CO_2 in reactor effluent	[0.05;0.11]
$x_{Prod,DMM}$ / -	Liquid mole fraction of DMM in reactor effluent	[0.01;0.08]
x_{Prod,H_2O} / -	Liquid mole fraction of water in reactor effluent	[0.03;0.14]

possible and relaxations as tight as possible in order to keep the optimization tractable. At the same time, accuracy requirements need to be met. To find the optimal trade-off between model accuracy and computational performance, sophisticated methods exist (e.g., ALAMO - Automatic Learning of Algebraic MOdels[49]). Such methods use a library of terms with a simple functional form to iteratively build a process or unit model until a desired accuracy is reached. The resulting models are tailored to fulfill the desired trade-off as good as possible. In this study, we do not intend to find the sweet spot of this trade-off, but rather show that deterministic global optimization for complex processes is also possible with off-the-shelf surrogate models. In fact, in the reduced-space problem formulation in MAiNGO (cf. Section 4.1), the optimization problem remains the same irrespective of the size of the surrogate model, i.e., the ANN, but it only affects the relaxation tightness and the model accuracy.[45] To ensure a reasonable accuracy, we performed a sensitivity analysis regarding the number of hidden layers and neurons per layer. The mean squared error (MSE) for all cases and all settings for the training of the ANN model are summarized in Fig. 2. For an acceptable maximum MSE of 10^{-4}, the ANN must have at least 10 neurons in total. Numerical experiments have shown that the ANN with 3 layers and 4 neurons for each layer results in a lower solution time than ANNs with the same size but different amount of layers or those with an even smaller size. This indicates that the number of layers has a significant influence on the relaxation tightness (cf., Section 4.2). The prediction capabilities of the final ANN model are demonstrated in Fig. 3a-3c exemplary for the split factors of H_2, CO_2, and DMM.

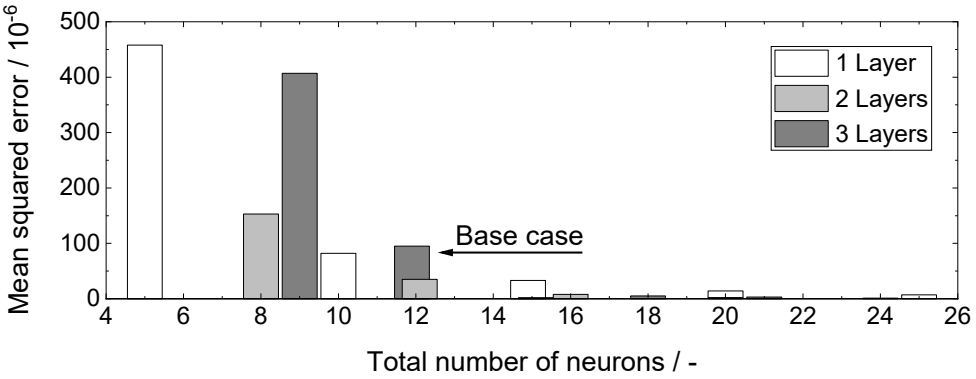

Figure 2 MSE averaged over all test set data points for ANNs with a different total number of neurons and a different number of hidden layers. For training, the Levenberg Marquardt training function, a hyperbolic tangent transfer function, a training ratio of 70 %, a validation ratio of 15 %, and a test ratio of 15 % was used.

Reactor model

For the reductive synthesis of DMM from methanol, H_2 and CO_2, there is no reaction kinetic model available in the open literature. As we want to find the maximum expected process performance, we consider reaction equilibrium throughout this study. Although the catalytic reaction (R1) takes place only in the liquid phase, the VLE within the reactor influences reaction equilibrium and vice versa. Therefore, the conversion of methanol is dependent on the ratio of H_2 to CO_2 in the gaseous reactor feed, as well as on the ratio of gas to liquid within the reactor. However, a corresponding sensitivity analysis (ESI Section 3) has shown that their influence on reactor performance is low. As mainly the reactor performance determines how much gas need to be supplied back to the system, the two ratios do not affect process design and exergy demand significantly. Therefore, a fixed gas composition corresponding reaction stoichiometry and a molar ratio of 1:1 for the amount of gas within the reactor is used for the optimization. For industrial application, the catalyst is assumed to be immobilized and therefore not withdrawn with the liquid reactor effluent containing only DMM, water, and unreacted methanol, H_2, and CO_2. With these assumptions, only the reactor pressure needs to be considered as input variable for the data-driven reactor model. The output variable has been chosen to be methanol conversion at simultaneous phase and reaction equilibrium to yield linear equations for the reactor process model. As unreacted H_2 and CO_2 remain in the liquid

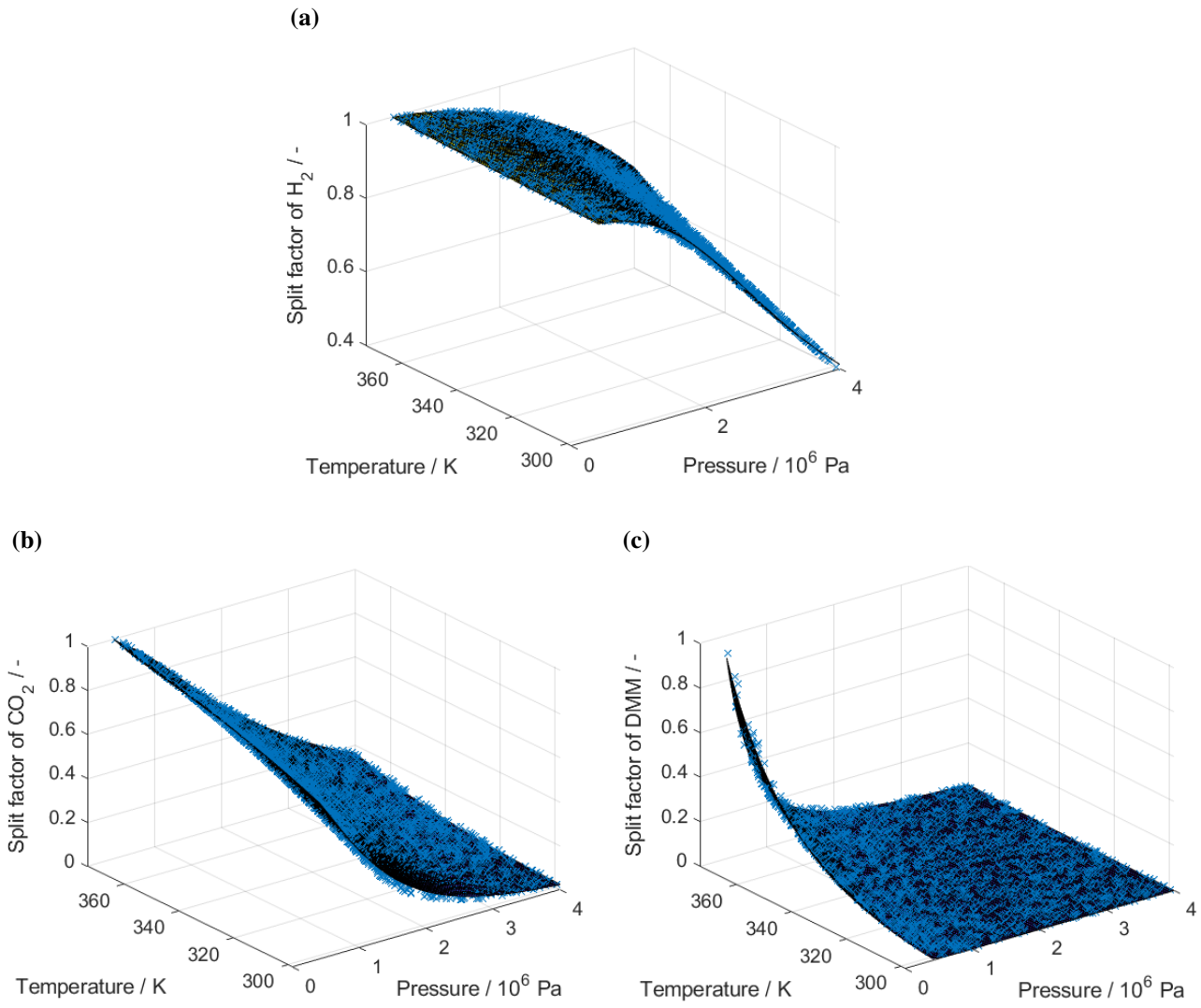

Figure 3 ANN prediction (surface) and PCP-SAFT training data (×) for the split factor of H_2 (Fig. 3a), CO_2 (Fig. 3b), and DMM (Fig. 3c) within the flash unit F1 as a function of temperature and pressure. An ANN with 3 layers and 4 neurons each with the settings summarized in the description of Fig. 2 is used for optimization. The plotted split factors are shown for the globally optimal molar reactor outlet composition (H_2: 0.018, CO_2: 0.091, DMM: 0.070, Methanol: 0.694, Water: 0.127).

phase at reaction equilibrium, also the K-values for these two components need to be considered as output variables.

For data generation, we use the same in-house Java implementation of the PCP-SAFT EOS as for the flash model. As it does not allow the simultaneous consideration of chemical reactions to this date and convergence is sensitive to initial values in the first place, an iterative procedure for

calculating the combined phase and chemical equilibrium is applied in Matlab. First, the phase equilibrium via the PCP-SAFT EOS determines how much gas dissolves in methanol. The resulting liquid phase composition $\tilde{x}_{\mathrm{Prod},i}$ (before reaction) is then used to solve the definition of the equilibrium constant

$$K(T) = \prod_{i \in C} \left(\frac{1}{\tilde{x}_{\mathrm{Prod},i}} \frac{\tilde{f}_i}{f_i^0} x_{\mathrm{Prod},i} \right)^{\nu_i} = \exp\left(\frac{-\Delta G_{\mathrm{R}}^0}{RT} \right) \tag{2}$$

for mole fractions $x_{\mathrm{Prod},i}$ (after reaction). The fugacities $\tilde{f}_i$ of each component i can be taken from the in-house Java implementation of the PCP-SAFT EOS, in which the standard state fugacity f_i^0 corresponds to standard pressure. Parameter ν_i is the stoichiometric coefficient of component i according to Reaction (R1). The standard Gibbs energy is calculated by

$$\Delta G_{\mathrm{R}}^0 = \sum_{i \in C} \nu_i \mu_i^0 = \sum_{i \in C} \nu_i \left(\int_{T_0}^{T} c_{\mathrm{p},i}^{\mathrm{iG}} dT' - T \int_{T_0}^{T} \frac{c_{\mathrm{p},i}^{\mathrm{iG}}}{T'} dT' + \Delta_{\mathrm{f}} H_i^{\mathrm{iG}}(T^0) - T \Delta_{\mathrm{f}} S_i^{\mathrm{iG}}(T^0, p^0) \right). \tag{3}$$

The ideal gas heat capacities $c_{\mathrm{p},i}^{\mathrm{iG}}$, the standard enthalpies of formation $\Delta_{\mathrm{f}} H_i^{\mathrm{iG}}$ and the standard molar entropies $\Delta_{\mathrm{f}} S_i^{\mathrm{iG}}$ are taken from the DIPPR 801 Database. The iterations eventually terminate once a threshold for the reaction extent has been reached.

The iterative procedure makes the application of the mechanistic reactor model computationally much more demanding compared to the mechanistic flash model resulting in long computation times per data point. Therefore, data-driven modeling using ANNs is not suitable because of the high number of required samples. Instead, given the small number of attainable samples, Gaussian processes[50] (GPs) represent a suitable alternative modeling approach. As we consider only a single input for the reactor model and for each of the three output variables its own GP, only a small set of samples is required to accurately model the behavior within the reactor. Similarly to the data-driven flash model, we performed a sensitivity analysis to determine the minimum complexity of the GP models to improve optimization tractability. As the GP model complexity scales with the number of data points, the most suitable trade-off between model accuracy and computational performance could be achieved with 6 data points (cf. Section 4.2). The corresponding MSEs are

summarized in Tab. 3. The prediction capabilities of the final GP models are demonstrated in Fig. 4a-4c for the equilibrium methanol conversion (C_{MeOH}) and the K-values of H_2 and CO_2. The large deviation between the simulation data calculated with the PCP-SAFT and the PC-SAFT EOS for these variables (Fig. 4) highlights the importance of the correct choice of the thermodynamic model for the reactor model.

Table 3 MSE for the GP models for C_{MeOH}, $K_{\mathrm{H_2}}$, and $K_{\mathrm{CO_2}}$ for a different number of considered training data points and the matern kernel with parameter 3/2 as the covariance function. The MSE corresponds to the test set only. The base case model is highlighted in gray.

	GP data points				
	4	6	8	10	12
$\mathrm{MSE}_{C_{\mathrm{MeOH}}}$ / 10^{-10}	140	6.5	1.1	0.5	0.6
$\mathrm{MSE}_{K_{\mathrm{H_2}}}$ / -	0.707	0.039	0.029	0.019	0.010
$\mathrm{MSE}_{K_{\mathrm{CO_2}}}$ / 10^{-4}	50	2.95	1.89	1.13	0.62

3.2 Process model

For the remaining process units, the consideration of a simpler thermodynamic model is justified as either only moderate operating pressures are considered or H_2 and CO_2 have already been separated. Furthermore, the remaining process units provide only approximate estimates for the exergy demand for compression and product purification, for which simple models are sufficient as they are expected to make up only a smaller share of the overall exergy demand.[27]

3.2.1 Distillation column sequencing using superstructure optimization

As the optimal distillation column sequence is generally dependent on the upstream process, a fixed sequence could lead to a suboptimal process design and operating conditions. Therefore, we consider superstructure optimization, which calculates the optimal distillation column sequence based on the optimal operating conditions of the upstream process. We consider only distillation columns D1-D3 for superstructure optimization as CO_2 should be separated first to minimize the need for low-temperature distillation (cf. Fig. 1). This results in only two possible sequences,

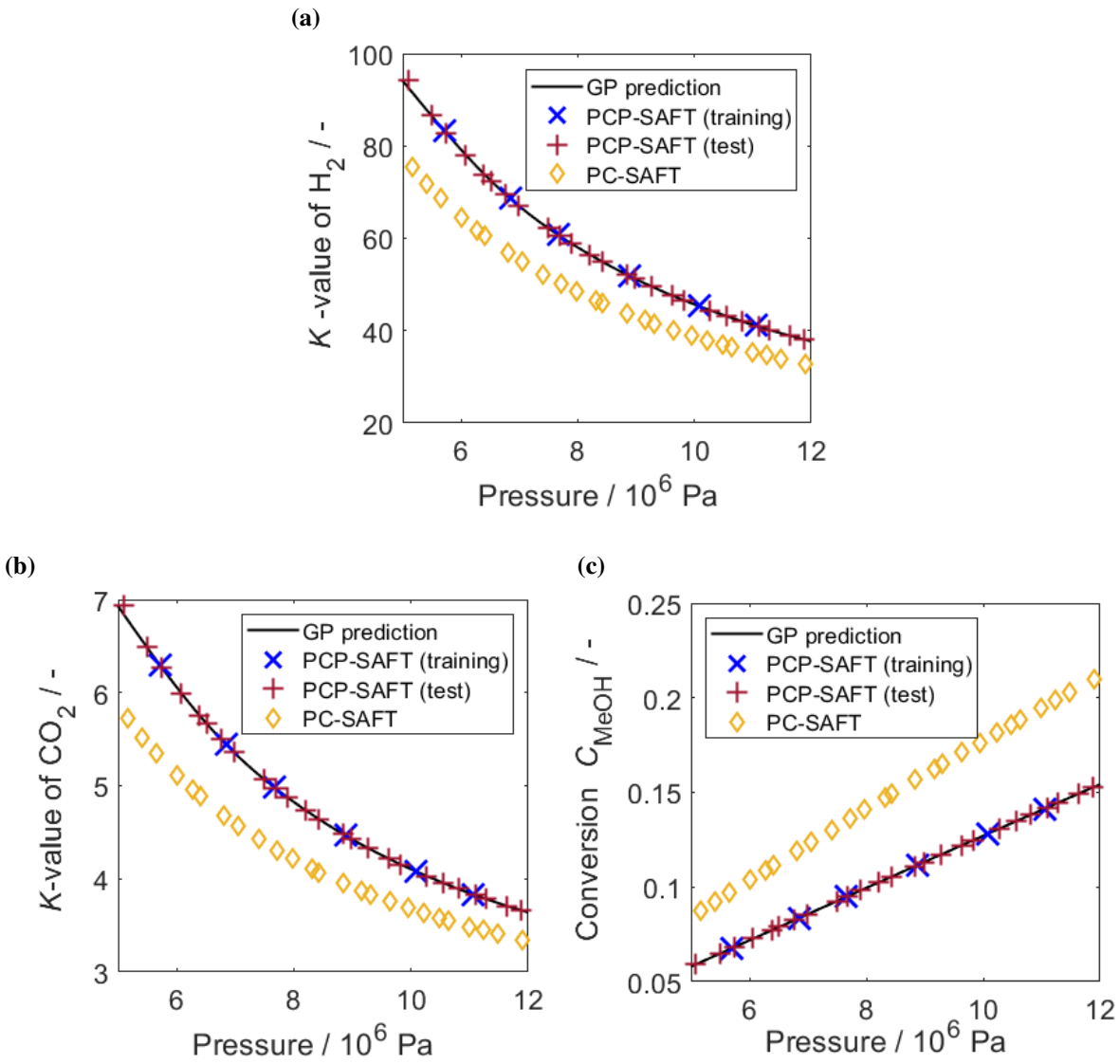

Figure 4 GP prediction (black line) and PCP-SAFT training data (×) for the K-values of H_2 (Fig. 4a) and CO_2 (Fig. 4b), and methanol equilibrium conversion C_{MeOH} (Fig. 4c) for the reactor unit as a function of pressure. The red plus markers (+) represent the PCP-SAFT test data used to calculate the MSE reported in Tab. 3 to demonstrate the validity of the GP model between training samples. The orange diamonds (◇) correspond to the PC-SAFT EOS and demonstrate the deviation between PC-SAFT and PCP-SAFT EOS for the underlying system.

which we could have simply considered as two separate nonlinear programs (NLP). We nevertheless formulate the separation process as a superstructure and demonstrate that solving the MINLP is cheaper than enumerating the two options as NLPs.

In this study, we use the state-equipment network[51] (SEN) superstructure representation, in

which we assign all separation tasks that have a cut between the same components to the same distillation column (separation cut SEN).[52,53] In contrast to the state-task network, this representation requires the smallest number of distillation column models, while keeping model equations comparatively simple.

For the separation cut SEN, the connection between columns can be fully described by two types of binary variables. Variable X_d^{F} indicates whether column d is connected to the global feed (i.e., the bottom product of column D0) to the superstructure. Variable $X_{l,j}^{\mathrm{s}}$ indicates whether the output stream s (distillate or bottom) of column l is connected to column j. In this work, the following equations are used to describe the connection between the distillation columns:

$$X_{2,3}^{\mathrm{B}} + X_{3,2}^{\mathrm{D}} = 1 \tag{4}$$

$$X_3^{\mathrm{F}} = X_{3,2}^{\mathrm{D}} \tag{5}$$

$$X_2^{\mathrm{F}} = X_{2,3}^{\mathrm{B}}. \tag{6}$$

Note that the distillate stream of column 1 is always fed to column 2 and the distillate stream from column 2 is always fed to column 1, regardless of the selected sequence (Fig. 1). To determine the feed flow rates to a column, all flow rates that can be fed to a column are multiplied by the corresponding binary variable to ensure that only the flow rates of the active connections are used (cf. *Direct MINLP* problem formulation in Burre et al.[54]). For example, the feed of component i to column 2 is given by

$$F_{2,i} = D_{1,i} + X_2^{\mathrm{F}}\, F_i + X_{3,2}^{\mathrm{D}}\, D_{3,i}\,, \tag{7}$$

where F_i is the global feed of component i and $D_{d,i}$ is the distillate flow rate of component i in column d.

For the distillation column models, the Underwood equations[55] are used. In order to apply the Underwood method to the azeotropic mixture considered in this work, the coordinate transformation presented by Liu et al.[56] is used. In their method, the azeotropic system is divided into subsystems which behave approximately like a non-azeotropic mixture. Within this transformation, the azeotropes are treated as pseudo-components. In this work, the subsystems are modeled as ideal mixtures and the vapor pressures of the pure components are determined using the extended Antoine equation. To determine the vapor pressures of the azeotropic pseudo-components, the Antoine parameters are fitted using data from flash calculations in Aspen Plus.

3.2.2 Miscellaneous models

The remaining units are modeled using simple process models to get an estimate on exergy demand while maintaining optimization tractability. For gas compression, we use a one-stage compressor model to keep the amount of optimization variables small and consider an isentropic and mechanical efficiency of 80 % and 90 %, respectively. This provides a rather conservative estimate, as the model overestimates the actual exergy demand slightly. In Section 4.2, we evaluate whether this simplification is reasonable. For the pumps, an isentropic efficiency of 90 % is considered. For the heat exchanger, we use the logarithmic mean temperature difference to approximate the thermodynamic mean temperature, which we use to compute exergy demand and excess. Heat integration is only approximated in the objective function by the simple summation of exergy demand and excess within the entire system. All parameters for the pure component property models (extended Antoine for vapor pressure, DIPPR-106 for heat of vaporization, DIPPR-107 for heat capacity) are taken from the Aspen Plus DB-PURE37 data base and for the Henry's constant correlation for H_2 and CO_2 from the APV110 HENRY-AP and BINARY data bases.

4 Global optimization

Most process models introduced in Section 3 are nonlinear. In addition to the nonconvex terms of the pure component property models, the hyperbolic tangent activation function applied in the ANN model as well as the covariance function of the GP model introduce nonconvexities into the process model. Also the Underwood equations for modeling the distillation columns within the superstructure and the corresponding discrete decisions therein are responsible for further nonconvex terms. Irrespective of the considered process models, the structure of the process itself with its recycle streams makes the resulting optimization problem nonconvex. To find the most promising process structure and operating conditions for reductive DMM production despite its nonconvex nature, global optimization is desirable.

The application of global optimization to large optimization problems is however challenging. Particularly, problems incorporating data-driven submodels usually exhibit a large number of optimization variables, which often makes the optimization problem not tractable for state-of-the-art deterministic global solvers. To still solve such problems to global optimality, our open-source deterministic global solver MAiNGO[32] effectively exploits the smaller problem size of the so-called *reduced-space* problem formulation,[57] in which the only optimization variables are the degrees of freedom and tear variables. By additionally considering tailored relaxations for the nonconvex terms of the process model, processes incorporating hybrid models could already be solved efficiently to global optimality.[45,46]

4.1 Problem formulation and objective function

The process model introduced in Section 3 is implemented in the programming language C++, in which all intermediate process variables are calculated as functions of the degrees of freedom, tear variables, and a few additional optimization variables (to avoid model equations yielding weak relaxations[58]). The degrees of freedom for the DMM production process depicted in Fig. 1 are the reactor pressure (P_{R1}), the temperature (T_{F1}) and pressure (P_{F1}) of the flash unit for gas recy-

cling, the binary decision variable for the choice of the optimal distillation sequence ($X^{\mathrm{B}}_{2,3}$), as well as the recoveries of the light and heavy key component ($\gamma_{\mathrm{LK},d}$ and $\gamma_{\mathrm{HK},d}$, respectively) of distillation columns D1 - D3 (Tab. 5). As column D0 is not part of the superstructure and separates pure CO_2 from the remaining liquid mixture, $\gamma_{\mathrm{LK},0}$ and $\gamma_{\mathrm{HK},0}$ are fixed to 1 and 0, respectively. Tear variables are introduced for each recycle stream and for process units that otherwise would need to be modeled by implicit functions.[57] The elimination of optimization variables using the model equality constraints (i.e., the reduced-space formulation in MAiNGO) results in a dramatic reduction in problem size. The resulting process model contains only 41 optimization variables, one of which is binary. It has 55 inequality and 31 equality constraints (Tab. 4). To facilitate the modeling procedure and benefit from tailored relaxations, we use the model libraries implemented in MAiNGO (e.g., enthalpy of vaporization, ideal gas enthalpy)[59] and MeLOn.[45,46] The model library MeLOn provides several machine learning models including ANNs and GPs, which are accessed by MAiNGO via a build-in interface. Corresponding model parameters from the training in Matlab are provided by an automatically generated csv- (ANN) or json-file (GP). Relaxations of all functions and their subgradients are automatically obtained from the MC++ library.[60]

The objective function of the optimization problem is the minimization of the net exergy demand

$$\begin{aligned}\dot{E}_{\mathrm{total}} = \dot{n}_{\mathrm{H_2}}\,\hat{e}_{\mathrm{H_2}} + \dot{n}_{\mathrm{MeOH}}\,\hat{e}_{\mathrm{MeOH}} + \sum_{c\in C} W_c + \sum_{p\in P} W_p + \sum_{h\in H} \dot{E}_{\mathrm{Q_{in},}h} + \sum_{d\in D} \dot{E}_{\mathrm{Q_{reb},}d} \\ - \sum_{h\in H} \dot{E}_{\mathrm{Q_{out},}h} - \sum_{d\in D} \dot{E}_{\mathrm{Q_{cond},}d} - \dot{E}_{\mathrm{Q_{R1,out}}},\end{aligned} \tag{8}$$

where $\dot{n}_{\mathrm{H_2}}$ and $\dot{n}_{\mathrm{MeOH}}$ is the net consumption of raw materials H_2 and methanol, respectively, $\hat{e}_{\mathrm{H_2}}$ and $\hat{e}_{\mathrm{MeOH}}$ is their molar exergy, W_{c} is the power input of compressor $c \in C$, W_p is the power input of pump $p \in P$, $\dot{E}_{\mathrm{Q_{in},}h}$ and $\dot{E}_{\mathrm{Q_{out},}h}$ is the exergy flow of the heat demand and excess for heat exchanger $h \in H$ and flash F1, $\dot{E}_{\mathrm{Q_{reb},}d}$ and $\dot{E}_{\mathrm{Q_{cond},}d}$ is the exergy flow of reboiler and condenser duties of distillation column $d \in D$, and $\dot{E}_{\mathrm{Q_{R1,out}}}$ is the exergy flow of excess heat from the reaction. The ambient temperature is assumed to be 25 °C to calculate the exergy flows.

The corresponding process model that can be given directly to the deterministic global solver MAiNGO is available as electronic supplement.

4.2 Results and discussion

The deterministic global solver MAiNGO[32] v0.5.0.2 employs a spatial branch-and-bound algorithm with several bound tightening techniques and uses the multivariate McCormick method[61,62] implemented in MC++[60] to obtain relaxations. The optimization is carried out with the parallel version of MAiNGO on an Intel Xeon Platinum 8160 processor using 40 cores. Both the relative and absolute optimality tolerance is set to 10^{-3}. To improve convergence, we use the following non-default settings in MAiNGO: First, we utilize a combination of an adaptation of the *Kelley's* algorithm[63] and a *n*-simplex algorithm to linearize relaxations instead of utilizing a midpoint linearization. Second, we selectively consider auxiliary variables (AVs) for repeated nonlinear terms to improve the tightness of the relaxations.[58] With this, the base case optimization problem considers *Kelley's* and *n*-simplex relaxation linearization, 84 AVs, 6 GP data points, and 3 layers and 4 neurons each layer. All characteristics of the optimization problem are summarized in Tab. 4.

The consideration of a special linearization strategy for relaxations reduces the number of nodes required for convergence significantly but in turn increases solution time per node. The consideration of AVs is key for convergence in the first place (Tab. 4 and Fig. 5). As MAiNGO treats the process model in the reduced-space as one function being dependent only on the degrees of freedom, tear variables, and a few additional optimization variables (cf. Section 4.1), the model relaxation is constructed from a sequence of mathematical operations (cf. McCormick method),[61] which results from the procedural concatenation of explicit model equations implemented in the C++ code. Within this sequence of mathematical operations some individual terms may appear repeatedly, which could weaken model relaxations.[62] To prevent this and still yield tight relaxations for the optimization in the reduced-space, we add certain selected AVs to benefit from both the reduced problem size and potentially tight relaxations from the auxiliary variable method[64,65] (AVM) typically employed by most state-of-the-art deterministic global solvers.[66,67]

Table 4 Problem size and numerical results for different objective functions, solver settings, and model detail. For all considered cases, the global optimal solution was found in the root node.

Objective function	$\min \dot{E}_{\text{total}}$					$\max \eta_{\text{Ex}}$
Solver settings / model detail	Base case	Midpoint linearization	No AVs	GP w 8 data points	ANN w 5 neurons	Base case
Number of						
Continuous variables	40	40	40	40	40	40
Discrete variables	1	1	1	1	1	1
Equality constraints	31	31	31	31	31	31
Inequality constraints	55	55	55	55	55	55
B&B nodes	2715	225887	456000[a]	45000[a]	37200[a]	38500[a]
Optimal objective value / MJ kg^{-1}	27.4	27.4	27.4[a]	27.4[a]	27.5[a]	91.9 %[a]
Lower bound of root node / MJ kg^{-1}	-16	-2.2 $\times 10^{9}$	-3.5 $\times 10^{11}$	-16	-61.3	-4.1 $\times 10^{8}$
CPU time per B&B node / s	7.7	0.4	0.6	6.4	7.7	7.5
Total CPU time / h	5.8	27.6	80[a]	80[a]	80[a]	80[a]
Rel. optimality gap / %	0	0	9.0 $\times 10^{7}$	0.8	0.5	1.6

[a] Optimization has reached the CPU limit of 80 CPUh.

The base case optimization problem is solved to global optimality in 5.8 CPUh or 2715 nodes (Tab. 4 and Fig. 5). The global solution is 2.08 MJ mol^{-1} or 27.4 MJ kg^{-1} net exergy demand per produced DMM, which corresponds to an exergy efficiency of 91.9 %. At the optimal operating conditions, a maximum reactor pressure of 120 bar is applied (Tab. 5) resulting in an equilibrium methanol conversion of 15.4 %. As the pressure variable is at its upper bound, an even higher pressure could result in an even better performance but also requires a more complex reactor design. With the one-stage compressor model, compression accounts for only 4 % of the total exergy demand, which would decrease even further if a multi-stage compressor model would be considered. As the optimal operating pressure is already at its upper bound, a multi-stage compressor model would not influence the optimal operating conditions. For a detailed process design at a later stage of development, however, a multi-stage compressor model should be considered.

On the modeling level, the choice of the objective function also has a significant influence on the optimization. Instead of minimizing net exergy demand (cf. Equation (8)), maximizing exergy efficiency

$$\eta_{\text{Ex}} = \frac{\dot{n}_{\text{DMM}}\, \hat{e}_{\text{DMM}} + \sum_{h\in H} \dot{E}_{\text{Q}_{\text{out},h}} + \sum_{d\in D} \dot{E}_{\text{Q}_{\text{cond},d}} + \dot{E}_{\text{Q}_{\text{R1,out}}}}{\dot{n}_{\text{H}_2}\, \hat{e}_{\text{H}_2} + \dot{n}_{\text{MeOH}}\, \hat{e}_{\text{MeOH}} + \sum_{c\in C} W_c + \sum_{p\in P} W_p + \sum_{h\in H} \dot{E}_{\text{Q}_{\text{in},h}} + \sum_{d\in D} \dot{E}_{\text{Q}_{\text{reb},d}}}, \tag{9}$$

the optimization does not converge within 80 CPUh (optimality gap of 1.6 %, Tab. 4). Also the complexity of each process model must be kept moderate to yield tight relaxations. Considering 8 instead of 6 data points for the Gaussian process reactor model, an optimality gap of 0.8 % remains. Considering 5 instead of 4 neurons for each of the 3 layers for the flash model, an optimality gap of 0.5 % remains (Tab. 4). If the two individual NLPs are solved, the total solution time (30.1 CPUh) exceeds that of the base case MINLP problem incorporating the superstructure model for distillation column sequencing significantly. This shows that a superstructure representation (SEN, cf. Section 3.2.1) together with a problem formulation (Direct MINLP[54] in a reduced-space) that both do not introduce many additional variables (here, only the binary variable $X^{\mathrm{B}}_{2,3}$) are promising for deterministic global optimization.

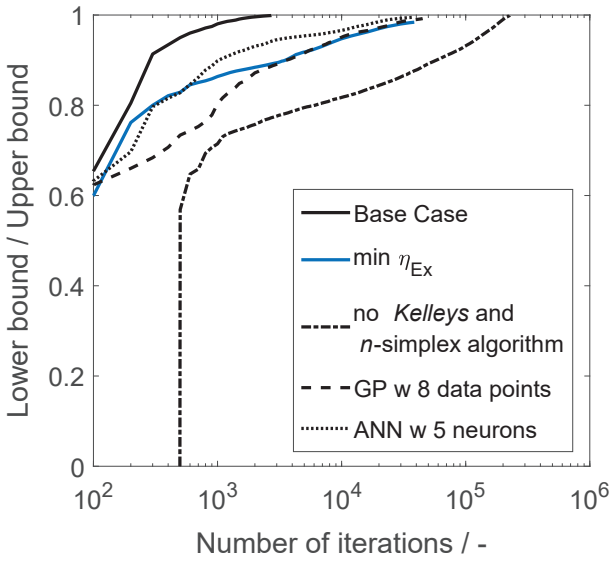

Figure 5 Convergence indicated by the ratio of the lower bound to the upper bound during the course of optimization for all considered cases. As the lower bound for the case without the consideration of AVs is low and does not increase within the course of optimization, it is not displayed in this diagram.

The flash unit operates at 4 bar and 42.0 °C to recycle 99.6 % H_2 and 77.6 % CO_2. The rest of the CO_2 is separated by the low-temperature distillation column D0, before DMM can be purified in the subsequent distillation column sequence. The optimal sequence is obtained for $X^{\mathrm{B}}_{2,3} = 1$, where column D2 first separates the azeotrope between methanol and DMM from the bottom

Table 5 Degrees of freedom for the reductive DMM production process (Fig. 1), their interval bounds and optimal values.

Degree of freedom	Description	Bounds	Optimal value
P_{R1} / bar	Pressure of reactor R1	[50, 120]	120
T_{F1} / °C	Temperature of flash F1	[25, 90]	42.0
P_{F1} / bar	Pressure of flash F1	[4, 10]	4
$X^{B}_{2,3}$ / -	Decision variable for connecting column D2 with D3 via the bottom product	{0, 1}	1
$\gamma_{LK,1}$ / -	Recovery of the LK component of column D1	[0, 1]	1
$\gamma_{HK,1}$ / -	Recovery of the HK component of column D1	[0, 1]	0
$\gamma_{LK,2}$ / -	Recovery of the LK component of column D2	[0, 1]	1
$\gamma_{HK,2}$ / -	Recovery of the HK component of column D2	[0, 1]	0
$\gamma_{LK,3}$ / -	Recovery of the LK component of column D3	[0, 1]	1
$\gamma_{HK,3}$ / -	Recovery of the HK component of column D3	[0, 1]	0.05

product of column D0. Then, DMM is separated from methanol in the pressure swing distillation comprising column D2 and D3 leading to a total share of exergy demand for separation of about 9.6 % of that for the entire process. This is in good agreement with values from literature (7 %[5] and 8 %[27], both decoupled from the upstream methanol process). However, the exergy demand for separation reported by Burre et al.[27] corresponds to a reactor pressure of 80 bar, for which a methanol conversion of 15.7 % was estimated using the NRTL thermodynamic model—the most suitable one to this date. Our calculations with the more accurate PCP-SAFT EOS show that this conversion can only be reached at even higher pressures.

The results show that the correct choice of the thermodynamic model for the reactor is crucial for process design and evaluation. To evaluate the accuracy of the simple Underwood model for distillation, we compare corresponding results with those obtained from a tray-to-tray model and the NRTL thermodynamic model (accurate for moderate pressures) in Aspen Plus. The results show that the total exergy demand is overestimated by the Underwood model only slightly by about 10 % (Tab. 6). The estimates for the individual distillation columns are, however, partly inaccurate for the Underwood model for this nonideal system. Although these inaccuracies do not have a significant influence on the overall performance of the process, it certainly has for a more detailed process design at a later stage of development. Therefore, more research should

be dedicated to the development of more accurate distillation models suitable for global flowsheet optimization.

Table 6 Comparison of the reboiler exergy demand for distillation column D0-D3 calculated with the Underwood model and the tray-to-tray model using the NRTL thermodynamic model in Aspen Plus (RadFrac).

Distillation column	Exergy demand / $MJ\,kg^{-1}$	
	Underwood (this study)	Tray-to-tray (Aspen Plus)
D0	0.23	0.24
D1	0.65	0.11
D2	0.23	0.25
D3	1.14	1.41
Total	**2.25**	**2.01**

5 Conclusion

The direct reduction of CO_2 belongs to the most hydrogen-efficient pathways for dimethoxymethane (DMM) production given its favorable reaction stoichiometry. Its need for a high reactor pressure makes the application of thermodynamic models available in the open literature however inaccurate, which has limited process design to the development of simple process models so far.

To enable reliable process design and ultimately advance efficient DMM production, we measured liquid equilibrium densities and fitted binary parameters for the PCP-SAFT equation of state (EOS) for the components involved in the reaction. Whereas this thermodynamic model was found to predict the vapor-liquid equilibrium of the system properly, it constitutes a major challenge for deterministic global optimization for process design. To benefit from both the high accuracy of the thermodynamic model and the potential of deterministic optimization, we developed data-driven thermodynamic models for process units that potentially operate at high pressures and contain significant amounts of H_2 and CO_2. The equilibrium-based reactor model is therefore approximated by Gaussian processes, while the flash unit for gas recycling is approximated by an artificial neural network. In combination with a superstructure model for distillation column sequencing and several recycling streams within the process, the resulting mathematical program is nonconvex. To still find the most favorable process design and operating conditions, we used our open-source deterministic global solver MAiNGO for optimization.

The capability of MAiNGO to exploit the small problem size of the so-called *reduced-space* problem formulation makes the optimization converge to the global optimum in 5.8 CPUh. To achieve this performance, several measures had to be taken: On the algorithm level, a combination of *Kelley's* and a *n*-simplex algorithm for linearize relaxations, as well as a hybrid of the McCormick and the auxiliary variable method had to be used. On the modeling level, a suitable trade-off between model complexity (regarding the data-driven models) and computational performance for global optimization had to be found. The resulting process performance for reductive DMM production from methanol, H_2, and CO_2 was found to be 91.9 % at an optimal reactor pressure of 120 bar. As only simple Underwood models were used for the distillation columns within

the superstructure, which have been found to succumb significant inaccuracies for the underlying system, future work should focus on the development of distillation models that are suitable for global optimization. To increase methanol conversion and decrease the exergy demand for separation, co-solvents could be considered to either enhance gas solubility or enable in-situ extraction of DMM from the reaction phase. In this regard, data on the phase equilibrium of reaction mixtures including different co-solvents and on their suitability for the ruthenium-catalyzed reaction would be beneficial. This data could be used to fit PCP-SAFT EOS parameters for a variety of multi-phase systems. Such parameters would then enable an optimization-based screening of co-solvents, which potentially increase reaction equilibrium and decrease energy demand for DMM purification.

Acknowledgement

The authors gratefully acknowledge funding by the German Federal Ministry of Education and Research (BMBF) within the project NAMOSYN: Nachhaltige Mobilität durch synthetische Kraftstoffe (FKZ 03SF0566P0). We thank Huixin Shi for implementing the basis of the overall process model in C++ and Sebastian Kaminski for providing the in-house Java implementation of the PCP-SAFT EOS.

References

(1) IEA, International Energy Agency, Energy Technology Perspectives 2017: Catalysing Energy Technology Transformation. 2017.

(2) Schemme, S.; Samsun, R. C.; Peters, R.; Stolten, D. Power-to-fuel as a key to sustainable transport systems – An analysis of diesel fuels produced from CO 2 and renewable electricity. *Fuel* **2017**, *205*, 198–221.

(3) Burre, J.; Bongartz, D.; Mitsos, A. Production of Oxymethylene Dimethyl Ethers from Hydrogen and Carbon Dioxide—Part II: Modeling and Analysis for $OME_{3\text{-}5}$. *Industrial & Engineering Chemistry Research* **2019**, *58*, 5567–5578.

(4) Held, M.; Tönges, Y.; Pélerin, D.; Härtl, M.; Wachtmeister, G.; Burger, J. On the energetic efficiency of producing polyoxymethylene dimethyl ethers from CO_2 using electrical energy. *Energy & Environmental Science* **2019**, *12*, 1019–1034.

(5) Deutz, S.; Bongartz, D.; Heuser, B.; Kätelhön, A.; Langenhorst, L. S.; Omari, A.; Walters, M.; Klankermayer, J.; Leitner, W.; Mitsos, A.; Pischinger, S.; Bardow, A. Cleaner production of cleaner fuels: wind-to-wheel – environmental assessment of CO_2-based oxymethylene ether as a drop-in fuel. *Energy & Environmental Science* **2018**, *11*, 331–343.

(6) Schmitz, N.; Burger, J.; Ströfer, E.; Hasse, H. From methanol to the oxygenated diesel fuel

poly(oxymethylene) dimethyl ether: An assessment of the production costs. *Fuel* **2016**, *185*, 67–72.

(7) Mahbub, N.; Oyedun, A. O.; Kumar, A.; Oestreich, D.; Arnold, U.; Sauer, J. A life cycle assessment of oxymethylene ether synthesis from biomass-derived syngas as a diesel additive. *Journal of Cleaner Production* **2017**, *165*, 1249–1262.

(8) Mahbub, N.; Oyedun, A. O.; Zhang, H.; Kumar, A.; Poganietz, W.-R. A life cycle sustainability assessment (LCSA) of oxymethylene ether as a diesel additive produced from forest biomass. *The International Journal of Life Cycle Assessment* **2018**, *24*, 881–899.

(9) Oyedun, A. O.; Kumar, A.; Oestreich, D.; Arnold, U.; Sauer, J. The development of the production cost of oxymethylene ethers as diesel additives from biomass. *Biofuels, Bioproducts and Biorefining* **2018**, *12*, 694–710.

(10) Bokinge, P.; Heyne, S.; Harvey, S. Renewable OME from biomass and electricity—Evaluating carbon footprint and energy performance. *Energy Science & Engineering* **2020**, *8*, 2587–2598.

(11) Zhang, X.; Oyedun, A. O.; Kumar, A.; Oestreich, D.; Arnold, U.; Sauer, J. An optimized process design for oxymethylene ether production from woody-biomass-derived syngas. *Biomass and Bioenergy* **2016**, *90*, 7–14.

(12) Pélerin, D.; Gaukel, K.; Härtl, M.; Jacob, E.; Wachtmeister, G. Potentials to simplify the engine system using the alternative diesel fuels oxymethylene ether OME_1 and $OME_{3\text{-}6}$ on a heavy-duty engine. *Fuel* **2020**, *259*, 116231.

(13) Omari, A.; Heuser, B.; Pischinger, S.; Rüdinger, C. Potential of long-chain oxymethylene ether and oxymethylene ether-diesel blends for ultra-low emission engines. *Applied Energy* **2019**, *239*, 1242–1249.

(14) Vertin, K. D.; Ohi, J. M.; Naegeli, D. W.; Childress, K. H.; Hagen, G. P.; McCarthy, C. I.; Cheng, A. S.; Dibble, R. W. Methylal and Methylal-Diesel Blended Fuels for Use in Compression-Ignition Engines. SAE Technical Paper Series, 1999.

(15) Härtl, M.; Gaukel, K.; Pélerin, D.; Wachtmeister, G. Oxymethylenether als potenziell CO_2-neutraler Kraftstoff für saubere Dieselmotoren Teil 1: Motorenuntersuchungen. *MTZ - Motortechnische Zeitschrift* **2017**, *78*, 52–59.

(16) Schmitz, N.; Ströfer, E.; Burger, J.; Hasse, H. Conceptual Design of a Novel Process for the Production of Poly(oxymethylene) Dimethyl Ethers from Formaldehyde and Methanol. *Industrial & Engineering Chemistry Research* **2017**, *56*, 11519–11530.

(17) Oestreich, D.; Lautenschütz, L.; Arnold, U.; Sauer, J. Production of oxymethylene dimethyl ether (OME)-hydrocarbon fuel blends in a one-step synthesis/extraction procedure. *Fuel* **2018**, *214*, 39–44.

(18) Oestreich, D. *Prozessentwicklung zur Gewinnung von Oxymethylenethern (OME) aus Methanol und Formaldehyd*; PhD Thesis, KIT Scientific Publishing, 2017.

(19) Voggenreiter, J.; Burger, J. Side Products in the Water-Tolerant Synthesis of Poly(oxymethylene) Dimethyl Ethers: Formation Kinetics and Implications for Process Design. *Industrial & Engineering Chemistry Research* **2021**, *60*, 2418–2429.

(20) Schmitz, N.; Breitkreuz, C. F.; Ströfer, E.; Burger, J.; Hasse, H. Separation of water from mixtures containing formaldehyde, water, methanol, methylal, and poly(oxymethylene) dimethyl ethers by pervaporation. *Journal of Membrane Science* **2018**, *564*, 806–812.

(21) Haltenort, P.; Hackbarth, K.; Oestreich, D.; Lautenschütz, L.; Arnold, U.; Sauer, J. Heterogeneously catalyzed synthesis of oxymethylene dimethyl ethers (OME) from dimethyl ether and trioxane. *Catalysis Communications* **2018**, *109*, 80–84.

(22) Breitkreuz, C. F.; Schmitz, N.; Ströfer, E.; Burger, J.; Hasse, H. Design of a Production Process for Poly(oxymethylene) Dimethyl Ethers from Dimethyl Ether and Trioxane. *Chemie Ingenieur Technik* **2018**, *90*, 1489–1496.

(23) Peter, A.; Fehr, S. M.; Dybbert, V.; Himmel, D.; Lindner, I.; Jacob, E.; Ouda, M.; Schaadt, A.; White, R. J.; Scherer, H.; Krossing, I. Towards a Sustainable Synthesis of Oxymethylene Dimethyl Ether by Homogeneous Catalysis and Uptake of Molecular Formaldehyde. *Angewandte Chemie* **2018**, *130*, 9605–9608.

(24) Peter, A.; Stebens, G.; Baumgärtner, J. F.; Jacob, E.; Mantei, F. K.; Ouda, M.; Krossing, I. Facile Two-Phase Catalysis: From Dimethoxymethane and Monomeric Formaldehyde towards Oxymethylene Ethers (OMEs). *ChemCatChem* **2020**, *12*, 2416–2420.

(25) Su, S.; Zaza, P.; Renken, A. Catalytic dehydrogenation of methanol to water-free formaldehyde. *Chemical Engineering & Technology* **1994**, *17*, 34–40.

(26) Thenert, K.; Beydoun, K.; Wiesenthal, J.; Leitner, W.; Klankermayer, J. Ruthenium-Catalyzed Synthesis of Dialkoxymethane Ethers Utilizing Carbon Dioxide and Molecular Hydrogen. *Angewandte Chemie* **2016**, *128*, 12454–12457.

(27) Burre, J.; Bongartz, D.; Deutz, S.; Mebrahtu, C.; Osterthun, O.; Sun, R.; Völker, S.; Bardow, A.; Klankermayer, J.; Palkovits, R.; Mitsos, A. Comparing pathways for electricity-based production of dimethoxymethane as a sustainable fuel. *Energy & Environmental Science* **2021**, *14*, 3686–3699.

(28) Carlson, E. C. Don't gamble with physical properties for simulations. *Chemical engineering progress* **1996**, *92*, 35–46.

(29) Gross, J. An equation-of-state contribution for polar components: Quadrupolar molecules. *AIChE Journal* **2005**, *51*, 2556–2568.

(30) Gross, J.; Vrabec, J. An equation-of-state contribution for polar components: Dipolar molecules. *AIChE Journal* **2006**, *52*, 1194–1204.

(31) Kleiner, M.; Sadowski, G. Modeling of Polar Systems Using PCP-SAFT: An Approach to Account for Induced-Association Interactions. *The Journal of Physical Chemistry C* **2007**, *111*, 15544–15553.

(32) Bongartz, D.; Najman, J.; Sass, S.; Mitsos, A. MAiNGO – McCormick-based Algorithm for mixed-integer Nonlinear Global Optimization, Technical Report. 2018; `http://permalink.avt.rwth-aachen.de/?id=729717`.

(33) Siebert, M.; Seibicke, M.; Siegle, A. F.; Kräh, S.; Trapp, O. Selective Ruthenium-Catalyzed Transformation of Carbon Dioxide: An Alternative Approach toward Formaldehyde. *Journal of the American Chemical Society* **2019**, *141*, 334–341.

(34) Schieweck, B. G.; Klankermayer, J. Tailor-made Molecular Cobalt Catalyst System for the Selective Transformation of Carbon Dioxide to Dialkoxymethane Ethers. *Angewandte Chemie* **2017**, *129*, 10994–10997.

(35) Thenert, K. M. Maßgeschneiderte Ruthenium-Katalysatoren für die stoffliche Nutzung von CO2 in Kombination mit molekularem Wasserstoff. Ph.D. thesis, Universitätsbibliothek der RWTH Aachen, 2018.

(36) Gross, J.; Sadowski, G. Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules. *Industrial & Engineering Chemistry Research* **2001**, *40*, 1244–1260.

(37) Aigner, M.; Echtermeyer, A.; Kaminski, S.; Viell, J.; Leonhard, K.; Mitsos, A.; Jupke, A. Ternary System CO_2/2-MTHF/Water—Experimental Study and Thermodynamic Modeling. *Journal of Chemical & Engineering Data* **2019**,

(38) Tumakaka, F.; Gross, J.; Sadowski, G. Thermodynamic modeling of complex systems using PC-SAFT. *Fluid Phase Equilibria* **2005**, *228-229*, 89–98.

(39) Leu, A.-D.; Chung, S. Y.-K.; Robinson, D. B. The equilibrium phase properties of (carbon dioxide + methanol). *The Journal of Chemical Thermodynamics* **1991**, *23*, 979–985.

(40) Gillespie, P.; Wilson, G. *Vapor-liquid equilibrium data on water-substitute gas components: N_2-H_2O, H_2-H_2O, CO-H_2O, H_2-CO-H_2O, and H_2S-H_2O*; 1980.

(41) Kling, G.; Maurer, G. Solubility of hydrogen in aqueous ethanolamine solutions at temperatures between 323 and 423 K. *Journal of Chemical & Engineering Data* **1991**, *36*, 390–394.

(42) DeVaney, W.; Berryman, J. M.; Kao, P. L.; Eakin, B. High temperature VLE measurements for substitute gas components. Gas Processors Association.

(43) Brunner, E.; Hültenschmidt, W.; Schlichthärle, G. Fluid mixtures at high pressures IV. Isothermal phase equilibria in binary mixtures consisting of (methanol + hydrogen or nitrogen or methane or carbon monoxide or carbon dioxide). *The Journal of Chemical Thermodynamics* **1987**, *19*, 273–291.

(44) Institute of Physical Chemistry, University of Cologne, ThermoC. 2018; `http://thermoc.uni-koeln.de/index.html`, Accessed: 2021-11-15.

(45) Schweidtmann, A. M.; Mitsos, A. Deterministic Global Optimization with Artificial Neural Networks Embedded. *Journal of Optimization Theory and Applications* **2018**, *180*, 925–948.

(46) Schweidtmann, A. M.; Bongartz, D.; Grothe, D.; Kerkenhoff, T.; Lin, X.; Najman, J.; Mitsos, A. Deterministic global optimization with Gaussian processes embedded. *Mathematical Programming Computation* **2021**,

(47) Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Networks* **1989**, *2*, 359–366.

(48) Nentwich, C.; Varela, C.; Engell, S. Optimization of chemical processes applying surrogate models for phase equilibrium calculations. 2019 International Joint Conference on Neural Networks (IJCNN). 2019.

(49) Cozad, A.; Sahinidis, N. V.; Miller, D. C. Learning surrogate models for simulation-based optimization. *AIChE Journal* **2014**, *60*, 2211–2227.

(50) Rasmussen, C. E. *Advanced Lectures on Machine Learning*; Springer Berlin Heidelberg, 2004; pp 63–71.

(51) Yeomans, H.; Grossmann, I. E. A systematic modeling framework of superstructure optimization in process synthesis. *Computers & Chemical Engineering* **1999**, *23*, 709–731.

(52) Yeomans, H.; Grossmann, I. E. Nonlinear disjunctive programming models for the synthesis of heat integrated distillation sequences. *Computers & Chemical Engineering* **1999**, *23*, 1135–1151.

(53) Mencarelli, L.; Chen, Q.; Pagot, A.; Grossmann, I. E. A review on superstructure optimization approaches in process system engineering. *Computers & Chemical Engineering* **2020**, *136*, 106808.

(54) Burre, J.; Bongartz, D.; Mitsos, A. Comparison of MINLP formulations for global superstructure optimization. *Optimization and Engineering* **2022**, Submitted for publication.

(55) Underwood, A. J. V. Fractional distillation of multicomponent mixtures. *Chemical Engineering Progress* **1948**, *44*, 603–614.

(56) Liu, G.; Jobson, M.; Smith, R.; Wahnschafft, O. M. Shortcut Design Method for Columns Separating Azeotropic Mixtures. *Industrial & Engineering Chemistry Research* **2004**, *43*, 3908–3923.

(57) Bongartz, D.; Mitsos, A. Deterministic global optimization of process flowsheets in a reduced space using McCormick relaxations. *Journal of Global Optimization* **2017**, *69*, 761–796.

(58) Najman, J.; Bongartz, D.; Mitsos, A. Linearization of McCormick relaxations and hybridization with the auxiliary variable method. *Journal of Global Optimization* **2021**, *80*, 731–756.

(59) Najman, J.; Bongartz, D.; Mitsos, A. Relaxations of thermodynamic property and costing models in process engineering. *Computers & Chemical Engineering* **2019**, *130*, 106571.

(60) Chachuat, B.; Houska, B.; Paulen, R.; Peric, N.; Rajyaguru, J.; Villanueva, M. E. Set-Theoretic Approaches in Analysis, Estimation and Control of Nonlinear Systems. *IFAC-PapersOnLine* **2015**, *48*, 981–995, https://omega-icl.github.io/mcpp/.

(61) McCormick, G. P. Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems. *Mathematical Programming* **1976**, *10*, 147–175.

(62) Tsoukalas, A.; Mitsos, A. Multivariate McCormick relaxations. *Journal of Global Optimization* **2014**, *59*, 633–662.

(63) J. E. Kelley, J. The Cutting-Plane Method for Solving Convex Programs. *Journal of the Society for Industrial and Applied Mathematics* **1960**, *8*, 703–712.

(64) Smith, E. M.; Pantelides, C. C. Global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering* **1997**, *21*, S791–S796.

(65) Tawarmalani, M.; Sahinidis, N. V. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming*; Springer US, 2002.

(66) Misener, R.; Floudas, C. A. ANTIGONE: Algorithms for coNTinuous / Integer Global Optimization of Nonlinear Equations. *Journal of Global Optimization* **2014**, *59*, 503–526.

(67) Achterberg, T. SCIP: solving constraint integer programs. *Mathematical Programming Computation* **2009**, *1*, 1–41.

Graphical abstract