

SUPPORTING CHEMICAL PROCESS DESIGN UNDER UNCERTAINTY

A. Wechsung^{1,2}, J. Oldenburg¹, J. Yu¹ and A. Polt^{1*}

¹BASF SE, 67056 Ludwigshafen, Germany.

E-mail: jan.oldenburg@basf.com; joanna.yu@basf.com; axel.polt@basf.com

²Currently: Massachusetts Institute of Technology,

Department of Chemical Engineering, 66-363, 77 Massachusetts Ave,
Cambridge, MA 02139, USA.

E-mail: awechsun@mit.edu

(Submitted: December 10, 2009 ; Revised: July 29, 2010 ; Accepted: July 30, 2010)

Abstract - A major challenge in chemical process design is to make design decisions based on partly incomplete or imperfect design input data. Still, process engineers are expected to design safe, dependable and cost-efficient processes under these conditions. The complexity of typical process models limits intuitive engineering estimates to judge the impact of uncertain parameters on the proposed design. In this work, an approach to quantify the effect of uncertainty on a process design in order to enhance comparisons among different designs is presented. To facilitate automation, a novel relaxation-based heuristic to differentiate between numerical and physical infeasibility when simulations do not converge is introduced. It is shown how this methodology yields more details about limitations of a studied process design.

Keywords: Chemical process design; Process modeling; Simulation and optimization; Uncertainty; Process performance measures; Convergence; Process Simulators.

INTRODUCTION

Process engineers are expected to design safe, dependable and cost-efficient processes. During the stage of conceptual process design, most design decisions are yet to be made, as the design will become more detailed with each decision. Each decision reduces the degrees of freedom of the process design. Initially, as the engineer begins to sketch the flowsheet, there are an incredibly large number of process alternatives. With each decision made, the designed process slowly assumes shape. Especially at an early stage of process design, each decision limits the range of future choices. Furthermore, each decision reduces the scope of the impact of future decisions. Thus, at the early stages of process design, one sets course for the future process. But one steers ahead while facing large

uncertainties, as typically detailed information on some aspects of the proposed flowsheet is still scarce. It is important to provide the decision-maker with as much guidance as possible to support his/her difficult task. Keeping in mind that one of the goals is to design cost-effective processes, when deciding between different flowsheets, the overall cost associated with a flowsheet is an important metric. It is necessary to provide costing information at every step as the process is designed.

Additional difficulty stems from the complexity inherent to technical systems. Generally speaking, chemical processes are complex networks. Modifications to one unit operation at one location may propagate through the network and its feedback loops as represented by, e.g., recycle streams to result in unforeseen consequences at very different units of the flowsheet. Experienced process

*To whom correspondence should be addressed

This is an extended version of the manuscript presented at the PSE 2009 - 10th International Symposium on Process Systems Engineering, 2009, Salvador, Brazil, and published in Computer Aided Chemical Engineering, vol. 27, p. 939-944.

engineers may be able to identify such chains of cause and effect, but they are typically not intuitive.

Common to early stages of process design is the lack of certain information. When work is begun on a new process design, research has identified a promising path and early tests on lab and pilot plants have provided first estimates of the necessary data. But the lack of experience gained from operating a production size plant and the presence of considerable error bars on the data obtained, contribute to an additional challenge. When uncertainty of a subset of model-determining parameters is present and incorporated in the modeled process, the complexity of the model increases further. In the end, impact of changes to underlying assumptions of the process model is even less intuitive—even to an experienced process engineer. Hence, providing engineers with a systematic method to foster understanding of implications of uncertain parameters on their current design will facilitate better decision-making in process design.

Uncertainties are introduced in process design in many ways. Insufficient knowledge about reaction pathways and kinetics contributes to uncertainty just like limited thermodynamic data for chemical components does; lack of experience when performing a scale-up with novel process equipment presents a source of uncertainty as does hard to predict fouling in heat exchangers, distillation columns and other unit operations. When existing processes are modified, intimate knowledge of the process is available to the decision-maker and one expects that uncertainty is reduced. But even in this case, uncertainty can stem from varying raw material purities or from inadequate cost estimates for feedstock.

As briefly mentioned, one decides between different process designs, which achieve the technical requirements, by picking the most economical one. Commonly, commercial deterministic process simulation and/or optimizing software (e.g., Aspen Plus) is used to design processes beginning at very early stages of the design process. Thus, it is necessary to incorporate estimated equipment capital cost as well as cost of process streams into process simulators, which are used to optimize flowsheets. The process engineer is provided with this information easily accessible alongside with the technical data in process simulators. Thus, tools primarily designed to aid in solving engineering problems can support the design of cost efficient processes. This possibility is, for example, implemented in i-TCM (Intelligent Total Cost Minimization), a program which allows a multiparameter process optimization of a plant's total costs using Aspen-EO-Optimizer (Wiesel and Polt, 2006).

In the light of recent developments, achieving the most economic design is even more important. The most ubiquitous feedstocks to the chemical industry, oil and gas, have increased significantly in price while demand has slumped recently, which sharply contrasts with previously experienced years of stable prices in the face of economic growth. Although such unexpected scenarios can never be completely ruled out, process design is not expected to overcompensate for such unforeseen risks or deviations from design conditions. A very common measure to account for risk stemming from uncertainty in design assumptions and from uncertainty due to prediction of uncertainty is overdesign. Though it may be a simple solution to the problem, it is a costly one, too. Nowadays, the designed process is expected to be less accommodating using this costly means of equipment overdesign.

In this contribution, a method is presented that supports the process engineer in the presence of uncertainty using the principles of Monte Carlo methods. To compare different processes, results from many simulations are combined to a quantitative measure stating how well a given design copes with uncertainty. To increase the number of converged simulations, which is a major obstacle of these methods, relaxations of process constraints are introduced.

The present contribution is organized as follows: First, different approaches presented in the literature are reviewed. Second, the dependability analysis is presented. Third, the relaxation-based approach used to increase the number of converged simulations is described. Fourth, a case study will illustrate the application of the method. Fifth, the paper is summarized and possible extensions are discussed.

CHEMICAL PROCESS DESIGN UNDER UNCERTAINTY

Literature Review

When all parameters to a process are known exactly, the optimal design for a given process can be obtained by performing a single optimization. In case of uncertain parameters with a known probability distribution, computing the optimal design is more involved as design of a chemical process turns into a two-stage problem (Malik and Hughes, 1979). This problem can be seen as an example of a stochastic problem. First, design

decisions are made without knowing what value the stochastic variables will assume. Then, operating decisions are made after the stochastic variables took a value. The modeling assumption is that random variables exist in the model. During the first stage decisions, the values of these random variables are not known and need to be characterized by probability density functions. After the first stage decisions have been made, a particular outcome for each of the random variables is realized and the second stage decisions can be made with the precise knowledge.

Translated to the problem of process design, the first stage decisions are made during the design phase. Then, the uncertain parameters are only known to lie within some bounds, possibly characterized by mean and variance or similar measures. The decisions to be made are design decisions, e.g., deciding between process alternatives and sizing process equipment. The second stage corresponds to the operating phase. Now, the plant has been built and is operated. All uncertainties are revealed and the goal of decisions made at this stage is to find the optimal operating conditions.

Solutions of stochastic programs are found by repeatedly solving the second stage problem with different realizations of the random variables. These problems are called scenarios. The outcome of each scenario is tracked and weighted according to its probability of occurrence. Based on the current knowledge about the process, different options are considered. The design decisions are made by weighing the likelihood of all outcomes happening, instead of just looking for the worst case, which lead to costly overdesign. These methods rely on the possibility to solve the problems posed by scenarios efficiently as a large number of scenarios are to be considered.

Expected values will have to be computed numerically, e.g., using stochastic techniques, leading to the repeated requirement to solve the inner optimization problem that selects for the operating conditions. Hence, determining an optimal design using rigorous optimization methods becomes computationally prohibitive when studying complex processes. One usually abandons the requirement to solve both optimization problems rigorously. It is common industry practice to perform the outer optimization, which selects the best flowsheet and sizes the equipment, manually based on engineering insight while using commercial process simulation and/or optimization tools, e.g., Aspen Plus, to solve the inner optimization problem.

More rigorous methods have been developed since the 1970's that focus on justifying and, if

possible, reducing process overdesign. Freeman and Gaddy suggested a new measure, which they termed dependability, to quantify how well a proposed flowsheet meets its specifications as measured by a certain process performance criterion under uncertainty (Freeman and Gaddy, 1975). Finding this measure requires the solution of an integral over the probability density distribution of the considered performance measure. The authors state that integrating will in most cases require numerical integration and suggest that Monte Carlo methods are suited. They give an example that consists of a simple process flowsheet and showcase the method. Though they introduced the concept of dependability first, their paper lacked a clear definition. Later, Pistikopoulos and Mazzuchi (1990) introduced the notion of stochastic flexibility, which is defined strikingly similar to dependability. Both concepts measure the probability that a process is feasible given a joint probability density distribution of the uncertain parameters. Similarly, Straub and Grossmann (1990) also used the notion of stochastic flexibility citing earlier work by Pistikopoulos and Mazzuchi. Both, Grossman and Pistikopoulos refer to earlier work by Kubic and Stein (1988), who introduced yet another term, design reliability. They actually cite the paper by Freeman and Gaddy, but also do not mention the notion of dependability.

Grossmann and Sargent suggested an optimization formulation that requires the process to remain feasible for all possible realizations of the uncertain parameters (Grossmann and Sargent, 1978). As outlined before, design decisions are held constant across all scenarios while operating decisions are adjusted to each scenario. The authors propose to approximate the expected value of the objective that is defined as an integral over the probability density distribution by a finite weighted sum of representative scenarios. They claim that the density function is typically not known very precisely so that a small number of samples may give a sufficient answer. These sample points should be selected to cover the likely range.

Grossmann *et al.* extended these ideas. The authors use the notion of flexibility, the property of a process to ensure feasible regions of operation for any realization of the uncertain parameters (Grossmann *et al.*, 1983). In the publication, optimization formulations to design a process for a fixed degree of flexibility are presented and generalized to give the optimal degree of flexibility. In the first case, the designer specifies the range of parameters for which the process is to remain feasible. In the second case, an optimal trade-off

between the cost of the plant and its flexibility is found. Here, a two-staged problem is formulated, which includes a feasibility constraint that is responsible for meeting all specifications for any realization of the uncertain parameters. The expected value of the cost subject to the feasibility constraint is minimized. The feasibility constraint is written as a max-min-max constraint that is difficult to compute. To remedy this problem, an index of flexibility is introduced that stems from the length of the side of the largest hypercube that can be inscribed into the feasible region of the process about the nominal point of design.

The methods presented by Grossmann and coworkers lead to nonlinear programming problems that can be of considerable size as each subproblem stemming from a sampled realization of the uncertain parameters contains a complete model of the flowsheet. Thus, typically decomposition methods such as Bender's decomposition are employed to make the problem computationally feasible. Recent numerical methods for solving these robust optimization problems are reviewed by, e.g., Diehl et al. (2008). In the paper, the authors present a generalized semi-infinite programming formulation to study the worst-case behavior of a system in the context of control, which can be interpreted to address similar questions as raised by uncertainty during the design phase. A local reduction approach is employed so that the inner problem can be replaced by its optimality conditions, which turns the generalized semi-infinite program into a locally reduced finite NLP. Note that these discussed formulations can lead to considerable overdesign, as they require by definition that the process remains feasible for any possible realization of the uncertain parameters. Thus, it yields process designs that include sufficient design reserves to cope with parameter variations within the specified bounds.

Common to the previously discussed methods is the need to evaluate expected values. As in most cases of practical importance no analytical solution can be found, different methods are used to solve the integral over the probability density distribution. One wide known method with a wide range of applications is Monte Carlo methods. Here, the integral is approximated by sampling points randomly from the range space according to their probability. The expected value follows then simply by adding the function values of sampled points (Spro, 1967).

Though in principle such methods can be applied to problems of arbitrary dimension, one also needs to consider the computational expenditure necessary to

find a converged solution. The original Monte Carlo methods assume that the points are sampled randomly according to their probability. Developments to improve the rate of convergence have led to so-called quasi-Monte Carlo methods that use deterministically determined series of samples. The key idea is to use additional information when selecting sample points. For example, Latin hypercube sampling techniques divide the sampled space into hypercubes of equal probability and require that from each hypercube only one point is sampled, thus forcing the algorithm to sample a more representative portion of the space in fixed time (McKay *et al.*, 1979).

Proposed Approach: Relaxation-Based Dependability Analysis

Building upon Freeman and Gaddy's approach, a Monte Carlo scheme is employed to study the impact of uncertain parameters on the design of a process plant. Utilizing i-TCM, processing equipment and its limitations are modeled in Aspen Plus using shortcut sizing methods. The simulation runs are controlled via an Excel interface. One of the greatest obstacles of automated process simulations is the difficulty to differentiate between physical infeasibility and numerical difficulties within the solver. A heuristic priority scheme is proposed that supports the user in significantly increasing the number of converged runs by augmenting the optimization with a tailored relaxation scheme. With this procedure, the process engineer gains insight into the limitations of the process without the need of interpreting the results of failed simulation runs—in itself a difficult and sometimes even hopeless task in times of equation-oriented simulation.

As discussed in the beginning, uncertainty can increase the difficulty of process design considerably. In this case, dependability analysis provides a very helpful tool to support process engineers in the design process. In combination with the novel approach presented in this paper and showcased in a case study, dependability analysis can enhance understanding of complex chemical processes significantly. Unlike several approaches proposed in the literature, relaxation-based dependability analysis is a tool to support daily process engineering work without the need for intensive user intervention and numerical tuning effort. Another key characteristic of the proposed method is that existing process models can be used for the analysis in straight forward manner via an easy-to-use interface.

DEPENDABILITY ANALYSIS

Keeping track of uncertain parameters and the implications of variations of them becomes increasingly difficult as their number increases. While one may be able to compare different designs qualitatively when uncertainty is restricted to one or two parameters and a narrow bandwidth, quantitative comparisons quickly become intractable. For one, the need to keep track of probabilities and different values for the objective make it a cumbersome problem to carry out manually. Thus, engineers typically design for a nominal point of operation and consider some perturbations manually. While this procedure may build intuition and thus may enable predictions about the performance when other parameter variations occur, it is certainly a challenging task for complex process flowsheets.

In order to move past just qualitative comparisons among competing designs, a quantitative measure that can be calculated automatically provides a great support tool to the process engineer. It allows ranking designs and then focusing on different behaviors over the range of studied parameter variations, thus freeing the process engineer from a manual task while supporting a much more complex task.

Dependability as a Quantitative Measure

Freeman and Gaddy (1975) propose such a measure that they termed dependability, defined as the probability that the process meets its specification,

$$D(x_d, s) = \int_{x_0} P(x_0) \delta(x_0, x_d, s) dx_0, \quad (1)$$

where x_d notes the design variables, s the process specifications, x_0 the uncertain parameters, $P(x_0)$ the probability density function for x_0 and $\delta(x_0, x_d, s)$ states if the design meets the specifications under the assumed uncertain parameters ($\delta = 1$), or if does not ($\delta = 0$). $P(x_0)$ will be known in advance, whereas $\delta(x_0, x_d, s)$ is typically an unknown function. The calculation of D is termed solving the outer problem in analogy to bilevel programs.

To determine $\delta(x_0, x_d, s)$, the so-called inner problem is solved. While the outer problem focuses on the uncertain parameter, the inner problem adjusts the operating conditions to ensure that a feasible process is found, if possible. Recall that the design is fixed so that the degrees of freedom of this problem

correspond to the adjustments possible once a design has been chosen and built such as reflux ratios in distillation columns or reactor temperatures. It is important to point out that there is no trivial method to determine $\delta(x_0, x_d, s)$ which is not accessible a priori. In practice, to solve the inner problem and determine feasibility of the design for the specific sample of the uncertain parameters, a commercial process simulator with its rigorous process models is used.

Before discussing more details of the procedure, it is worthwhile to note certain aspects of dependability. A different interpretation of it can be given by the percentage of time when the plant is operating in specification (Freeman and Gaddy, 1975). In many cases, $D < 1$ does not necessitate that, on average, the process will not be able to meet the specification, for which it has been designed. On the contrary, in certain cases, a plant with a smaller additional design margin may still be able to make up for times of reduced production when the varying parameters are in its advantage at a later time. Also, $D < 1$ does not entail that the plant is down at some instances throughout the year, though product quality or quantity may be reduced. In contrast, methods requiring the process to stay within its specifications at all times (e.g., Grossmann *et al.*, 1983) lead to greater equipment overdesign to fulfill these conditions. This is more important when the uncertainty stems from parameters that continue to vary when the process is operational, such as feedstock purity conditions. In contrast to uncertainties in, e.g., kinetic constants whose value is unknown, but actually does not vary, it is not necessary to design processes that are always capable of delivering nominal capacity.

Determining Dependability Using a Quasi Monte Carlo Method

In order to obtain an estimate for the dependability, the integral in Eq. 1 has to be solved. When examining the integrand, one notes that δ is a complicated function and it is in most cases not known upfront. On the other hand, the probability density function for the uncertain parameters is specified when the problem is set up using the available information such as confidence intervals on experimental data or known variations in raw materials. Though the integral is too complex to be solved analytically, it can be solved by numerical means. Monte Carlo (MC) methods are widely used to solve such—typically multi-dimensional—integrals in many different contexts and also have been applied to supplement decision-making (Sprow, 1967).

Instead of using numerical quadrature formulas like Gaussian quadrature that construct a converging series, Monte Carlo methods approximate integrals by different means. They randomly sample the function value at different locations in parameter space. In particular, the accuracy of the result follows from the law of large numbers and is independent of problem dimensionality (Caflisch, 1998). However, its rate of convergence can be slow, especially when highly accurate results are required.

In contrast to random Monte Carlo methods, which use randomly selected samples in parameter space according to the probability distribution function, quasi-Monte Carlo methods uses a deterministic number sequence that is chosen explicitly to increase the rate of convergence (Morokoff and Caflisch, 1994). In contrast to the random selection, a deterministic sequence of points can ensure that the most representative sample of the parameter space is obtained. It has been pointed out in the literature that these sampling methods only need to visit a small fraction of the space to obtain a representative answer (Bernardo *et al.*, 1999). As Monte Carlo methods are often applied to problems where each function evaluation at a sample point is computationally expensive, as is the case in the problem studied here, simulations only on the order of hundreds to thousands can be performed in reasonable time. On the other hand, the sampled space is multidimensional for most problems in industrial applications so that there is great need to ensure that the available computational effort is used to explore as much of the space as possible. Thus, quasi-Monte Carlo methods provide a more efficient means for numerical integration here.

Latin hypercube sampling, an example of a quasi-Monte Carlo method, divides parameter space in each dimension in equally probable sections and samples only once from each section. It has been shown that quasi-Monte Carlo methods are capable of outperforming random Monte Carlo methods in terms of obtaining more accurate results with the same number of evaluations (Morokoff and Caflisch, 1994). However, the authors also point out that quasi-Monte Carlo methods become less advantageous as the dimensionality of the problem increases due to increases in discrepancy. This can be explained as follows: deterministic sampling sequences are constructed using geometrical arguments in order to guarantee a favorable spacing. However, as the dimensionality of the problem increases, sequences suggested in the literature show repetitious, correlated behavior in projections of some dimensions. This can be overcome by increasing the number of samples

taken. For the problem considered here, the number of uncertain parameters, which are studied simultaneously, is small so that the discussed problem does not impair the method.

To overcome this shortage, a Hammersley sequence was used in this contribution to obtain the deterministic number sequence for the quasi-Monte Carlo method. Diwekar and Kalagnanam (1997) proposed this low-discrepancy sequence that can easily be expanded to higher dimensions. They provide a description of an algorithm that constructs the sequence, which has been implemented in this work. Furthermore, the authors also presented evidence for favorable convergence properties of their proposed sequencing method.

Other authors have applied different numerical integration methods to the problem of design under uncertainty (e.g., Straub and Grossmann, 1990 and Bernardo *et al.*, 1999). As Bernardo *et al.* remark, the number of required samples grows exponentially with the dimensionality of the problem when quadrature methods, e.g., Gaussian quadrature, are used. They also note that a sampling technique, i.e., a quasi-Monte Carlo method, is more adequate for problems with a larger number of uncertain parameters. Thus, for a small number of parameters as considered in the case study, quadrature techniques may be more efficient, but more complex scenarios will require a quasi-Monte Carlo technique and thus it was decided to utilize the latter throughout.

Implementation of Dependability Analysis

Dependability is accessible to determination with the tools available to process engineers, a key feature of this approach to ease industrial application. In contrast to rigorous methods discussed above that depend on specialized optimization software, the stochastic approach can be carried out using process simulation tools in combination with standard office software. The method presented above can be implemented in commercial process optimization tools, in this case Aspen Plus, in connection with an external controller for the Monte Carlo simulations, here Microsoft Excel. At each sampled point, a simulation of the process flowsheet with modified parameters is conducted. Simulations and their data in Aspen Plus are accessible from Excel using an ActiveX automation server interface. Additionally, routines to determine the Hammersley sequence, to transform this uniform sequence to the specified probability distribution of the uncertain parameter, and to control runs in Aspen Plus are implemented in Visual Basic for Applications (VBA) in Excel. Furthermore, post-

processing and visualizing of accumulated data is also included in the Excel workbook. In detail, for each individual Aspen Plus run, information on the values of the sampled parameters, selected computed variables, and status of the run are reported to the user as a reference. These data are aggregated and reported for each variable as mean and standard deviation. Furthermore, dependability and fraction of converged simulations are provided.

RELAXATION-BASED DEPENDABILITY ANALYSIS OF FLOWSHEET SIMULATIONS

In early attempts of this work, it was quickly realized where major obstacles lie in this approach. The need to repeatedly perform process flowsheet simulations requires great autonomy of the software. But every user of flowsheeting software is aware of the difficulty inherent to converging these simulations. Flowsheets are therefore carefully constructed by increasing complexity slowly. To address this difficulty when varying parameters repeatedly, the need for additional heuristics to guide the software to a converged solution had been identified.

For further discussion, a distinction will be made between reasons that can cause a flowsheet simulation or optimization to fail. Firstly, the flowsheet may be physically infeasible, for example energy balances cannot be satisfied with the given specifications, separations are physically infeasible at the attempted conditions or the intended reaction pathways fail to deliver the necessary conversion. In this situation, the designed process is clearly unable to perform its task. Secondly, a flowsheet simulation may fail due to numerical reasons such as slow convergence, numerical instability or difficulty to identify a feasible solution. However, process simulators such as Aspen Plus do not provide results that differentiate between these reasons for which a simulation did not converge. This task is left up to the user's experience to decide and select how to move forward in resolving this situation.

Relaxations as an Aid in Differentiating between Numerical and Physical Infeasibility

A novel idea is introduced here to overcome the lack of robustness when solving the optimization problem for the process flowsheet automatically: A heuristic based on relaxations is used. Since only converged results—regardless of their success—provide insight into physical constraints of the problem, it is essential to reduce infeasibility due to

numerical reasons without requiring user interaction. Otherwise, when nonconverged runs are excluded, the statistic can be severely biased and report too optimistic values as simulations are more likely to converge for perturbation towards physically “easier” parameter values.

As outlined, non-convergence of a flowsheet can be either due to numerical issues when solving the flowsheet or physical infeasibility of the designed process subject to the assumed conditions. While the latter provides the engineer with useful information about the design, the former is unfavorable as numerical problems conceal the question of physical feasibility. In general, it is a very time-consuming, manual task to converge a process flowsheet. In this case, such a path is infeasible as it will be necessary to perform on the order of hundreds or thousands simulations for one specific process design. The novel idea that can greatly boost convergence—key to providing reliable information about the process—is a heuristic used to prioritize competing restrictions using insights from daily operations. When operating a plant, the most important objective is to keep a plant running that is tied into a large production network as it serves both as a consumer of intermediates from other processes and produces feedstock for even other plants. Thus, if fluctuations in uncertain parameters cannot be overcome by the implemented control loops, it is more important to keep production quality on target than produce the required quantity. Although cost-effectiveness is important, shutting down a process can have costly repercussions on a larger scale. This creates several staggered goals for the process to meet: If the uncertainties cause the process to fail the most stringent target, there are still looser restrictions to meet. Thus, by providing these relaxations of the original specifications, one can increase convergence, albeit to less stringent targets.

Lastly, it should be noted that constraint relaxation in the context of this paper differs from the idea suggested by Bernardo *et al.* (1999). They introduced a strategy to simplify the integration to determine the expected value of the objective function that they also call constraint relaxation. For partially feasible solutions, they suggest to penalize the solution by adding a penalty term to the objective. Here, constraint relaxation is a method to increase the number of convergent process simulations.

Algorithmic Details

According to this ranking the most important priority is to produce substances of desired purity,

the second priority is to achieve the desired quantity of product, and lastly, the operating costs are to be minimized. Hence, the optimization problem within Aspen Plus is set up to maximize product subject to purity constraints and includes an upper limit on the produced amounts.

The starting point of all simulations is a converged simulation supplied by the user that has been set up for the nominal point of design. The uncertain parameters are adjusted each time according to the sampled values. First, the simulation is started from the converged flowsheet. If this flowsheet converges, the results are directly used when calculating the expected value. Otherwise, either physical infeasibility or numerical problems lead to convergence problems. In this case, bounds of the optimization problem can be selectively relaxed to user-specified new values when the original problem does not converge. Here, the intention is to increase the range of feasibility. If the problem converges in a second attempt with the relaxed bounds, this solution is used as the starting point for a homotopy method. After each successful convergence of the optimization problem, the relaxed bounds will be consecutively retightened until the original constraints are reached. If the flowsheet can be converged with the original constraints in place, the results are used in the calculation of the integral. Otherwise, the results do not participate in the process evaluation, but they are noted in the detailed output to the user. Apparently, one is attempting to operate the process just beyond its physical limit in this case, which is useful information when studying the feasibility limits of a proposed process design.

Dependability can be viewed as a condensed measure of the ability of the process design to cope with varying conditions. In combination with heuristics to improve convergence of the flowsheet, it can be determined automatically. Certainly, more converged runs increase the statistical foundation for the measure of dependability. Furthermore, more detailed information about feasibility limits also gives the process engineer a better understanding of physical limitations of the process.

CASE STUDY: ETHYLENE OXIDE SYNTHESIS

To illustrate the above-described method, a process to synthesize ethylene oxide is studied that was set up based on available information in the literature (Onken and Behr, 1996; Rebsdats and

Mayer, 2005). The process is modeled and solved in Aspen Plus using the equation-oriented (EO) simulation mode.

Setting up simulation for Relaxation-based Dependability Analysis

First, an optimal process design—in terms of minimized overall cost—is found using short-cut equipment sizing and costing methods to provide the baseline for the analysis of effects of uncertain parameters (Biegler *et al.*, 1997; Wiesel and Polt, 2006). The dimensions resulting from this optimization are subsequently adjusted to account for empirical overdesign factors. Then, physical limitations of used equipment are included in the model using simplified physical measures, such as F-factors for distillation columns, similar to those used in the sizing procedures. As mentioned earlier, the optimization problem to solve for the dependability analysis is to maximize product quantity subject to unbounded feed streams, product purity restrictions and the above designed plant with its modeled physical limitations, e.g., an upper limit on the product quantity.

In this case study, uncertainty in conversion rates used to model the reaction kinetics as well as fluctuating raw material purities are regarded. It is assumed that all parameters follow a triangular distribution with specified lower and upper limits as well as a mode. The number of simulations, that are performed sampling the uncertain parameters according to the above-described quasi-Monte Carlo technique, is increased until the results converge to 10^{-1} . It is found that 1000 simulations are required, which is in agreement with results reported by Bernardo *et al.* (1999). Overall, the simulations require computation time on the order of several hours.

Results of the Dependability Analysis

In the studied case, more than 95% of all simulations converged; 18% required lower product quantity to converge whereas more than 77% complied with the process specifications. Therefore, the dependability of the design is 0.77. Figure 1 shows that these different result areas are rather cleanly separated from each other except for a few numerical artifacts.

In the reported case, the plant is capable of producing on average close to 98% of nominal product quantity when neglecting samples for which no information is available due to nonconverged simulation runs.

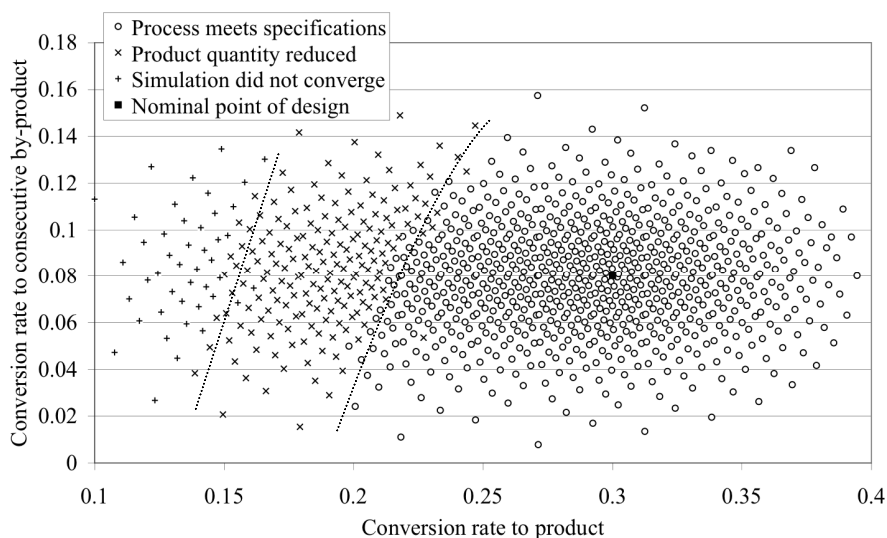


Figure 1: Plot of simulation results show the influence of two uncertain parameters on process dependability. Bounds of the dependable range are approximated and included for illustration.

Discussion

As Figure 1 shows, the process is well behaved in proximity to its nominal point of design. However, as conversion rates deviate in unfavorable directions, the limits of the process are tested and exceeded as denoted by decrease of product. When the studied parameters vary even more, numerical issues inhibit further studies, as convergence is lost.

Although the dependability of the designed process appears to be fairly low, the design is capable of nearly achieving nominal product quantity. The discrepancy between low dependability and high actual product quantity can be explained by the only gradual decline of product output as the process is operated at points beyond its name plate production bound. Here, the importance of as many as possible converged simulations becomes obvious since conclusions can only be drawn from results of converged results. Therefore, if one were to not account for and not to include this region between fulfilling specifications and no convergence in numerical results for averaged process state variables, the actual performance of the plant would be greatly underestimated, resulting in the seemingly need for greater equipment overdesign.

CONCLUSION AND FUTURE WORK

With relaxation-based dependability analysis, an approach has been proposed in this work to support

process design when some parameters are not known exactly. One important aspect of the methodology is the interconnection of process simulator and simple user interface to encapsulate and automate repeated Aspen Plus runs. Equally important are the proposed heuristic means of relaxation to differentiate between numerical problems and physical limitations. These lead to an increase in converged runs which allow this approach to generate non-trivial insight into complex chemical processes and to provide a quantitative measure of sensitivity of the designed process to uncertain parameters. Hence, the proposed method can provide help to process engineers when designing processes to meet uncertain process conditions on one the hand while, on the other hand, limiting unnecessary process overdesign.

The method of relaxation-based dependability analysis can be applied to study effects of manifold uncertain influences to a process. Though in the case study presented uncertainties have been restricted to process parameters, the method can be easily extended to incorporate uncertain cost coefficients, e.g., to study the impact of fluctuating raw material costs. Likewise, the proposed priority ranking need not be applicable for all processes; it can be adjusted to meet different priorities without further implications to the methodology.

Monte Carlo methods rely on performing many simulations. Thus, there will be many simulations in close proximity to the base case as these are very likely and, in most cases, these simulations will converge and report a positive outcome. It would be

very helpful to be able to identify this range of converged simulations with positive outcomes based on already performed simulations. If such identification can be made with sufficient confidence, simulations within this range may be skipped and recorded as successful, thus freeing computational efforts that can be focused on the more difficult regions of parameter space.

Another important issue for practical applications is to design good graphical representations of results obtained with multiple varying parameters. Once the number of parameters is increased past three, only projections onto a lower dimensional space can be viewed, thus losing information. Hence, complex interactions between multiple parameters are difficult to identify. Here, statistical methods can be helpful and should be investigated.

NOMENCLATURE

$D(x_d, s)$	dependability of design x_d under specifications s
$\delta(x_0, x_d, s)$	signifies if design x_d meets specifications s under assumed uncertain parameters x_0 ($\delta = 1$) or not ($\delta = 0$).
$P(x_0)$	probability density function for x_0
s	process specifications
x_0	uncertain parameters
x_d	design variables

REFERENCES

- Bernardo, F. P., Pistikopoulos, E. N. and Saraiva, P. M., Integration and computational issues in stochastic design and planning optimization problems. *Ind. Eng. Chem. Res.*, 38, No. 8, 3056 (1999).
- Biegler, L. T., Grossmann, I. E. and Westerberg, A. W., *Systematic methods of chemical process design*. Prentice Hall, Englewood Cliffs (1997).
- Cafflich, R. E., Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7, 1 (1998).
- Diehl, M., Gerhard, J., Marquardt, W. and Mönnigmann, M., Numerical solution approaches for robust nonlinear optimal control problems. *Comp. Chem. Eng.*, 32, No. 6, 1279 (2008).
- Diwekar, U. M. and Kalagnanam, J. R., Efficient sampling technique for optimization under uncertainty. *AIChE J.*, 43, No. 2, 440 (1997).
- Freeman, R. A. and Gaddy J. L., Quantitative overdesign of chemical processes. *AIChE J.*, 21, No. 3, 436 (1975).
- Grossmann, I. E. and Sargent, R. W. H., Optimum design of chemical plants with uncertain parameters. *AIChE J.*, 24, No. 6, 1021 (1978).
- Grossmann, I. E., Haleman, K. P. and Swaney, R. E., Optimization strategies for flexible process design. *Comput. Chem. Eng.*, 7, No. 4, 439 (1983).
- Kubic, W. I. and Stein, F. P., A theory of design reliability using probability and fuzzy sets. *AIChE J.*, 34, No. 4, 583 (1988).
- Malik, R. K. and Hughes, R. R., Optimal design of flexible chemical processes. *Comput. Chem. Eng.*, 3, No. 1-4, 473 (1979).
- McKay, M. D., Beckman, R. J. and Conover, W. J., A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21, No. 2, 239 (1979).
- Morokoff, W. J. and Cafflich, R. E., Quasi-random sequences and their discrepancies. *SIAM J. Sci. Comput.*, 15, No. 6, 1251 (1994).
- Pistikopoulos, E. N. and Mazzuchi, T.A., A novel flexibility analysis approach for processes with stochastic parameters. *Comp. Chem. Eng.*, 14, No. 9, 991 (1990).
- Onken, U. and Behr, A., *Chemische Prozeßkunde - Lehrbuch der technischen Chemie (Vol. 3)*. Georg Thieme Verlag, Stuttgart (1996).
- Rebsdatt, S. and Mayer, D., Ethylene Oxide. In: *Ullmann's Encyclopedia of Industrial Chemistry (7th edition)*, Ullmann, F. et al. Wiley-VCH Verlag, Weinheim (2005).
- Sprow, F. B., Evaluation of research expenditures using triangular distribution functions and monte carlo methods. *Ind. Eng. Chem.*, 59, No. 7, 35 (1967).
- Straub, D. A. and Grossmann, I. E., Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties. *Comp. Chem. Eng.*, 14, No. 9, 967 (1990).
- Wiesel, A. and Polt, A., Conceptual steady state process design in times of value based management. *Proceedings of PSE 9 and ESCAPE 16*, 799 (2006).