Integrating Different Fidelity Models for Process Optimization: A Case of Equilibrium and Rate-based Extractive Distillation using Ionic Liquids

Ashfaq Iftakher[1], Ty Leonard[1] and M. M. Faruque Hasan[1,2,*]

[1]Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843-3122, USA.

[2]Texas A&M Energy Institute, Texas A&M University, College Station, TX, 77843, USA.

Abstract

This work integrates different fidelity models for optimizing an extractive distillation process. Equilibrium and rate-based approaches are used for modeling phase separation in extractive distillation. The equilibrium-based approach assumes vapor-liquid equilibrium at each stage, while the rate-based approach rigorously considers mass and heat transfer rates, giving more accurate but complex representation. Challenges arise with low-volatility, high-viscosity solvents like ionic liquids, which can drive the system far from equilibrium. We compare equilibrium and rate-based models for separating R-410A, an azeotropic mixture of R-32 and R-125, using an ionic liquid entrainer. Analyzing over 4,300 simulations with dimensionality reduction and topological analysis, we find that while predictions from the two models exhibit similar trends, the overestimation in equilibrium-based purities often leads to infeasible process designs. With these insights, we propose a hybrid optimization method that combines equilibrium/mechanistic model-based optimization with rigorous rate-based evaluation for feasible and optimal designs for ionic liquid-assisted extractive distillation.

Keywords: Hybrid Optimization, Process Design, Extractive Distillation, Ionic Liquid, Refrigerant

*Correspondence should be addressed to M.M. Faruque Hasan at hasan@tamu.edu, Phone: (979) 862-1449.

1 Introduction

Depending on the fidelity and complexity, often more than one mathematical models are developed and solved to predict and optimize a system's behavior.[1] An example is vapor-liquid equilibrium (VLE)-based chemical separations (e.g., absorption, distillation, reactive distillation, and extractive distillation) where different models are regularly used to design unit operations and intensified separation systems.[2,3] These models range from equilibrium-based models to intricate rate-based models, thereby providing a spectrum of predictive capabilities that reflect different assumptions about reality and our ability to handle complexities at various scales. With the advancement in computational hardware and the proliferation of both experimental and model-based data, a critical question arises: How can we exploit the strengths of the individual models to improve both the computational efficiency and the predictive accuracy of process design and optimization schemes? Under this context, hybrid optimization frameworks that integrate mechanistic and data-driven models offer a structured way to utilize the accuracy of detailed models while benefiting from the efficiency of simpler ones. (Note the differences between hybrid modeling approaches and hybrid optimization frameworks. In hybrid modeling, the model is developed by combining fundamental knowledge or first principles with observations or data.[4] In hybrid optimization, we exploit multiple models that may include first principles-based models, data-driven models, and hybrid surrogate models[5,6]).

Hybrid optimization can be particularly useful in the context of extractive distillation processes that combine solvent-based extraction with thermally driven distillation in a single column.[7,8] Extractive distillation has shown great promise to separate azeotropic or close-boiling mixtures that are difficult to separate through conventional distillation. Typically, a solvent, known as an entrainer is added to the mixture which disrupts the azeotrope and selectively absorbs one or more components[9,10], thereby making the separation of the target compound more attainable. Both the equilibrium and rate-based models have been used in the past for designing extractive distillation processes. For example, Benyounes et al.[11] analyzed energy efficiency using equilibrium models for acetone-methanol-water and acetone-chloroform-benzene systems. Errico et al.[12] applied equilibrium models for ethanol-water separation using ethylene glycol. Mendoza and Riascos[13] used a rate-based model to determine the minimum energy required for ethanol dehydration. More recently, Fontana et al.[14] compared both models for anhydrous

bioethanol production, finding the rate-based design required 62% more stages and 8.4% higher costs. Quijada-Maldonado et al.[15] developed a pilot scale process for water-ethanol separation using an ionic liquid (IL) as entrainer, and found that the rate-based model predicted the separation performance of the pilot plant within a 10% relative error. Considerable works exist on the design and analysis of extractive distillation processes, which can be found elsewhere[2,7,16–21]. Optimization of extractive distillation processes mostly involves equilibrium-based model due to its relatively simpler construct (reviews on extractive distillation process optimization and control can be found elsewhere[22,23]). For example, Kossack et al.[9] used mixed-integer models for deterministic optimization of extractive distillation processes for a set of entrainers. Skiborowski et al.[24] reported a hybrid evolutionary-deterministic optimization approach. Monjur et al.[25] used equilibrium models and performed process optimization for refrigerant separation.

The choice of correct models is important to accurately determine the appropriate feed and solvent stages of an extractive distillation column. The equilibrium model considers that every stage is at equilibrium, thereby making it error-prone in the case of subcooled or superheated feed streams that may enter the column far off from the equilibrium. Also, in most cases, the temperatures of the feed and the solvent are different than the temperature of the stage where it is introduced. For such cases, rate-based models may be more appropriate. The rigorous modeling of mass and heat transfer phenomena makes the rate-based approaches more appropriate for modeling complex separations, like those involving azeotropic or close-boiling mixtures and viscous entrainers such as ILs.[26] However, uncertainties related to transport properties limited the use of rate-based models for industrial practice in the past. Therefore, one should be careful about the use of appropriate and experimentally validated mass transfer correlations and transport property models when using the rate-based approach for specific mixtures and packing. Nonetheless, several works concluded that the rate-based model is to be preferred over the equilibrium model since the Murphree tray efficiency is difficult to predict.[27] It is important to capture the non-ideal behavior introduced by entrainers like ILs that exhibit negligible vapor pressure, as well as the nonlinear transport properties that affect the interfacial mass transfer.[26] This significantly affects the performance of computational optimization when the rate-based model is used. Even for process simulation, sophisticated and computationally intensive approaches are needed to avoid convergence issues.[28,29] Therefore, there is an interest in investigating whether the two models can be used

interchangeably or not. Indeed, it is now established for reactive distillation that the predictions from the equilibrium and rate-based models are in general similar.[30] Furthermore, when the vapor-liquid interfacial area is very large, the profiles obtained from the two models become identical. While these are important findings, it is currently unknown whether the same conclusions hold true for other intensified systems, such as extractive distillation that utilize solvents with special thermodynamic and transport properties such as ILs. Even if the predictions are similar, it is important to further elucidate how these similarities can be exploited for process design and optimization.

In this work, we ask the question: How can we exploit the strengths of both the equilibrium and the rate-based models in terms of computational efficiency and prediction accuracy when optimizing an extractive distillation process? If successful, this would lead to efficient model-based process optimization while incorporating accurate predictions. To that end, we hypothesize that the equilibrium model can be used as efficient surrogate or proxy of the original design problem thereby providing a near-optimal estimate of decision variables (preliminary optimization step), while the rate-based model provides an evaluation of the actual process behavior at the optimal design condition (validation step). As we will discuss in detail in Section 2, indeed the equilibrium model can be used to capture the trends in process behavior as functions of key decision variables. In particular, we conduct over 4,300 simulations to compare performance predictions of both the equilibrium and the rate-based models across a wide range of design scenarios. Our observations indicate that, although the trends are similar, the equilibrium model generally provides an overestimation of key performance metrics, such as product purity, which necessitates a further refinement and validation under rate-based considerations.

Building on this, a key novelty of this work then entails to a new hybrid constraint optimization scheme (Section 3). This approach avoids direct optimization of the complex rate-based model, making the overall scheme computationally efficient and more tractable. We are particularly interested in the design and optimization of an extractive distillation process for separating R-410A, an equimass azeotropic refrigerant mixture of R-32 and R-125, using 1-Ethyl-3-methyl-imidazolium thiocyanate ([EMIM][SCN]) as an IL entrainer.

To summarize, the novel contributions of this paper include:

- A systematic approach to compare models in terms of their prediction similarity involving high-dimensional design space,

- Extensive similarity analysis between the equilibrium and the rate-based extractive distillation models using dimensionality reduction techniques, distance metrics, and topological invariants, such as Principal Component Analysis (PCA), Wasserstein distance, Cosine similarity, and Euler Characteristic,
- A mechanistic/data-driven hybrid algorithm for constraint design optimization, which utilizes novel constraint refinement and update rules. Under this setting, the rate-based model is evaluated through a black-box simulator, and the equilibrium model is used as a surrogate. This allows us to navigate a simpler model space using Bayesian optimization while ensuring feasible rate-based design.

The rest of the article is organized as follows: Section 2 describes the process flow diagram, presents a comparison of the simulation results obtained based on equilibrium and rate-based models, and uses several data-analytic metrics for similarity analysis. Based on the analysis and observation, Section 3 presents a hybrid algorithm for process optimization. Lastly, Section 4 includes a summary of key observations and concluding remarks.

2 Comparison of Equilibrium and Rate-based Model Predictions

For modeling phase separation in an extractive distillation column, two approaches are available in the literature: the equilibrium model and the rate-based (nonequilibrium) model.[30] The equilibrium model[31] simplifies the calculation by incorporating the MESH (Material balances, Equilibrium relations, Summation checks, and Heat balances) equations. The vapor-liquid equilibria (VLE) is calculated using activity coefficient models (e.g., NRTL) or through the equation of states (EOS) such as Peng-Robinson. Stage-wise energy balance is performed by considering the latent heat of vaporization and sensible heat changes. To account for the deviation from equilibrium in practical operation in tray and packed columns, the tray efficiency and the height equivalent of a theoretical plate (HETP) are used. The rate-based models[31–33], on the other hand, assume the VLE only at the interface and capture actual stage efficiencies through explicit modeling of mass transfer rates across the vapor-liquid interface using the Maxwell-Stefan equation or the

Fickian model for diffusivities, and two-film mass transfer coefficients that account for both vapor and liquid side resistances.[31] Rate-based models also capture heat transfer effects by incorporating heat transfer resistances and temperature gradients across the phases, which can be significant in highly non-ideal mixtures. In addition to mass and energy balance equations, this approach includes transport between vapor and liquid phases. The equations incorporate both interfacial and bulk concentrations, resulting in a more comprehensive representation of the column behavior. For interested readers, the complete equilibrium and rate-based model equations can be found elsewhere[30,31,34].

In this section, we describe the process configuration for R-410A separation using [EMIM][SCN]. [EMIM] [SCN] is selected because of its prior use and promise as a solvent for the separation of R-410A mixture. It is also one of the most extensively experimentally explored ionic liquids for this application.[35] In subsequent sections, we first conduct a base design simulation to analyze column behavior between equilibrium and rate models, followed by a parametric study for three key design variables: R-410A feed stage, reflux ratio, and IL flow rate. Finally, we perform a simulation-based comparison across various design scenarios to characterize prediction similarity using several data-analytic techniques. All simulations, analysis, and process optimization are conducted on a Dell Windows system with an Intel(R) Core(TM) i7-13700 2.10 GHz CPU and 32 GB of memory.

2.1 Process Configuration and Base Design Simulation

For R-410A separation using [EMIM][SCN], we consider a process configuration shown in Figure 1 that is similar to Monjur et al.[25] and Viar et al.[26] In our work, we consider 100 kg/hr R-410A feed flow rate that is fed to the extractive distillation column at 1 MPa and 278 K.[26] The feed is introduced as a liquid because it enhances separation performance.[36] [EMIM][SCN], which has a density of 1113.9 kg/m^3 and a viscosity of 20.79 mPa.s at 303.15 K,[37] consists of an imidazolium-based cation core and a thiocyanate anion, with one methyl and one ethyl group attached to the cation core. Its high selectivity for R-32 over R-125 makes it suitable for R-410A separation.[26] It is introduced to the extractive distillation column at 1 MPa and 288.15 K. Due to the negligible vapor pressure of IL, the IL feed stage is positioned at the top of the column to ensure maximum gas-liquid contact.[38] The decision to operate the extractive distillation column at 1 MPa follows the recommendation by Viar et al.[26] Since [EMIM][SCN] is R-32 selective, the extractive distillation column separates

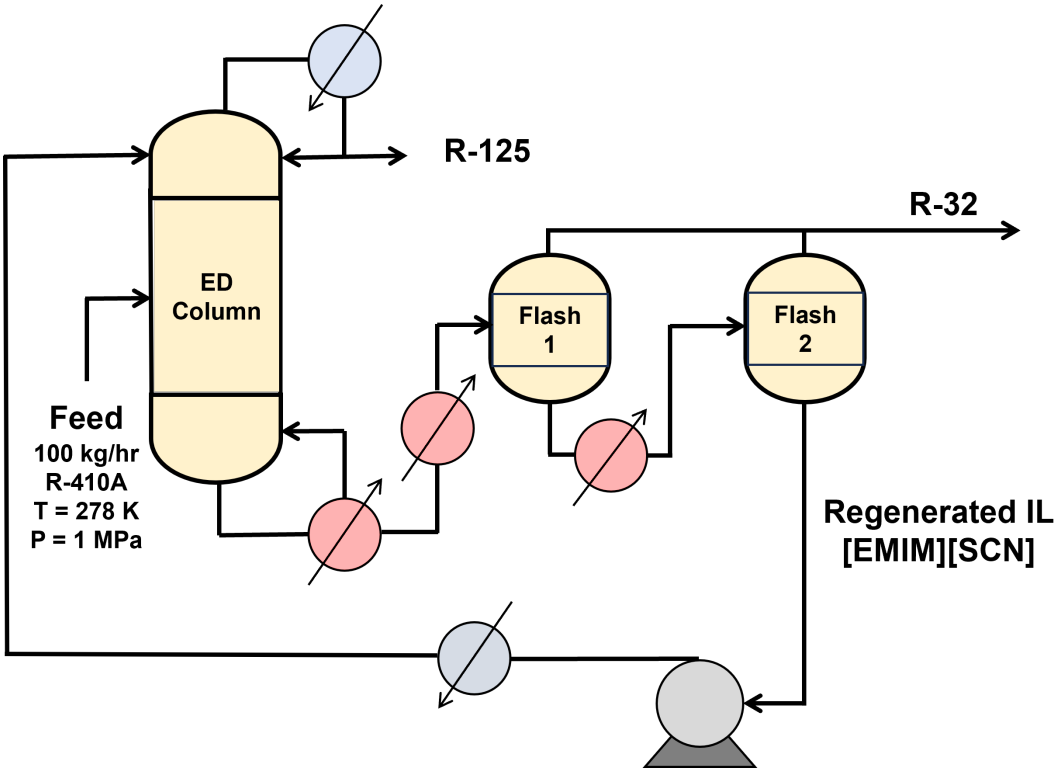

Figure 1: Base process flow diagram of extractive distillation process for R-410A separation using [EMIM][SCN].

high-purity R-125 as a distillate from with a specified flow rate of 50 kg/hr. The IL rich with R32 stream exits at the bottom, which is then fed to the two flash separators. These two sequential flash separators are used for solvent regeneration and high-purity separation of R-32. A recycling and solvent reconditioning system is used for recirculation of the IL to the extractive distillation column. The vapor from both the flash separators are mixed and collected as high-purity R-32. The pressure and temperature of the bottom outlet from the second flash (with high purity [EMIM][SCN]) are adjusted using a pump and a cooler and then recirculated to the extractive distillation column.

We perform the process simulation in Aspen Plus V14. Several pure component and mixture properties are required for process simulation. Tables 1-3 summarize such essential properties like normal boiling point, critical temperature, pressure, molar volume, compressibility factor, acentric factor, and temperature-dependent properties such as heat

Table 1: Molecular and Thermophysical Properties of [EMIM][SCN]

Property	Value
Molar mass - M [g/mol]	169.240
Boiling point - T_b [K]	717.300
Critical temperature - T_c [K]	1013.600
Critical pressure - P_c [bar]	22.260
Critical volume - V_c [cm^3/mol]	666.400
Acentric factor - ω	0.393
Compressibility factor - Z_c	0.176

Table 2: Temperature-dependent correlations for thermophysical properties.

Property	**Units**	**Equation**
Ideal Gas Heat Capacity	kJ/(kmol·K)	$C^g_{p,i}(T) = \alpha_{1i} + \alpha_{2i}T + \alpha_{3i}T^2 + \alpha_{4i}T^3$
Liquid Heat Capacity	kJ/(kmol·K)	$C^l_{p,i}(T) = \alpha_{1i} + \alpha_{2i}T + \alpha_{3i}T^2 + \alpha_{4i}T^3$
Heat of Vaporization	kJ/mol	$\Delta H_{vap_i}(T) = \alpha_{1i}\left(\frac{1-T/T_{ci}}{1-\alpha_{2i}/T_{ci}}\right)^{0.38}$
Liquid Viscosity	mPa·s	$\eta^l_i(T) = \alpha_{1i}\exp\left(\frac{\alpha_{2i}}{T}\right) + \alpha_{3i}$
Liquid Vapor Pressure	atm	$p^l_i(T) = \exp\left(\alpha_{1i} + \frac{\alpha_{2i}}{T+\alpha_{3i}}\right)$
Liquid Surface Tension	mN/m	$\sigma^l_i = \alpha_{1i} + \alpha_{2i}T$

capacity, vapor pressure, viscosity, and surface tension. We use the NRTL activity coefficient model. It is important to ensure that the NRTL model accurately predicts the phase equilibria. For that, we collect experimental solubility data of R-32 and R-125 in [EMIM][SCN] at 303.15 K from Asensio-Delgado et al.[37,44] and simulate the NRTL model using the parameters listed in Table 4. As shown in Figure 2, the NRTL parameters fit the experimental solubility well. Also, at a fixed pressure P, the absorbed amount of R-32 is greater than R-125, suggesting R-32 selectivity of [EMIM][SCN].

Next, we simulate the equilibrium and rate-based models for a base design, with key design variables listed in Table 5. We use the RadFrac block in Aspen Plus V14 for its ability to handle both model types and various packing options. Mellapak 750Y Sulzer-structured packing is chosen for its high surface area, as noted by Meindersma et al.[46] and Sridhar et al.[47]. We note that the rate-based model predictions may significantly deviate based on our choice of mass and heat transfer correlations. In this work, based on the recommendation of Quijada-Maldonado et al.[15], we modeled mass transfer using

Table 3: Coefficients for the property models for [EMIM][SCN].

Property	α_1	α_2	α_3	α_4	**Reference**
Ideal Gas Heat Capacity	-558.580	5.870	-0.018	2×10^{-5}	[39]
Liquid Heat Capacity	-39.425	3.678	-0.014	2×10^{-5}	[39,40]
Heat of Vaporization	142.2	298	-	-	[41]
Liquid Viscosity	3.44×10^{-4}	3304.06	2.474	-	[42]
Liquid Vapor Pressure	-1×10^{35}	-	-	-	
Liquid Surface Tension	79.253	-0.087	-	-	[43]

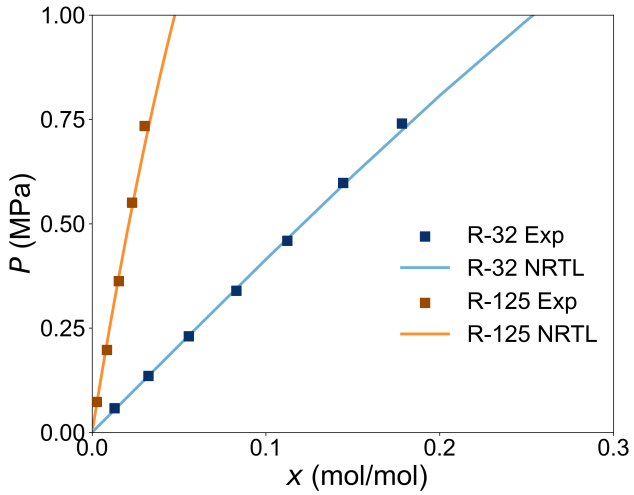

Figure 2: Solubility of R-32 and R-125 at 303.15 K in [EMIM][SCN]. Square points depict experimental data points. Solid lines depict solubility prediction from the NRTL model.

the Bravo et al. correlation[48], and heat transfer using the Chilton-Colburn method[31]. In the base case simulation, the equilibrium-based design meets the purity constraint (R-125 mass purity is 0.9953 in the distillate), while the rate-based design does not achieve the purity (R-125 mass purity is 0.9899 in the distillate). Figure 3 shows the composition and temperature profiles for the base design. Although the profiles are generally similar, the equilibrium model estimates higher temperatures in the stripping section and higher vapor mass fractions in the rectifying section compared to the rate-based model. Next, to achieve a rate-based design that satisfies the purity requirement, the packing height was increased by adjusting the HETP parameter. An HETP value of 0.35 m was determined

Table 4: NRTL parameters.

Compound 1	Compound 2	α	τ_{12}^0	τ_{21}^0	τ_{12}^1	τ_{21}^1	Reference
R-32	[EMIM][SCN]	0.2	0	0	1025.6	-210.04	[44,45]
R-125	[EMIM][SCN]	0.2	187.11	23.863	-51435	-6466.3	[37,45]

Table 5: Process specifications for the base case.

Parameters	Value
R-410A Feed flow rate (kg/hr)	100
IL flow rate (kg/hr)	800
Composition (mass fraction): R-32	0.5
Composition (mass fraction): R-125	0.5
Thermodynamic fluid package	NRTL
Reflux Ratio	2
Distillate Flow Rate (kg/hr)	50
Column Diameter (m)	0.7
Flash-1 Pressure (MPa)	0.1
Flash-2 Pressure (MPa)	0.01
Flash-1 Temperature (K)	313
Flash-2 Temperature (K)	313
R-410A feed stage (N_f)	11
IL feed stage (N_s)	2
Total stage including reboiler and condenser (N_t)	18
Explicit Constraints	
R-125 mass fraction in the outlet	≥ 0.995
R-32 mass fraction in the outlet	≥ 0.995

to be suitable and was used in all subsequent results for rate-based simulations.

2.2 Parametric Study on the Effects of Key Design Variables

To investigate the qualitative and quantitative differences between the equilibrium and rate-based model predictions, we perform a parametric study on three key design variables: R-410A feed stage, reflux ratio, and solvent flow rate. We also quantify process performance by calculating the process energy consumption. We use specific energy consumption (SEC)[49] as the objective function, calculated as shown in Equation 1.

$$SEC = \frac{|E_{reboiler}| + |E_{condenser}| + |E_{flash1}| + |E_{flash2}| + |E_{pump}| + |E_{cooler}|}{Feed\ flow\ rate} \quad (1)$$

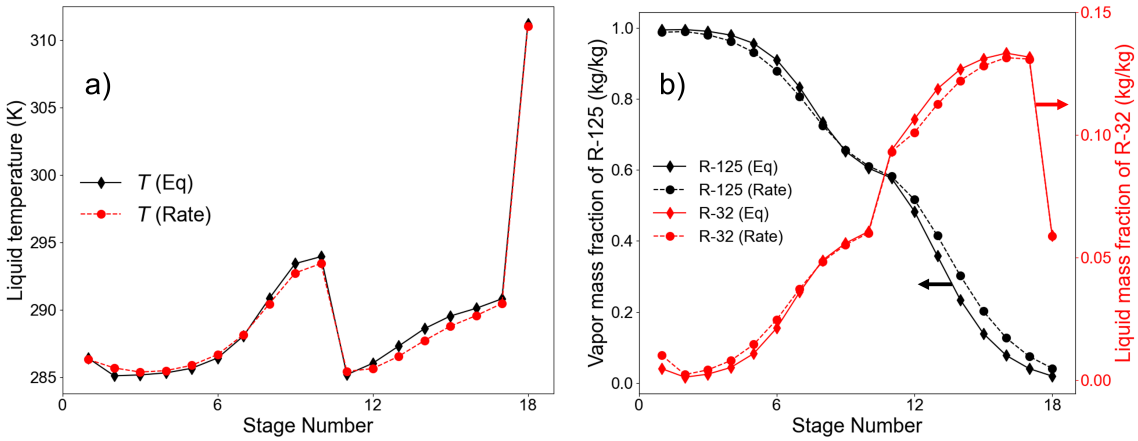

Figure 3: Temperature and concentration profiles along the extractive distillation column for the base design. (a) Temperature, and, (b) Composition.

where, SEC is calculated by first summing all process duties (in kW), and then normalizing by the R-410A feed flow (in kg/hr). Thus, SEC (in kW-hr/kg R-410A) indicates the energy required to process 1 kg of R-410A. While other objective functions like equivalent work (in kj/kg) have previously been used for process optimization[25], our parametric study and the hybrid optimization strategy (described in Section 3) can be adapted to them as well.

We first vary the R-410A feed stage from 3 to 17, keeping other variables fixed to the base design values (see Table 5). Figure 4 shows that the equilibrium model consistently predicts higher mass purity than the rate-based model. Both models identify an optimal feed stage where SEC is minimized and separation purity is maximized. However, the rate-based model does not meet the required purity, while the equilibrium model suggests a design that meets the required purity. Next, we vary the reflux ratio from 0.5 to 7, while keeping other design variables fixed to Table 5. Figure 5 shows that increasing the reflux ratio generally increases mass purity for both models, though the equilibrium model predicts higher purities. For the equilibrium model, purity asymptotically approaches a critical value, while the rate-based model shows an inflection point. That is, mass purity tends to increase until a maximum reflux ratio is achieved, and after that, the mass purity decreases with the increase of reflux ratio. Energy consumption increases with reflux ratio. The rate-based model again fails to meet the required purity, unlike the equilibrium model. Finally, we vary the flow rate of [EMIM][SCN] from 500 kg/hr to 1200 kg/hr, keeping other

variables fixed to Table 5. Figure 6 shows that increasing the solvent flow rate generally increases mass purity and energy consumption for both models. Here, both models suggest designs that meet the required purity, though at different IL flow rates.

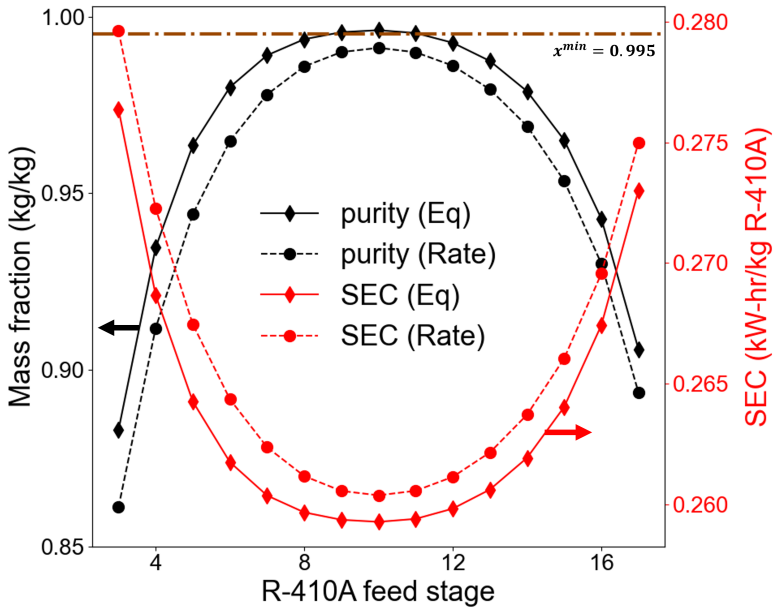

Figure 4: Effect of R-410A feed stage on mass purity and energy consumption (Reflux Ratio = 2, IL solvent to feed ratio = 8).

Based on these studies, we observe a similarity between the equilibrium and rate-based models. Despite similar trends, the rate-based design often fails to achieve the required purity. This leads us to hypothesize that practical design of IL-assisted extractive distillation necessitates validation using the rate-based model. To address the challenge of optimizing the computationally intensive rate-based model, we pose the following question: *Can we exploit the similarity between the equilibrium and rate-based models to guide the design optimization of equilibrium-based IL-assisted extractive distillation, ensuring purity constraints are met under rate-based considerations?* To answer this, next, we explore the design space to validate the similarity between these models across a wide range of conditions.

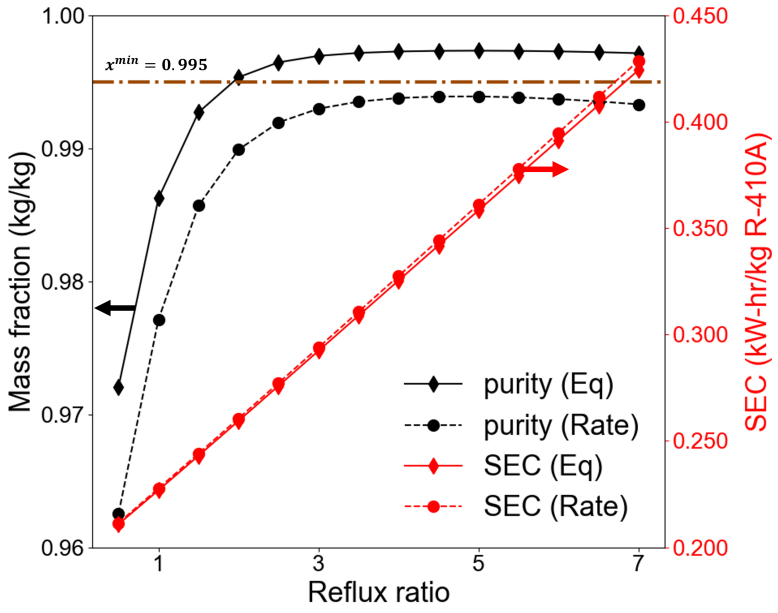

Figure 5: Effect of reflux ratio on mass purity and energy consumption (R-410A feed stage = 11, IL solvent to feed ratio = 8).

2.3 Similarity Analysis

We conduct a space-filling simulation for both the equilibrium and rate-based models. We note that the objective of this space-filling simulation is to understand and characterize the similarity between the two models. These simulation results are not explicitly used for process optimization in Section 3. Table 6 lists the input variables along with their bounds, while R-125 mass purity, R-32 mass purity, and SEC are selected as measured variables. We consider a total of 4383 simulations for which both the equilibrium model and rate-based model successfully converged in Aspen Plus. Figure 7 shows the results for the measured variables, i.e., mass purity and specific energy consumption. Note that the mass purity of R-125 is equal to the mass purity of R-32 (since the distillate flow rate is specified to be 50 kg/hr). We observe that the equilibrium model consistently predicts higher mass purity than the rate-based model. Additionally, energy consumption tends to increase with higher mass purity.

After sampling the design space, we characterize the similarity between the equilibrium

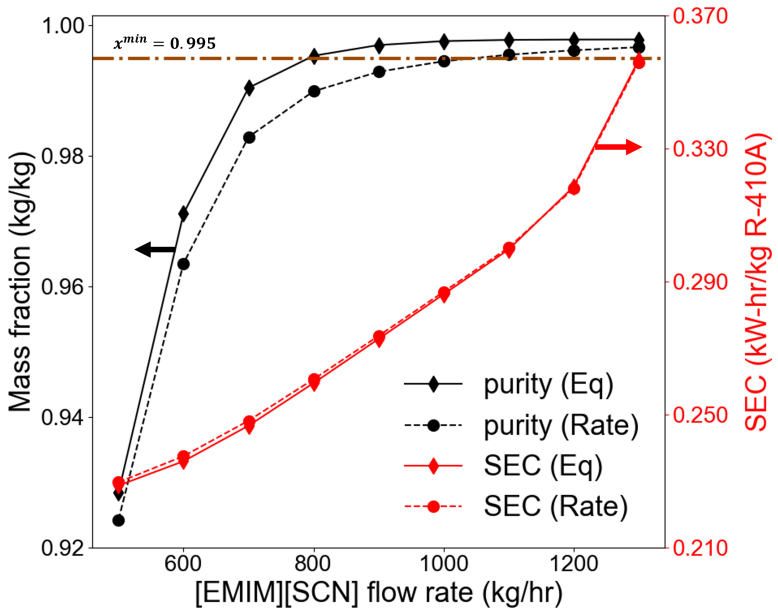

Figure 6: Effect of Ionic Liquid Flow Rate on mass purity and energy consumption (R-410A feed Stage = 11, Reflux Ratio = 2).

and rate-based model datasets. Given the high dimensionality (7 input variables and 3 measured variables), we use principal component analysis (PCA) to project the data into two dimensions for visualization. PCA uses eigenvalue decomposition of the data covariance matrix.[50] to transform high-dimensional data into principal components, capturing the maximum variance. Figure 8 shows the PCA results, highlighting a high degree of overlap between the two models. The explained variance ratios for the first two principal components are 0.32 and 0.15, respectively, indicating that together they capture approximately 47% of the total variance in the dataset. We also compute loadings of PCA on the two principal components (see Section S1 of the Supplementary Material) that verify that reflux ratio, solvent flow rate, and feed stage are the three key features that contribute to the variability.

While PCA uses linear transformation for dimensionality reduction and visualization of the variance, it may not be as effective for highly nonlinear datasets. For example, in our case, PCA captures less than 50% of the total variance, suggesting that it may not fully capture the topological structure of our high-dimensional data. Therefore, to gain

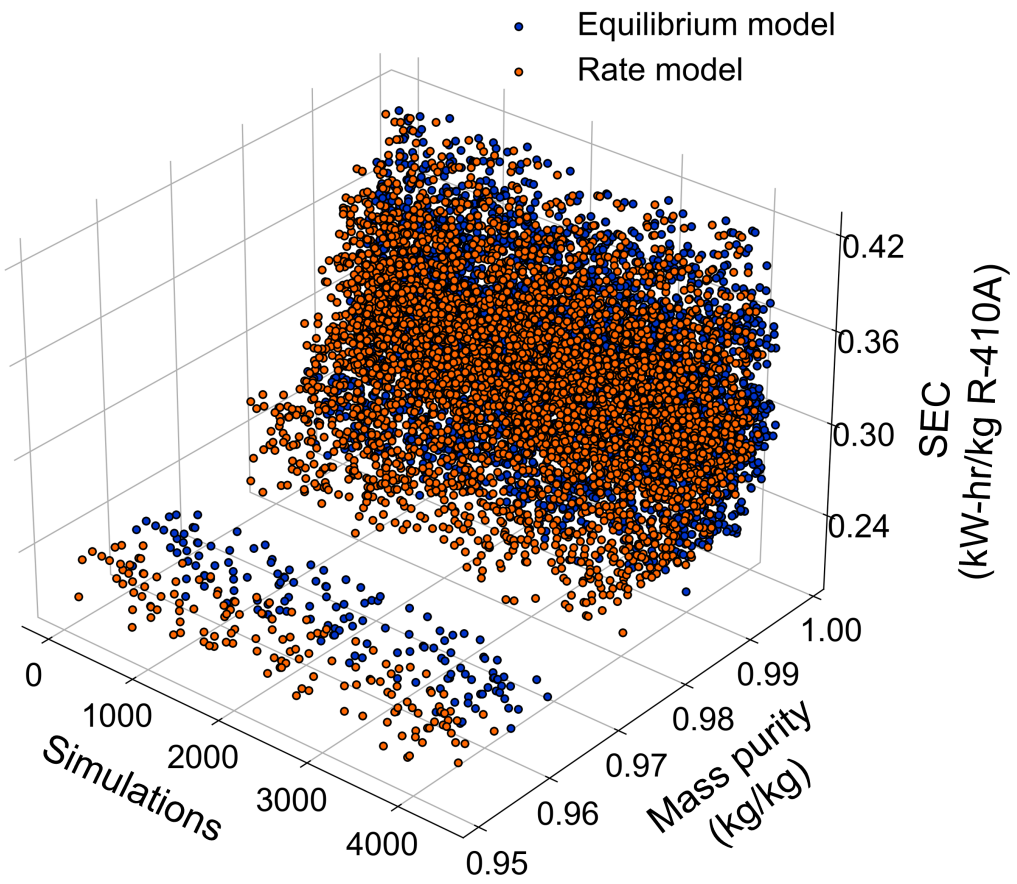

Figure 7: Design space of mass purity and energy consumption between the equilibrium model and the rate model. Simulation numbers are randomized.

deeper insights into the structural similarities and differences between the equilibrium and rate-based models, we employed Euler characteristic[51], a topological invariant that summarizes the shape and connectivity of the structure for understanding non-linear relationships in high-dimensional space - in our case, the 10-dimensional input-output space of our dataset. The Euler characteristic, χ is computed via the Euler-Poincaré formula: $\chi = V - E + F$, where V, E, and F are the numbers of vertices, edges, and faces, respectively. In high-dimensional data analysis, the Euler characteristic captures underlying connectivity through the filtration value, which represents a scale parameter in the Vietoris-Rips complex[52] that determines the degree of connectivity of the data points.

Table 6: Variable bounds for design space exploration

Input Variable	**Bounds**
x_1, R-410A feed stage	$9 \leq x_1 \leq 12$
x_2, Reflux ratio	$2 \leq x_2 \leq 5$
x_3, Solvent flow rate	$700 \leq x_3 \leq 1000$
x_4, Flash 1 Temperature (K)	$310 \leq x_4 \leq 320$
x_5, Flash 1 Pressure (MPa)	$0.1 \leq x_5 \leq 0.2$
x_6, Flash 2 Temperature (K)	$310 \leq x_6 \leq 320$
x_7, Flash 2 Pressure (MPa)	$0.01 \leq x_7 \leq 0.03$

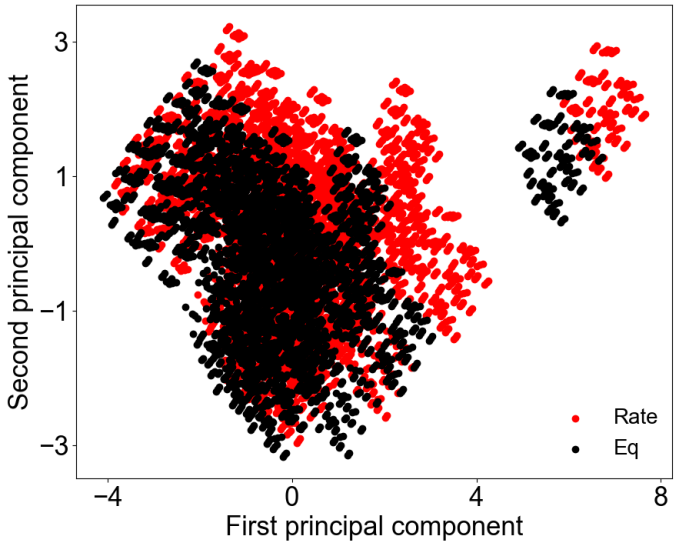

Figure 8: Principal component analysis of the design space between the rate model and the equilibrium model.

The filtration value is a dimensionless number that represents a distance or similarity threshold. At lower filtration values, only closely related design points (those with very similar equilibrium and rate-based model predictions for design variables such as reflux ratio, feed stage, and solvent flow rate) are connected. As the filtration value increases, connections are made between more distant points, revealing larger-scale similarities or differences between the models. The progression of Euler characteristic, χ values with increasing filtration reflects the evolution of complexity and connectivity of the design space

at different scales.[53] Further details on the Euler characteristic and the computational topology in general, can be found elsewhere[53–55]. In our work, we compute the Euler characteristic using the Python Gudhi 3.9.0 library[56].

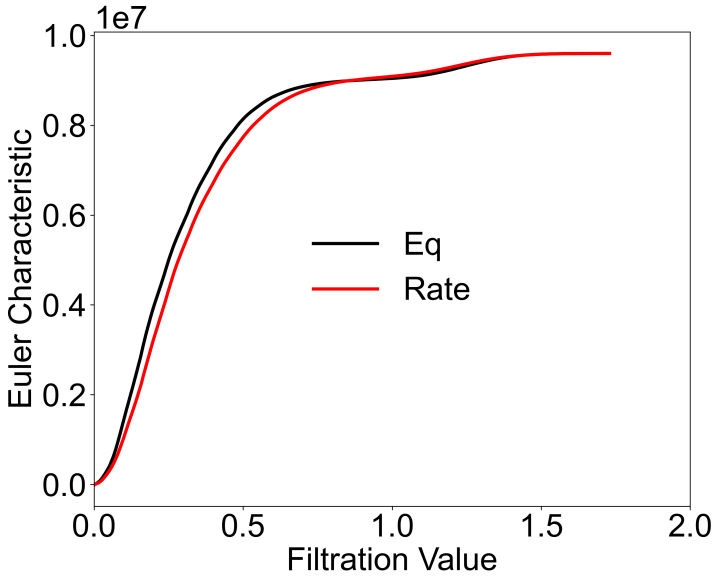

Figure 9: Euler characteristic of the design space between the rate model and the equilibrium model.

As shown in Figure 9, The Euler Characteristic curves for both models follow similar trajectories over the entire range of filtration values. At low filtration values, the high Euler characteristic values suggest many disconnected components, indicating that equilibrium and rate-based models differ in their predictions for small, localized regions of the design space. In practical terms, this indicates that, for certain combinations of process variables (such as at a specific reflux ratio and IL flow rate), the models give different predictions regarding product purity. However, as the filtration value increases, more design points become interconnected, and the Euler Characteristic begin to stabilize, suggesting that the models show similar trends over broader design ranges. This physically represents that both the equilibrium and the rate-based models converge on similar predictions across broader ranges of the design space. For instance, both models might start to agree on the trend of increasing product purity with increasing solvent flow rate. Finally, the Euler Characteristic values essentially become identical for filtration values > 1.5 indicating that

no new topological features are emerging. This stabilization implies that, beyond a certain scale, the overall structure of the design space, as predicted by both models, is similar when viewed on a larger scale. In practical terms, this means that while the models might differ in specific predictions, they are consistent in identifying similar global trends across the entire design space.

Next, we quantify the distributional similarity between the two models by computing the Wasserstein distance. This is a measure of the minimum cost of transforming one distribution of datasets into the other, considering the geometry of the data.[57] Here, we consider only the measured variables and construct 3-dimensional point clouds of equilibrium and rate-based datasets. That is, each point $x_i \in \mathcal{D} \subset \mathbb{R}^3$, where x_i represents a 3-dimensional point in the set $\mathcal{D}$, consists of R-125 mass purity, R-32 mass purity, and SEC. We then compute the Wasserstein distance $= 0.1284$ using the Pot: Python optimal transport library[58]. Since the Wasserstein distance is low, we can infer that the two datasets are distributionally similar. While the Wasserstein distance provides a useful metric for comparing the overall distribution of outputs from the equilibrium and rate-based models, it does not account for point-wise differences in predictions. This means that, even though the Wasserstein distance is low in our work, the models could still produce significantly different outputs for specific inputs. To further investigate this issue, We compute the Cosine similarity[59] between each type of measured variable. Note that cosine similarity directly assesses point-wise directional agreement between the models. We find that the Cosine similarity of R-125 mass purity (kg/kg) between the rate model and the equilibrium model is > 0.99, Cosine similarity of R-32 mass purity (kg/kg) between the rate model and the equilibrium model is > 0.99 and the SEC between the rate model and the equilibrium model is also > 0.99, validating both distributional and point-wise similarity between the equilibrium and rate-based models.

While using PCA, Euler characteristic, Wasserstein distance, and cosine similarity may seem redundant, each serves a distinct purpose to help up understand the similarities between the equilibrium and rate-based models from linear, topological, and geometric perspectives. For example, PCA helped identify that both models predict similar directions of change in the design space, but PCA may miss non-linear interactions between the design variables which could be critical when small variations in design parameters lead to significant differences in purity prediction. In our case, the two principal components of PCA explained less than 50% of the variance, indicating that it doesn't fully capture

the topological structure of our 10-dimensional data. To address this, we used Euler characteristic to summarize the shape and connectivity of the data. Wasserstein distance then quantified geometric differences between the output distributions. Specifically, if the Wasserstein distance was found to be significant, this would indicate that while the models might agree on trends (as seen in PCA), they differ considerably in terms of the absolute values they predict. Lastly, cosine similarity measured directional agreement between the output vectors of the two models, thereby ensuring both distributional and point-wise similarity.

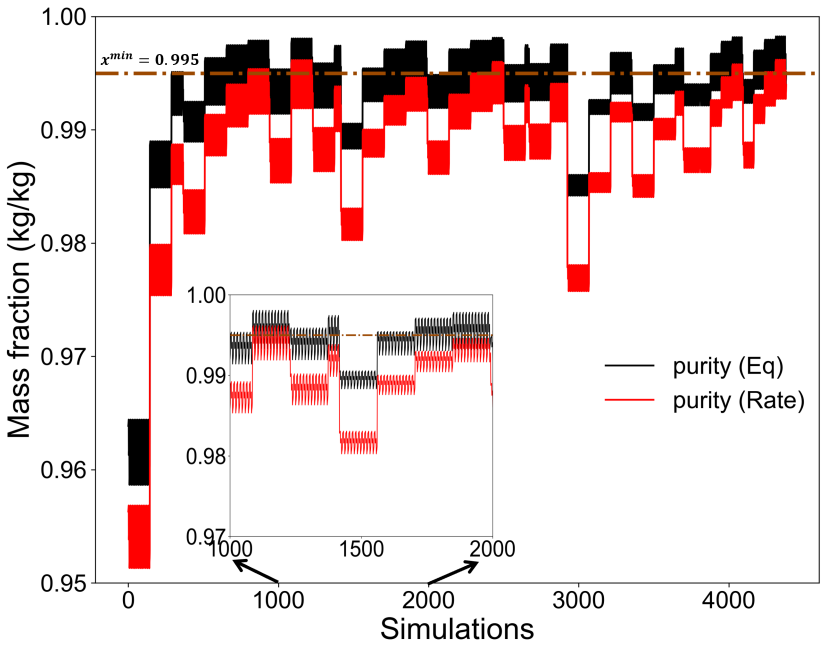

Figure 10: Mass purity for the design space. Mass purity estimates of the equilibrium model tend to be greater than the estimates of the rate model, but they follow similar trend.

Finally, we analyze only the primary measured variable of interest (that is, mass purity) in Figure 10. Several observations are outlined: i) the trends are consistent across the entire domain of interest, ii) ordered combinations of the input parameters lead to nested clustering, and iii) the magnitude of deviation is not consistent, as changing design parameters by the same amount does not necessarily correspond to the same proportional

change in the rate and equilibrium models. This prompts us to sort the equilibrium-based purity in increasing order (see Figure 11) to investigate the relative difference in the infeasible and feasible zones. Here, 'infeasible' and 'feasible' refer to converged Aspen Plus simulations where the product purity was less than 99.5% and greater than 99.5%, respectively.

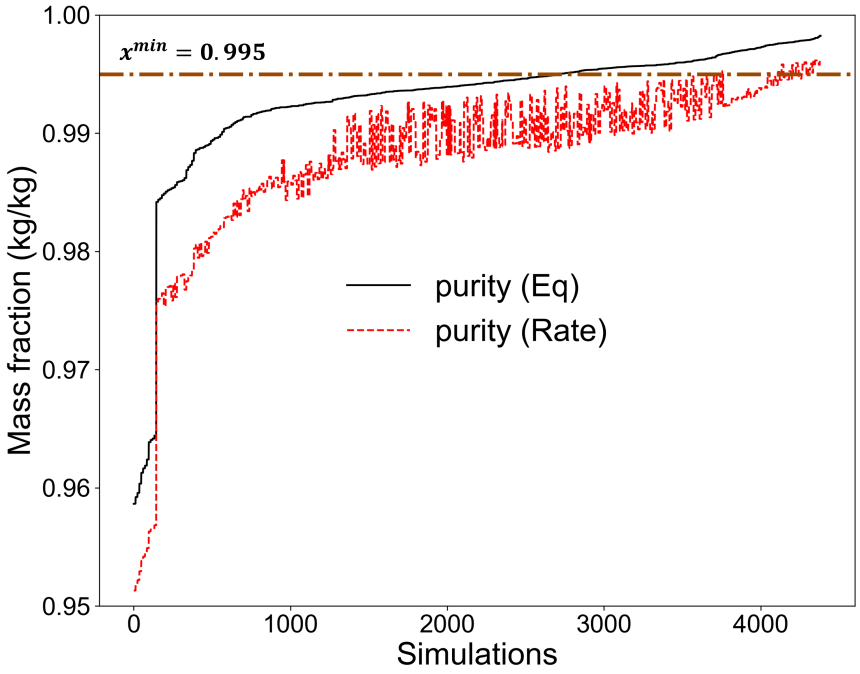

Figure 11: Parametric study of mass purity (simulations are sorted based on the increasing order of equilibrium based mass purity).

From Figure 11, we observe the following: 1) when the rate-based model is feasible, the equilibrium model is also feasible, 2) The mass purity data points for the equilibrium model lie above those of the rate-based model across the entire design space. This analysis supports our claim that equilibrium-based models can guide the optimization for feasible rate-based design. Consider an optimization scheme at iteration k, where the optimized design based on the equilibrium model is infeasible for the rate-based model. We can then use the deviation in mass purity prediction between the equilibrium and rate-based models to tighten the purity constraint for the equilibrium model and re-optimize at iteration $k+1$.

From Figure 11, we can infer that the solution for iteration $k+1$ will lie to the *right* of the solution at iteration k. Thus, as iterations progress, a feasible solution under rate-based consideration will be attained, given that a feasible solution exists within the domain of interest. With this observation, we describe a hybrid optimization approach in Section 3.

3 Hybrid Optimization

Consider the following constrained process optimization problem P:

$$\begin{aligned} \min_{\boldsymbol{x}} \ & f(\boldsymbol{x}) \\ \text{s.t.} \ & \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}, \\ & \boldsymbol{h}(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min}, \end{aligned} \tag{P}$$

where $\boldsymbol{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}$ is a vector of $n(= n_1 + n_2)$ variables, $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$ represents a set of physical constraints ($g_i(x) = 0$, where $i = 1, 2, ..., m$) representing the system behavior or reality, and $\boldsymbol{h}(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min}$ represents a set of externally imposed operational constraints ($h_j(\boldsymbol{x}) \geq \epsilon_j^{\min}$, where $j = 1, 2, ..., q$) on the system performance. Each constraint in $\boldsymbol{h}(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min}$ must be satisfied to ensure that the process performance metrics, such as minimum product purity, maximum allowable emission, etc., are met. The overall system has $(n-m)$ degrees of freedom or decision variables to be optimized. The objective function f is minimized, which gives a measure of the quality of the solution (e.g., overall process energy consumption or cost).

Here, $f(\boldsymbol{x})$, $\boldsymbol{g}(\boldsymbol{x})$, and $\boldsymbol{h}(\boldsymbol{x})$ are functions for which, depending on the assumptions made on the reality, more than one models are available. For brevity, we will also use $f(\boldsymbol{x})$, $\boldsymbol{g}(\boldsymbol{x})$, and $\boldsymbol{h}(\boldsymbol{x})$ to represent a *high-fidelity* model with accurate prediction of the system behavior. We assume that the high-fidelity model can be accessed via an oracle or black-box simulator. Even if the entire model is known, we will treat it as an input-output simulator.

Next, we assume the availability of an equation-oriented *surrogate* model P-S:

$$\begin{aligned} \min_{\boldsymbol{x}} \ & f^s(\boldsymbol{x}) \\ \text{s.t.} \ & \boldsymbol{g}^s(\boldsymbol{x}) = \boldsymbol{0}, \\ & \boldsymbol{h}^s(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min} + \boldsymbol{\delta}, \end{aligned} \tag{P-S}$$

where $\boldsymbol{f^s}$, $\boldsymbol{g^s}$, and $\boldsymbol{h^s}$ are alternative representations $\boldsymbol{f}$, $\boldsymbol{g}$, and $\boldsymbol{h}$, respectively. A key characteristic of P-S is that it provides a fast but approximate estimation of the system's performance measures. The parameter $\boldsymbol{\delta} \in \mathbb{R}^q$ enforces additional tightness in the constraints of the surrogate Problem P-S so that its solution is most likely feasible for the original Problem P.

First, we optimize the surrogate problem P-S. Then, we evaluate P at the solution of P-S, and update parameters $\boldsymbol{\delta}$ based on the violation of the constraints $\boldsymbol{h}(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min}$. This iterative approach allows us to progressively tighten constraint $\boldsymbol{h^s}(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min} + \boldsymbol{\delta}$ in the surrogate space to ensure that a feasible solution in the original space is eventually attained, provided such a solution exists within the domain of interest. We describe these steps in Algorithm 1.

Let $\boldsymbol{x}_k^*$ represent the optimal solution of the Problem P-S at iteration k. The surrogate optimization and the corresponding black-box evaluation cover several cases:

Case 1: $\boldsymbol{h}_k^s(\boldsymbol{x}_k) < \boldsymbol{\epsilon}^{min} + \boldsymbol{\delta}_k, \quad \forall \boldsymbol{x}$, *i.e., the Problem P-S does not have a feasible solution at iteration* k.
In this case, we restore feasibility (Lines 7 and 8 of Algorithm 1), by defining $\delta_j^{'} = \max_{l \in \Omega_j} \delta_{j,l}$, where $\Omega_j = \{p : p < k \quad \& \quad h_{j,p}^s(\boldsymbol{x}_p^*) \geq \epsilon_j^{min} + \delta_{j,p}\}$. Essentially, $\delta_j^{'}$ is the maximum of all such δ_j from previous iterations p ($p < k$), such that the surrogate optimization was feasible. Next, we slightly perturb the j-th operational constraint around $\delta_j^{'}$ by setting $\delta_{j,k+1} = \delta_j^{'} + \epsilon_1$, where ϵ_1 is a small number and we use $\epsilon_1 = 5e^{-4}$.

Case 2: $\boldsymbol{h}_k(\boldsymbol{x}_k^*) < \boldsymbol{\epsilon}^{min}$, *i.e.,* $\boldsymbol{x}_k^*$ *is not a feasible solution to the Problem P as evaluated through black-box simulation of the models in P.*
In this case, we further constrain the operational constraint, $\boldsymbol{h}^s$ (Line 17 of Algorithm 1), of Problem P-S at the next iteration $k+1$, i.e., we impose $\boldsymbol{h}_{k+1}^s(\boldsymbol{x}) \geq \boldsymbol{\epsilon}^{min} + \boldsymbol{\delta}_{k+1}$, where $\boldsymbol{\delta}_{k+1}$ is updated as follows:

$$\boldsymbol{\delta}_{k+1} \leftarrow \left(1 - \boldsymbol{\epsilon}^{min}\right) - \left(1 - \boldsymbol{\delta}_k - \boldsymbol{\epsilon}^{min}\right) \exp - \left[\frac{\boldsymbol{\epsilon}^{min} - \boldsymbol{h}_k(\boldsymbol{x}_k^*)}{1 - \boldsymbol{\epsilon}^{min}}\right] \tag{2}$$

Property 1: $\boldsymbol{\delta}_{k+1} \geq \boldsymbol{\delta}_k$ *for all* k *in Case 2.* This ensures that if infeasibility is encountered in Problem P, then in the next iteration, additional tightness is enforced in Problem P-S.

Algorithm 1 Hybrid optimization scheme for feasible constrained black-box solution

$\boldsymbol{\delta}_0 \leftarrow 0$, $\boldsymbol{S} \leftarrow \emptyset$, $k \leftarrow 0$, $m \leftarrow 0$, $F \leftarrow +\infty$ ▷ Initialization
while $k \leq N^{max} \vee m < 2$ **do**
  $\boldsymbol{x}_k^* = \arg\min$ P-S ▷ Surrogate optimization
  if $\boldsymbol{x}_k^* = \emptyset$ **then** ▷ Surrogate infeasibility
    $m \leftarrow m + 1$
    for $j = 1, 2, ..., q$ **do**
      $\delta_j^{'} = \max_{l \in \Omega_j} \delta_{j,l}$, where $\Omega_j = \{p : p < k \ \ \& \ \ h_{j,p}^s(\boldsymbol{x}_p^*) \geq \epsilon_j^{min} + \delta_{j,p}\}$
      $\delta_{j,k+1} = \delta_j^{'} + \epsilon_1$ ▷ Feasibility restoration
    end for
  else if $\boldsymbol{h}_k^s(\boldsymbol{x}_k^*) \geq \boldsymbol{\epsilon}^{min} + \boldsymbol{\delta}_k$ **then** ▷ Surrogate feasibility
    $m \leftarrow 0$
    Evaluate $f_k(\boldsymbol{x}_k^*), \boldsymbol{g}_k(\boldsymbol{x}_k^*), \boldsymbol{h}_k(\boldsymbol{x}_k^*)$ ▷ Black box sampling
    if $\boldsymbol{h}_k(\boldsymbol{x}_k^*) < \boldsymbol{\epsilon}^{min}$ **then** ▷ Black box infeasibility
      for $j = 1, 2, ..., q$ **do**
        $S_j \leftarrow S_j \cup \{h_{j,k}^s(\boldsymbol{x}_k^*)\}$
      end for
      $\boldsymbol{\delta}_{k+1} \leftarrow \left(1 - \boldsymbol{\epsilon}^{min}\right) - \left(1 - \boldsymbol{\delta}_k - \boldsymbol{\epsilon}^{min}\right) \exp - \left[\frac{\boldsymbol{\epsilon}^{min} - \boldsymbol{h}_k(\boldsymbol{x}_k^*)}{1 - \boldsymbol{\epsilon}^{min}}\right]$ ▷ Update δ
    else if $\boldsymbol{h}_k(\boldsymbol{x}_k^*) \geq \boldsymbol{\epsilon}^{min}$ **then** ▷ Black box feasibility
      $\boldsymbol{\gamma}_k^U = \max(\boldsymbol{S})$
      $\boldsymbol{\delta}_{k+1} \leftarrow \boldsymbol{\gamma}_k^U - \boldsymbol{\epsilon}^{min} + \epsilon_2$
      if $f_k(\boldsymbol{x}_k^*) \geq F$ **then** ▷ Current solution worse than incumbent solution
        Break
      end if
      $F \leftarrow f_k(\boldsymbol{x}_k^*)$ ▷ Update incumbent solution
    end if
  end if
  $k = k + 1$
end while

Proof: Define $\lambda_{j,k} = \frac{\epsilon_j^{min} - h_{j,k}(\boldsymbol{x}_k^*)}{1-\epsilon_j^{min}}$. Observe that, for $j = 1, 2, .., q$; $\delta_{j,k+1} = (1 - \epsilon_j^{min}) - (1 - \epsilon_j^{min})\exp(-\lambda_{j,k}) + \delta_{j,k}\exp(-\lambda_{j,k})$. Since for all k in Case 2, $\lambda_{j,k} \geq 0$, This implies $(1 - \epsilon_j^{min})\exp(-\lambda_{j,k}) \leq (1 - \epsilon_j^{min})$ which implies $\delta_{j,k+1} = (1 - \epsilon_j^{min})(1 - \exp(-\lambda_{j,k})) + \delta_{j,k}\exp(-\lambda_{j,k})$. Since $\delta_{j,k} \leq 1 - \epsilon_j^{min}$, this then follows $\delta_{j,k+1} \geq \delta_{j,k}(1 - \exp(-\lambda_{j,k})) + \delta_{j,k}\exp(-\lambda_{j,k}) = \delta_{j,k}$. ■

Essentially, the refinement of the operational constraint in iteration $k+1$ depends on the constraint violation in iteration k. Since Problem P evaluated at the solution of Problem P-S violates the constraint, $\lambda_{j,k} > 0$. As $\lambda_{j,k}$ increases, $\exp(-\lambda_{j,k})$ decreases, meaning $\boldsymbol{\delta}_{k+1}$ is proportional to the constraint violation - allowing larger adjustments for significant violations and smaller ones as feasibility is approached.

Case 3: $\boldsymbol{h}_k(\boldsymbol{x}_k^*) \geq \boldsymbol{\epsilon}^{min}$, *i.e., the solution of the Problem P-S is feasible for the Problem P.*

In this case, we have obtained a feasible solution to the original problem with high-fidelity model. The next steps in Algorithm 1 (Lines 19 and 20) are added to see if the solution can be improved any further if it is already not a local optimal solution. Specifcially, we relax the purity constraint by defining $\boldsymbol{\gamma}_k^U = \max(\boldsymbol{S})$, where for $j = 1, 2, ..., q$; $S_j = \{h_{j,p}^s(\boldsymbol{x}_p^*) : p < k \ \ \& \ \ h_{j,p}(\boldsymbol{x}_p^*) < \epsilon_j^{min}\}$. The set S_j essentially collects all such values for the $j-$th operational constraint for which the surrogate model P-S is feasible but the corresponding Black-based evaluation is infeasible (Line 15 in Algorithm 1). Then $\boldsymbol{\delta}_{k+1} = \boldsymbol{\gamma}_k^U - \boldsymbol{\epsilon}^{min} + \epsilon_2$, where $\epsilon_2 = 1e^{-4}$.

Termination: The termination criteria for the Algorithm 1 are as follows:

- Incumbent solution comparison: At iteration k, if the current feasible solution is worse than the incumbent solution, the algorithm terminates and returns the current best incumbent solution. Otherwise, the current solution is set as the new incumbent solution, ensuring a monotonically decreasing incumbent solution.
- Successive infeasibility: The algorithm terminates if it encounters two successive infeasible solutions in the surrogate optimization step at iterations $k-1$ and k. This criterion ensures that even after slightly perturbing the purity constraint around $\boldsymbol{\delta}^{'}$,

if feasibility is not restored, the algorithm stops.

- Maximum iterations: The algorithm runs up to a specified maximum number of iterations, N^{max}, ensuring a finite-time solution, i.e., $k \leq N^{max}$.

3.1 Results for the extractive distillation process design under rate-based considerations

We demonstrate the application of Algorithm 1 to achieve a feasible [EMIM][SCN]-assisted extractive distillation process under rate-based considerations. Our key idea is to use the computationally efficient equilibrium model as a surrogate (Problem P-S) for the optimization of the rate-based extractive distillation process (Problem P). This combines the equation-oriented optimization of design and operational decision variables using the equilibrium model (mechanistic) with the simulation-based data-driven optimization using the rate-based model for sampling and validation.

For the case study, the continuous variables are the reflux ratio and solvent flow rate, while the R-410A feed state is an integer variable. The objective function is SEC, where the goal is to achieve an optimized process that corresponds to the design with the least energy consumption. The constraints $\boldsymbol{g}^s(\boldsymbol{x}) = \mathbf{0}$ define the MESH equations for the mass, equilibria and energy balance involving all process units in the flowsheet. The operational constraint $h^s(\boldsymbol{x})$ in this case is the purity constraint for R-32 and R-125. ϵ^{min} is the minimum purity required, which in our work is set to be 0.995. We optimize the equilibrium-based model P-S using a Bayesian Optimization scheme that directly interacts with an Aspen Plus file, where the MESH equations corresponding to the extractive distillation process are inherently built into the equilibrium model. Thus, Algorithm 1 leverages Aspen Plus's built-in capabilities to ensure that all physical constraints including MESH equations are respected during each iteration of the optimization process.

Within the Bayesian optimization scheme, we leverage Gaussian Processes (GP)[60] to predict the mean (μ) and standard deviation (σ) of the objective function, guiding the selection of new sampling points through acquisition functions that balance exploration and exploitation[61–63]. In our implementation, we used GP for prediction and expected improvement as the acquisition function to determine the next sampling point, all in a Python 3 environment[64]. We note that the use of GP in Bayesian optimization does not require the predictions from the equilibrium model to follow a Gaussian distribution.

Table 7: Hybrid optimization for large design space

			P-S optimization					P evaluation	
i	δ	$\epsilon^{min}+\delta$	FS	RR	SF	x_{R125}^{EQ}	SEC^{EQ}	x_{R125}^{Rate}	SEC^{rate}
1	0	0.9950	11	2.2213	944.21	0.9975	0.2860	0.9943	0.2869
2	0.0006	0.9956	10	2.7282	737.01	0.9959	0.2750	0.9905	0.2767
3	0.0030	0.9980	-	-	-	-	-	-	-
4	0.0011	0.9961	11	1.7509	1000	0.9974	0.2778	0.9939	0.2785
5	0.0018	0.9968	11	2.1960	912.16	0.9973	0.2808	0.9937	0.2818
6	0.0025	0.9975	11	2.1880	953.07	0.9975	0.2861	0.9943	0.2869
7	0.0028	0.9978	**11**	**4.2992**	**928.97**	**0.9979**	0.3527	**0.9958**	**0.3546**
8	0.0026	0.9976	11	2.1557	978.28	0.9976	0.2884	0.9946	0.2892
9	0.0027	0.9977	**11**	**4.2992**	**928.97**	**0.9979**	0.3527	**0.9958**	**0.3546**

Instead, GPs define a prior over functions, which is updated with observed data to form a posterior distribution that best fits the data. This method is well-established and widely applied, including in chemical engineering, without needing the underlying data to follow a Gaussian distribution.[61,65] After the Bayesian Optimization of Problem P-S, we evaluate the rate-based model (Problem P) at its solution. Note that the rate-based model is implemented within Aspen Plus that includes detailed mass and energy balance equations, as well as transport property models (given in Table 3) to simulate/evaluate Problem P at the solution of Problem P-S. We would like to note that not every simulation in Aspen Plus may converge. The flowsheet convergence issues often stem from the presence of a recycle loop in the flowsheet and the presence of added nonlinearity due to complex mass transfer and property correlations in the rate-based model. If we encounter cases for which the equilibrium or rate-based flowsheet does not converge, we assign a high penalty for that design condition, so that in the subsequent iterations, the same problematic design condition is not explored. While this procedure allows the algorithm to continue, we admit that it may restrict the search only to the points where the simulations are numerically stable.

We first consider a larger design space with the following optimization variable bounds: R-410A feed stage $[5-15]$, Reflux ratio $[0.5-5]$, and [EMIM][SCN] flow rate $[500-1000]$ kg/hr. The progression of the algorithm and the results for all iterations are given in Table 7. Here, column 1 refers to the iteration number, column 2 refers to the magnitude of constraint refinement, and column 3 refers to the imposed purity constraint on the

equilibrium model. Columns 4-6 refer to the solution obtained by optimizing the equilibrium model. These columns report feed stage, reflux ratio, and solvent flow rate, respectively. Columns 7 and 8 report the purity and total duty predicted by the equilibrium model. Finally, columns 9 and 10 report the rate-based purity and total duty evaluated at the same design condition.

In iteration 1, the equilibrium model satisfies the purity constraint, whereas the rate-based model does not satisfy the minimum purity. Therefore, iteration 1 leads to Case 2. We update $S = \{0.9975\}$. Also, since the rate-based evaluation does not meet the required purity, the purity constraint of the equilibrium model for the next iteration 2 is computed using Equation 2 ($\delta_2 = 0.0006$). At the end of iteration 2, we observe that the equilibrium model satisfies the purity constraint but the rate model does not satisfy the minimum purity. This again leads to Case 2. We then update $S = \{0.9975, 0.9959\}$ and compute $\delta_3 = 0.003$. In iteration 3, we notice that the equilibrium solution is infeasible, that is, no feasible solution has been found in iteration 3 such that the equilibrium model satisfies purity constraint $\epsilon^{min} + \delta_3 \geq 0.998$. This leads to Case 1. So, we restore feasibility by noting that $\delta^{'} = 0.0006$. Therefore, $\delta_4 = 0.0006 + 0.0005 = 0.0011$. The iteration progresses, and in iteration 7, we encounter the first case where the rate-based model is feasible. This leads to Case 3. Therefore, we first update the incumbent solution to be equal to the SEC obtained at iteration 7. Also, $\max(S) = 0.9975$ which implies $\delta_8 = 0.0026$. Finally, the algorithm terminates in 9 iterations after observing that the feasible rate-based design obtained in this iteration does not improve the incumbent solution which was obtained in iteration 7. The algorithm gives the following optimal values of the design variables with feasible rate-based design: R-410A feed stage = 11, Reflux ratio = 4.3, [EMIM][SCN] flow rate = 928.97 kg/hr.

For the hybrid optimization scheme, each iteration required 50 equilibrium-based and 1 rate based process simulations. The scheme converged after 9 iterations, thereby requiring a total of 450 equilibrium-based simulations and 9 rate-based simulations. The overall runtime of the optimization scheme, including all simulations, was $\approx$ 2036 s. To compare, we also performed a Bayesian optimization using only the rate-based model. Interestingly, even after 460 simulations and more than 40 minutes of runtime, the scheme did not find a feasible process condition to achieve the minimum product purity. This suggests the usefulness of the hybrid approach. We also check the sensitivity of this solution against different mass transfer correlations available through Aspen Plus, which are given in Section

Table 8: Hybrid optimization for smaller design space

			P-S optimization					P evaluation	
i	δ	$\epsilon^{min}+\delta$	FS	RR	SF	x_{R125}^{EQ}	SEC^{EQ}	x_{R125}^{Rate}	SEC^{rate}
1	0	0.9950	10	2.1701	781.79	0.9961	0.2625	0.9911	0.2637
2	0.0027	0.9977	**11**	**3.0592**	**985.81**	**0.9978**	0.3194	**0.9958**	**0.3205**
3	0.0012	0.9962	9	2	873.44	0.9966	0.2692	0.9925	0.2700
4	0.0026	0.9976	**11**	**3.3104**	**891.30**	**0.9977**	0.3149	**0.9950**	**0.3164**
5	0.0017	0.9967	9	2	890.98	0.9968	0.2715	0.9929	0.2724
6	0.0028	0.9978	**11**	**2.7728**	**987.91**	**0.9978**	0.3102	**0.9956**	**0.3112**
7	0.0018	0.9968	11	2.3177	842.08	0.9968	0.2755	0.9926	0.2766
8	0.0030	0.998	-	-	-	-	-	-	-
9	0.0029	0.9979	**11**	**4.5297**	**957.47**	**0.9979**	0.3642	**0.9961**	**0.3661**

S2 of the Supplementary Material.

We further test Algorithm 1 on a smaller design space with the following bounds: R-410A feed stage $[9-12]$, Reflux ratio $[2-5]$, and [EMIM][SCN] flow rate $[700-1000]$ kg/hr. For this smaller design, the progression of the algorithm and all the iteration results are given in Table 8. The algorithm terminates at iteration 9 since the feasible rate-based design in iteration 9 is worse than the incumbent solution (obtained in iteration 6). Thus, the optimized design variables are as follows: R-410A feed stage = 11, Reflux ratio = 2.77, [EMIM][SCN] flow rate = 987.91 kg/hr. The process flowsheet corresponding to this optimized design is shown in Figure 12.

4 Conclusions

We performed space-filling simulations to compare the predictions of the equilibrium and the rate-based models for an ionic liquid-assisted extractive distillation process to separate refrigerants R-32 and R-125 from R-410A mixtures. While different data analytic metrics provide different information, all of them strongly indicated that the predictions from both models have similar trends. The differences in absolute values might be a result of using viscous and low-volatile liquid-based as entrainer. Interestingly, we observed that the estimate of the product purities from the equilibrium model tends to be greater than the rate model. This might stem from the gas diffusion rates, dominant transport properties, and other factors, which are not considered in the equilibrium model, thereby leading

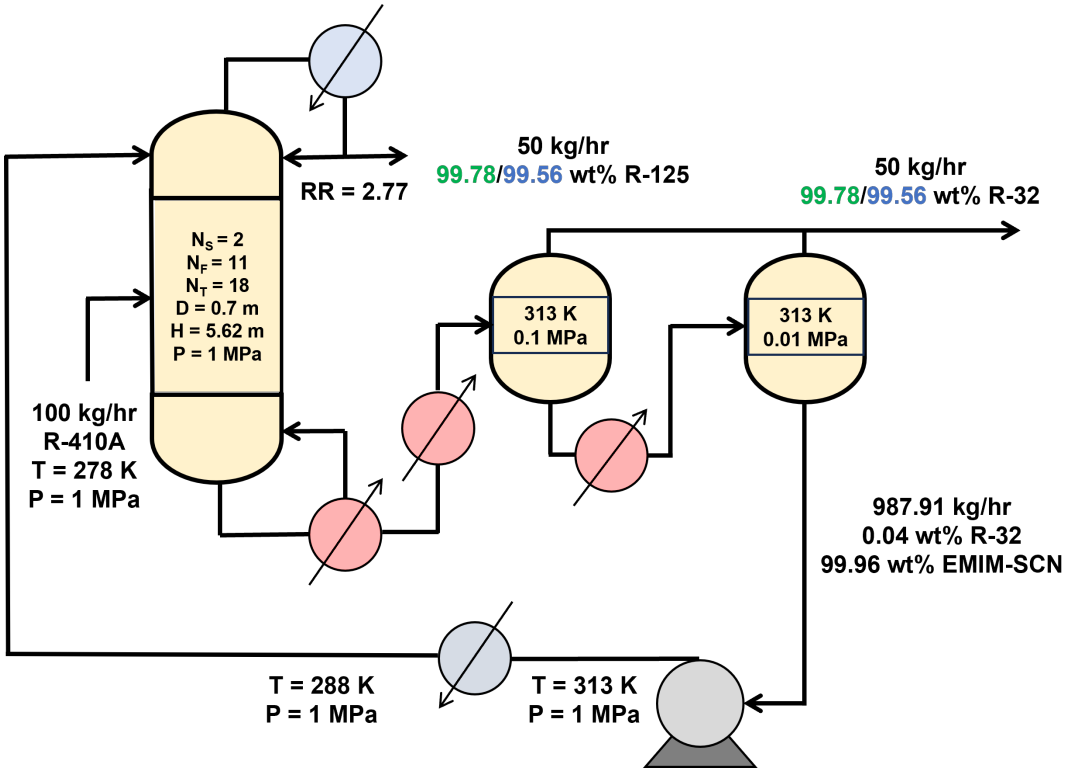

Figure 12: Optimized flowsheet for the smaller design space. Mass purity from the equilibrium-based model is indicated in green, while mass purity from the rate-based model is indicated in blue.

to a potential overestimation of the concentration of R-32 and R-125 in the bulk ionic liquid. The overestimation complicates design under strict purity constraints (e.g., R-32 requires a purity of 99.5 wt% or more). For instance, designs that are feasible based on the equilibrium model can still be infeasible based on rate-based considerations. To that end, we have reported a hybrid optimization scheme that was effective in exploiting the similarity in trends to find feasible process designs with improved energy consumption. In general, we observed that the designs derived from the equilibrium model could serve as excellent initial guesses for practical design purposes, thereby providing benefits through faster convergence and in silico validation.

We used Bayesian optimization for the equilibrium model, but future work could employ global optimization methods for guaranteed optimal flowsheets, although the algorithmic procedure presented here would remain similar. Our study focused on a fixed ionic liquid,

but the qualitative trends predicted by the two models are also expected to be similar for other ionic liquids. Also, the procedure presented in this work is not limited to only extractive distillation but may be extended to integrating multiple fidelity models in various process optimization problems. High-fidelity models can be thought of as Problem P whereas low-fidelity models can serve as Problem P-S However, it may still be necessary to first establish similarity between different fidelity models to ensure a feasible solution. Also, currently, we do not have a convergence guarantee to the optimal solution, and thus, the proposed approach can be viewed as a heuristic approach that yields near-optimal or local solutions.

An interesting, albeit ambitious approach would integrate the proposed hybrid optimization methodology with computer-aided molecular and process design, simultaneously optimizing the solvent choice and process design variables. Lastly, the methodology has the potential to significantly reduce simulation/computation time through optimization, compared to performing many space-filling simulations.

Acknowledgement

The authors gratefully acknowledge support from the NSF CAREER award CBET-1943479 and the NSF EFRI DChem 2029354. Part of the research was conducted with the computing resources provided by Texas A&M High Performance Research Computing.

References

[1] Venkatasubramanian, V. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal* **2019**, *65*.

[2] Tian, Y.; Demirel, S. E.; Hasan, M. M. F.; Pistikopoulos, E. N. An overview of process systems engineering approaches for process intensification: State of the art. *Chemical Engineering and Processing-Process Intensification* **2018**, *133*, 160–210.

[3] Iftakher, A.; Mansouri, S. S.; Nahid, A.; Tula, A. K.; Choudhury, M. S.; Lee, J. H.; Gani, R. Integrated design and control of reactive distillation processes using the driving force approach. *AIChE Journal* **2021**, *67*, e17227.

[4] Sansana, J.; Joswiak, M. N.; Castillo, I.; Wang, Z.; Rendall, R.; Chiang, L. H.; Reis, M. S. Recent trends on hybrid modeling for Industry 4.0. *Computers & Chemical Engineering* **2021**, *151*, 107365.

[5] Caballero, J. A.; Grossmann, I. E. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE Journal* **2008**, *54*, 2633–2650.

[6] Bradley, W.; Kim, J.; Kilwein, Z.; Blakely, L.; Eydenberg, M.; Jalvin, J.; Laird, C.; Boukouvala, F. Perspectives on the integration between first-principles and data-driven modeling. *Computers & Chemical Engineering* **2022**, *166*, 107898.

[7] Doherty, M. F.; Malone, M. F. *Conceptual design of distillation systems*; McGraw-Hill, 2001.

[8] Petlyuk, F. B. *Distillation theory and its application to optimal design of separation units*; Cambridge University Press, 2004.

[9] Kossack, S.; Kraemer, K.; Gani, R.; Marquardt, W. A systematic synthesis framework for extractive distillation processes. *Chemical Engineering Research and Design* **2008**, *86*, 781–792.

[10] Qi, J.; Li, Y.; Xue, J.; Qiao, R.; Zhang, Z.; Li, Q. Comparison of heterogeneous azeotropic distillation and energy-saving extractive distillation for separating the acetonitrile-water mixtures. *Separation and Purification Technology* **2020**, *238*, 116487.

[11] Benyounes, H.; Shen, W.; Gerbaud, V. Entropy flow and energy efficiency analysis of extractive distillation with a heavy entrainer. *Industrial & Engineering Chemistry Research* **2014**, *53*, 4778–4791.

[12] Errico, M.; Rong, B.-G.; Tola, G.; Spano, M. Optimal synthesis of distillation systems for bioethanol separation. Part 1: Extractive distillation with simple columns. *Industrial & Engineering Chemistry Research* **2013**, *52*, 1612–1619.

[13] Mendoza, D. F.; Riascos, C. A. *Computer Aided Chemical Engineering*; Elsevier, 2009; Vol. 27; pp 789–794.

[14] Fontana, M.; Marchesan, A. N.; Maciel Filho, R.; Maciel, M. R. W. Extractive distillation to produce anhydrous bioethanol with choline chloride with urea (1: 2) as a solvent: a comparative evaluation of the equilibrium and the rate-based models. *Chemical Engineering and Processing-Process Intensification* **2021**, *168*, 108580.

[15] Quijada-Maldonado, E.; Aelmans, T.; Meindersma, G. d.; de Haan, A. Pilot plant validation of a rate-based extractive distillation model for water–ethanol separation with the ionic liquid [emim][DCA] as solvent. *Chemical engineering journal* **2013**, *223*, 287–297.

[16] Knapp, J. P.; Doherty, M. F. Minimum entrainer flows for extractive distillation: A bifurcation theoretic approach. *AIChE Journal* **1994**, *40*, 243–268.

[17] Laroche, L.; Bekiaris, N.; Andersen, H. W.; Morari, M. The curious behavior of homogeneous azeotropic distillation—implications for entrainer selection. *AIChE journal* **1992**, *38*, 1309–1328.

[18] Levy, S. G.; Doherty, M. F. Design and synthesis of homogeneous azeotropic distillations. 4. Minimum reflux calculations for multiple-feed columns. *Industrial & engineering chemistry fundamentals* **1986**, *25*, 269–279.

[19] Laroche, L.; Andersen, H.; Morari, M.; Bekiaris, N. Homogeneous azeotropic distillation: comparing entrainers. *The Canadian Journal of Chemical Engineering* **1991**, *69*, 1302–1319.

[20] Wahnschafft, O. M.; Westerberg, A. W. The product composition regions of azeotropic distillation columns. 2. Separability in two-feed columns and entrainer selection. *Industrial & Engineering Chemistry Research* **1993**, *32*, 1108–1120.

[21] Iftakher, A.; Monjur, M. S.; Hasan, M. F. An Overview of Computer-aided Molecular and Process Design. *Chemie Ingenieur Technik* **2023**, *95*, 315–333.

[22] Gerbaud, V.; Rodriguez-Donis, I.; Hegely, L.; Lang, P.; Denes, F.; You, X. Review of extractive distillation. Process design, operation, optimization and control. *Chemical Engineering Research and Design* **2019**, *141*, 229–271.

[23] Sun, S.; Lü, L.; Yang, A.; Shen, W.; others Extractive distillation: Advances in conceptual design, solvent selection, and separation strategies. *Chinese Journal of Chemical Engineering* **2019**, *27*, 1247–1256.

[24] Skiborowski, M.; Rautenberg, M.; Marquardt, W. A hybrid evolutionary–deterministic optimization approach for conceptual design. *Industrial & engineering chemistry research* **2015**, *54*, 10054–10072.

[25] Monjur, M. S.; Iftakher, A.; Hasan, M. M. F. Separation process synthesis for high-gwp refrigerant mixtures: Extractive distillation using ionic liquids. *Industrial & Engineering Chemistry Research* **2022**, *61*, 4390–4406.

[26] Viar, M.; Asensio-Delgado, S.; Pardo, F.; Zarca, G.; Urtiaga, A. In the quest for ionic liquid entrainers for the recovery of R-32 and R-125 by extractive distillation under rate-based considerations. *Separation and Purification Technology* **2023**, *324*, 124610.

[27] Lee, J.-H.; Dudukovic, M. A comparison of the equilibrium and nonequilibrium models for a multicomponent reactive distillation column. *Computers & chemical engineering* **1998**, *23*, 159–172.

[28] Javaloyes-Anton, J.; Kronqvist, J.; Caballero, J. A. Simulation-based optimization of distillation processes using an extended cutting plane algorithm. *Computers & Chemical Engineering* **2022**, *159*, 107655.

[29] Pattison, R. C.; Baldea, M. Equation-oriented flowsheet simulation and optimization using pseudo-transient models. *AIChE Journal* **2014**, *60*, 4104–4123.

[30] Peng, J.; Lextrait, S.; Edgar, T. F.; Eldridge, R. B. A comparison of steady-state equilibrium and rate-based models for packed reactive distillation columns. *Industrial & engineering chemistry research* **2002**, *41*, 2735–2744.

[31] Taylor, R.; Krishna, R. *Multicomponent mass transfer*; John Wiley & Sons, 1993; Vol. 2.

[32] Sundmacher, K.; Hoffmann, U. Development of a new catalytic distillation process for fuel ethers via a detailed nonequilibrium model. *Chemical Engineering Science* **1996**, *51*, 2359–2368.

[33] Higler, A.; Taylor, R.; Krishna, R. The influence of mass transfer and mixing on the performance of a tray column for reactive distillation. *Chemical engineering science* **1999**, *54*, 2873–2881.

[34] Pradhan, S.; Kannan, A. Simulation and analysis of extractive distillation process in a valve tray column using the rate based model. *Korean Journal of Chemical Engineering* **2005**, *22*, 441–451.

[35] Baca, K. R.; Al-Barghouti, K.; Wang, N.; Bennett, M. G.; Matamoros Valenciano, L.; May, T. L.; Xu, I. V.; Cordry, M.; Haggard, D. M.; Haas, A. G.; others Ionic Liquids for the Separation of Fluorocarbon Refrigerant Mixtures. *Chemical Reviews* **2024**,

[36] Finberg, E. A.; Shiflett, M. B. Process Designs for Separating R-410A, R-404A, and R-407C Using Extractive Distillation and Ionic Liquid Entrainers. *Industrial & Engineering Chemistry Research* **2021**, *60*, 16054–16067.

[37] Asensio-Delgado, S.; Viar, M.; Pardo, F.; Zarca, G.; Urtiaga, A. Gas solubility and diffusivity of hydrofluorocarbons and hydrofluoroolefins in cyanide-based ionic liquids for the separation of refrigerant mixtures. *Fluid Phase Equilibria* **2021**, *549*, 113210.

[38] Chen, H.-H.; Chen, M.-K.; Chen, B.-C.; Chien, I.-L. Critical assessment of using an ionic liquid as entrainer via extractive distillation. *Industrial & Engineering Chemistry Research* **2017**, *56*, 7768–7782.

[39] Štejfa, V.; Rohlíček, J.; Červinka, C. Phase behaviour and heat capacities of selected 1-ethyl-3-methylimidazolium-based ionic liquids II. *The Journal of Chemical Thermodynamics* **2021**, *160*, 106392.

[40] Ge, R.; Hardacre, C.; Jacquemin, J.; Nancarrow, P.; Rooney, D. W. Heat capacities of ionic liquids as a function of temperature at 0.1 MPa. Measurement and prediction. *Journal of Chemical & Engineering Data* **2008**, *53*, 2148–2153.

[41] Kabo, G. J.; Paulechka, Y. U.; Zaitsau, D. H.; Firaha, A. S. Prediction of the enthalpies of vaporization for room-temperature ionic liquids: Correlations and a substitution-based additive scheme. *Thermochimica acta* **2015**, *609*, 7–19.

[42] Freire, M. G.; Teles, A. R. R.; Rocha, M. A.; Schroder, B.; Neves, C. M.; Carvalho, P. J.; Evtuguin, D. V.; Santos, L. M.; Coutinho, J. A. Thermophysical characterization of ionic liquids able to dissolve biomass. *Journal of Chemical & Engineering Data* **2011**, *56*, 4813–4822.

[43] Zhang, Q.; Li, M.; Zhang, X.; Wu, X. The thermodynamic estimation and viscosity, electrical conductivity characteristics of 1-alkyl-3-methylimidazolium thiocyanate ionic liquids. *Zeitschrift für Physikalische Chemie* **2014**, *228*, 851–867.

[44] Asensio-Delgado, S.; Pardo, F.; Zarca, G.; Urtiaga, A. Enhanced absorption separation of hydrofluorocarbon/hydrofluoroolefin refrigerant blends using ionic liquids. *Separation and Purification Technology* **2020**, *249*, 117136.

[45] Asensio-Delgado, S.; Pardo, F.; Zarca, G.; Urtiaga, A. Absorption separation of fluorinated refrigerant gases with ionic liquids: Equilibrium, mass transport, and process design. *Separation and Purification Technology* **2021**, *276*, 119363.

[46] Meindersma, G. W.; Quijada-Maldonado, E.; Aelmans, T. A.; Hernandez, J. P. G.; de Haan, A. B. *Ionic Liquids: Science and Applications*; ACS Publications, 2012; pp 239–257.

[47] Sridhar, P.; Araujo, J.; Rodrigues, A. Modeling of vanillin production in a structured bubble column reactor. *Catalysis today* **2005**, *105*, 574–581.

[48] Bravo, J. L. Mass transfer in gauze packings. *Hydrocarbon processing* **1985**, *64*, 91–95.

[49] Kwon, H.; Do, T. N.; Kim, J. Energy-efficient liquid hydrogen production using cold energy in liquefied natural gas: Process intensification and techno-economic analysis. *Journal of Cleaner Production* **2022**, *380*, 135034.

[50] Jolliffe, I. T. Principal component analysis. *Technometrics* **2003**, *45*, 276.

[51] Smith, A.; Zavala, V. M. The Euler characteristic: A general topological descriptor for complex data. *Computers & Chemical Engineering* **2021**, *154*, 107463.

[52] Zomorodian, A. Fast construction of the Vietoris-Rips complex. *Computers & Graphics* **2010**, *34*, 263–271.

[53] Edelsbrunner, H.; Harer, J. L. *Computational topology: an introduction*; American Mathematical Society, 2022.

[54] Smith, A.; Runde, S.; Chew, A. K.; Kelkar, A. S.; Maheshwari, U.; Van Lehn, R. C.; Zavala, V. M. Topological analysis of molecular dynamics simulations using the euler characteristic. *Journal of Chemical Theory and Computation* **2023**, *19*, 1553–1567.

[55] Smith, A.; Laubach, B.; Castillo, I.; Zavala, V. M. Data analysis using Riemannian geometry and applications to chemical engineering. *Computers & Chemical Engineering* **2022**, *168*, 108023.

[56] Project, G. GUDHI: Simplicial Complexes and Persistent Homology Library. https://pypi.org/project/gudhi/, 2024.

[57] Villani, C.; others *Optimal transport: old and new*; Springer, 2009; Vol. 338.

[58] Flamary, R.; Courty, N.; Gramfort, A.; Alaya, M. Z.; Boisbunon, A.; Chambon, S.; Chapel, L.; Corenflos, A.; Fatras, K.; Fournier, N.; others Pot: Python optimal transport. *Journal of Machine Learning Research* **2021**, *22*, 1–8.

[59] Xia, P.; Zhang, L.; Li, F. Learning similarity with cosine similarity ensemble. *Information sciences* **2015**, *307*, 39–52.

[60] MacKay, D. J. *Information theory, inference and learning algorithms*; Cambridge university press, 2003.

[61] Jones, D. R.; Schonlau, M.; Welch, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global optimization* **1998**, *13*, 455–492.

[62] Frazier, P. I. A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811* **2018**,

[63] Gandhi, A.; Zantye, M. S.; Hasan, M. F. Cryogenic energy storage: Standalone design, rigorous optimization and techno-economic analysis. *Applied Energy* **2022**, *322*, 119413.

[64] Van Rossum, G.; Drake, F. L.; others *Python reference manual*; Centrum voor Wiskunde en Informatica Amsterdam, 1995; Vol. 111.

[65] Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* **2015**, *104*, 148–175.