Deterministic Global Optimization of Multistage Membrane Gas Separation Using Surrogate Models

Marius Hörnschemeyer\textsuperscript{a*}, Christian Kunde\textsuperscript{b}

\textsuperscript{a}Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany
\textsuperscript{b}Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany
hoernschemeyer@mpi-magdeburg.mpg.de

Abstract

This paper deals with deterministic global optimization of multistage membrane gas separation processes that are described by spatially distributed models. The computational tractability of the optimization problem is improved by approximating the spatially distributed models with data-driven surrogates. The resulting problems are solved globally using BARON/GAMS. The binary separation of a mixture containing $\text{CO}_2$ and $\text{CH}_4$ is considered as a case study for multistage membrane gas separation processes. The influence of the feed composition on the globally optimal multistage configurations is studied.

Keywords: gas permeation, multistage, binary separation, deterministic global optimization, surrogate modeling.

1. Introduction

Selective membranes enable the separation of gas mixtures without energy-intensive phase changes. The product purity and recovery demands for gas separations are often beyond the limits of single stage systems. To overcome these limitations, configurations comprising multiple stages of membrane-based gas separation modules and recycles can be employed (Ismail et al., 2015). The optimization of such processes typically leads to the solution of nonlinear problems with continuous and discrete variables. Utilizing deterministic global optimization to solve those mixed-integer nonlinear programs (MINLP) avoids suboptimal local solutions that may otherwise result in poor process design decisions. However, the large number of complex nonlinear expressions and the auxiliary variables make detailed models of spatially distributed processes, e. g. based on finite volume discretization, unfavorable for deterministic global optimization.

In this contribution, replacing a spatially distributed model of a membrane module by a data-driven surrogate is studied as an option to trade some model accuracy for better computational tractability of the global optimization. The required data for the surrogate generation is provided by simulation runs with a fully discretized model of the membrane-based separation process. The developed optimization method is demonstrated for the membrane-based removal of $\text{CO}_2$ from biogas. A parameter study is carried out to investigate the influence of the feed composition on the globally optimal configurations of membrane modules. This way, the findings on multistage binary separations in (Kunde and Kienle, 2018) are extended to membrane-based gas separation.
2. Model description

2.1. Spatially distributed model / Membrane module model

The membrane module model describes the purification of a binary gas mixture due to a selective flow through a membrane. A membrane module operating in countercurrent flow pattern is modeled. The feed flow $N^f_\alpha$ enters the module and is transported along the membrane surface, while a part of it permeates through the membrane. The remaining gas on the retentate side leaves the module via the retentate flow $N^{R}_{\alpha,out}$. The permeated gas is collected at the permeate side, transported in the opposing direction to the bulk flow at the retentate side and leaves via the permeate flow $N^{P}_{\alpha,out}$. The permeating flux $j^m_\alpha$ is modeled according to the solution-diffusion-mechanism (Wijmans and Baker, 1995) assuming a constant permeance $Q_\alpha$ of the membrane. The fast permeating gas component A is enriched at the permeate side, while the slow permeating component B is enriched at the retentate side. Acting as a driving force for the flux

$$j^m_\alpha = Q_\alpha (x^R_\alpha P^R - x^P_\alpha P^P), \quad \alpha = A, B$$

the difference in partial pressure between both sides is determined using ideal gas behavior with mole fractions $x_\alpha$ and constant total pressures at retentate side $P^R$ and permeate side $P^P$. The permeate flow is increased by increasing the membrane area $A^m$. The opposing bulk flows at retentate and permeate side are assumed to be driven only by advection and ideally mixed except in their flow direction. The bulk flows are modeled using mass balances coupled by the permeating flux $j^m_\alpha$ of Eq. (1).

$$\frac{\partial}{\partial z} N^R_\alpha (z) = \frac{-A^m}{L} j^m_\alpha (z), \quad N^R_\alpha (z = 0) = N^R_\alpha$$

$$\frac{\partial}{\partial z} N^P_\alpha (z) = \frac{A^m}{L} j^m_\alpha (z), \quad N^P_\alpha (z = L) = 0$$

A numerical solution for the spatially distributed model is obtained by using the finite volume method as shown in Figure 1. Linear concentrations profiles are assumed in each control volume. The first-order upwind scheme is applied to the differential equation system, which is discretized by $N_z = 150$ control volumes for each flow channel, i.e. the retentate side and the permeate side.

2.2. Superstructure model

The superstructure model is adopted from Kunde and Kienle (2018) and allows to optimize cascades as well as a more complex superstructure called network (see Figure 2). Superstructures with membrane modules $s = 1, \ldots, N_{mo}$ are considered. Both superstructures have one overall feed $N^F_\alpha$, an overall retentate outlet $N^{RR}_\alpha$, enriched in component B, and an overall permeate outlet $N^{PP}_\alpha$, enriched in component A. Each module $s$ has a feed flow $N^F_{\alpha,s}$ and two product flows $N^{RR}_{\alpha,s}$ and $N^{PP}_{\alpha,s}$. No stream splitting is allowed. In the presented case study, minimum product purities $x^R_B$ and $x^P_A$ are required at the overall retentate and permeate outlet.
2.3. Objective function

As proposed by Kunde and Kienle (2018) a general performance indicator is chosen to serve as an objective function. The permeation effort is defined as the ratio of total amount of permeate to overall amount of feed. It is increased by repeated permeation, which requires larger membrane areas and additional gas compression. The permeation effort is therefore also an indicator for economic viability.

3. Surrogate modeling

Data-driven surrogate methods are applied here to obtain surrogates that accurately approximate the original model but require much less computational effort. The original model is treated as a black box with input and output data.

3.1. In- and outputs

In- and output variables are selected from the original model. This determines the properties of the input-output data and thereby affects accuracy and size of the surrogate. Here, we use the feed composition $x^A_f$ and permeate yield $Y^P$ as inputs.

$$x^F_A = \frac{N^P_A}{\sum_{\alpha=A,B} N^P_\alpha} \quad \text{and} \quad Y^P = \frac{\sum_{\alpha=A,B} N^P_\alpha}{\sum_{\alpha=A,B} N^P_\alpha} \quad \text{(4)}$$

At high retentate purities the flux $j^m_\alpha$ permeating through the membrane contains predominantly component B, thus reducing the permeate purity as well as the amount of retentate product. Therefore, combinations of $x^F_A$ and $Y^P$ that lead to a retentate purity well above product specifications are excluded from the data that is used to train the surrogate models. The corresponding input space $\Omega$ is defined as

$$\Omega = \left\{ x^F_A \in [0,1], Y^P \in [0,1] \left| x^F_A \geq \kappa_1 Y^P + \kappa_2, \ x^F_A \geq \kappa_3 Y^P + \kappa_4 \right. \right\} \quad \text{(5)}$$

with linear inequalities that remove most of the undesired input combinations.
The choice of the surrogate output is based on the authors' previous evaluation of a number of alternative model formulations. Only the best performing model formulation from this evaluation is presented here. The model equations for each membrane module read

\[
N_B^{R} = (1 - Y_B^{P})N_B^{P},
\]

\[
N_A^{R} = Y_B^{P}(N_A^{P} + N_B^{P}) - N_A^{P},
\]

\[
N_A^{F} = N_A^{R} + N_B^{P}, \quad \alpha = A, B
\]

with values for \( f \) obtained from the spatially distributed membrane module model. The mapping \( f \) is considered to be the original model and approximated by a surrogate model \( \tilde{f}(x) \), with \( x = (x_A^P, Y_B^P)^T \), constructed as a sum of a trend function \( g(x) \) and a weighted sum of radial basis functions \( \phi(x, x) \).

\[
\tilde{f}(x) = g(x) + \sum_{i=1}^{N_S} w_i \phi(x_i, x)
\]

The trend function \( g(x) \) is calibrated on all sample points \( (x_i, f(x_i)) \) by least square fitting. The remaining difference between \( g(x) \) and \( f(x) \) is approximated by the radial basis functions. We choose a linear radial basis function \( \phi(x_i, x) = \|x - x_i\| \) with a weighted euclidean distance. The distance weights are empirically selected to reduce the maximum approximation error over all sample locations \( x_i \). The reference points \( (x_i, f(x_i)) \), \( i = 1, \ldots, N_S \), of the radial basis function model are selected from the set of sample points \( S \) by an adaptive greedy sampling method that minimizes the maximum approximation error. In each step of the adaptive algorithm, the weights \( w_i \) are obtained by solving a linear equation system resulting from the interpolation conditions at the reference points \( \tilde{f}(x_i) = f(x_i) \).

### 3.2. Surrogate method

A broad range of data-driven surrogate methods is available in the literature. For noise-free data sets, such as generated by deterministic simulation models, Razavi et al. (2012) recommend interpolation methods. In the work at hand a radial basis function model is chosen. The radial basis function model \( \tilde{f}(x) \), with \( x = (x_A^P, Y_B^P)^T \), is constructed as a sum of a trend function \( g(x) \) and a weighted sum of radial basis functions \( \phi(x, x) \).

The error measures \( e_{\text{max}} \) and \( e_{\text{mean}} \) do not account for error propagation to other variables that are important for the overall process. Thus we define additional error measures \( e_{\text{max}}^R, e_{\text{max}}^P, e_{\text{mean}}^R \) and \( e_{\text{mean}}^P \) for the product purities by replacing \( f \) in Eq. (8) with \( x_A^R \) and \( x_A^P \), respectively.
4. Results

The present case study considers biogas upgrading as a separation of CO₂ (component A) from CH₄ (component B). Table 2 comprises all parameters and process requirements necessary to generate sample points and to carry out the superstructure optimization. The input space Ω contains 6262 equally spaced sample points, to which an empirically selected trend function is fitted. To keep the computational effort for parameter studies with many optimization runs feasible, we chose \( N_r = 50 \) reference points. The resulting approximation of the retentate outlet purity for a single module is shown in Figure 3. Table 1 presents the achieved approximation accuracy. Note that errors in process variables, such as the product purities, can be larger than the approximations errors of the surrogate due to error propagation.

A parameter study is carried out to investigate the influence of the overall feed composition \( \mathbf{x}_A \) on the minimized permeation effort and corresponding optimal structures. The results are obtained using the solver BARON 15.9.22. (Tawarmalani and Sahinidis, 2005) and the subsolvers CPLEX and CONOPT in GAMS 24.2.1. Options are kept at default values except setting the relative termination tolerance to \( 1 \times 10^{-3} \). Optimizations of the cascade and network superstructure are performed at thirty values of \( \mathbf{x}_A \). Each superstructure is optimized for 1, 2 and 3 modules. Solution times are in the order of minutes for each run. The lowest permeation efforts \( E_p \) achieved by these superstructures are depicted in Figure 4. There, the roman numerals indicate regions of \( \mathbf{x}_A \) that yield different optimal cascades. In region I, a two-staged cascade with a feed at stage \( s = 2 \) is optimal, whereas in region II, a two-staged cascade with a feed at stage \( s = 1 \) is optimal. The feed compositions in region III allow a single membrane module to be optimal. The permeation effort of optimal network configurations is significantly lower than that of optimal cascades at feed compositions close to the border between region I and II. These reductions are achieved by a single three-staged configuration, shown in Figure 4. At the feed composition \( \mathbf{x}_A = 0.236 \), a maximum relative reduction of 15.9% between an optimal cascade with \( E_p = 0.466 \) and the optimal network configuration with \( E_p = 0.392 \) is observed.

Table 1: Accuracy of the model formulation that includes the surrogate.

<table>
<thead>
<tr>
<th>error</th>
<th>( e_{max} )</th>
<th>( e_{mean} )</th>
<th>( e_{max}^x )</th>
<th>( e_{mean}^x )</th>
<th>( e_{max}^z )</th>
<th>( e_{mean}^z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>0.0119</td>
<td>0.0018</td>
<td>0.0105</td>
<td>0.0041</td>
<td>0.0014</td>
<td>0.0018</td>
</tr>
</tbody>
</table>
Table 2: Process parameters and requirements based on (Scholz et al., 2015).

<table>
<thead>
<tr>
<th>parameter</th>
<th>permeance $Q_a$</th>
<th>permeance $Q_b$</th>
<th>retentate pressure $p^R$</th>
<th>permeate pressure $p^P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>$2.01 \times 10^{-8}$ mol s$^{-1}$ m$^{-1}$ Pa</td>
<td>$3.35 \times 10^{-10}$ mol s$^{-1}$ m$^{-1}$ Pa</td>
<td>$16 \times 10^5$ Pa</td>
<td>$1 \times 10^5$ Pa</td>
</tr>
<tr>
<td>variable</td>
<td>purity retentate $x^R$</td>
<td>purity permeate $x^P$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>domain</td>
<td>$[0.96,1]$</td>
<td>$[0.96,1]$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Conclusion

In this contribution, the application of surrogate modeling allows for extensive parameter studies using deterministic global optimization, despite the high computational cost of the original spatially distributed process model. Results of the parameter study strongly suggest that the findings on multistage binary separation in (Kunde and Kienle, 2018) also extend to membrane-based gas separation: In certain parameter regions, countercurrent cascades are suboptimal compared to the alternative structure depicted in Figure 4. These parameter regions are located where the feed position of optimal cascades changes.

References


