Stochastic Framework for Modular Design Generation

Marianthi G. Ierapetritou* and Vishal Goyal
Rutgers University
Piscataway, NJ 08854

Abstract

One of the important considerations in plant design is to be able to accommodate changing product demands of future markets. All of the existing approaches for design optimization aim at creating customized designs in which production capacity have been specified to meet certain market demand. This paper presents a module-based approach to integrate probabilistic demand data-analysis and robust design optimization stages, that have been traditionally performed separately, to generate a set of designs that span the demand space in a cost-optimal fashion. A case study of design of a reactor-separator system is presented to illustrate the applicability of the proposed approach.

Keywords

Stochastic Optimization, Modular-Design, Data-Analysis

1 Introduction

Today's marketplace reality is characterized by rapid, uncertain, and continuous changes. Since design specifications of chemical plants, such as product demands, ambient conditions or model parameters normally vary during process operation, an optimal design must not only be economically optimal but also capable to operate in steady-state for a range of variable conditions that may be encountered. Competing markets enforces a high degree of market adaptation and orientation necessary.

With respect to the way the uncertainty is handled, three different approaches are known, such as, (a) the scenario-based approach (Grossmann and Sargent, 1978); (b) the stochastic approach (Pistikopoulos and Ierapetritou, 1995) and (c) the parametric approach (Pertsinidis et al., 1998). The first two approaches are based on characterization of the uncertain parameter space by considering either discrete scenarios or stochastic distributions, assuming that some information regarding the uncertainty is provided either in the form of most expected nominal point or specific range of values or in the form of a probability distribution function. In the parametric framework, no assumptions are made on the uncertainty model and the design optimization problem is solved parametrically over the uncertain demand space resulting in a map over the uncertainty-design space.

To handle the trade-off associated with the expected cost and its variability, Mulvey et al., (1995) proposed the concept of robustness. A decision is termed robust if the actual cost of the realized scenario remains "close" to the optimal expected cost of all the scenarios. Since then, a number of papers have been published involving robustness (among others, Suh and Lee, 2001), all varying in the definition of the robustness function and the process being optimized.

The motivation for this paper comes from realizing that significant economic savings can be achieved if standardized designs can be developed by taking into account the customer demand space. The relevance of the work origins from the fact that developing modular designs would be substantial cheaper for a manufacturer as developing a portfolio of designs would save considerable time and money and for a customer, various design alternatives would be available, based on cost and flexibility. Thus, the objectives of this paper is to introduce an unified data analysis-design optimization framework for determining robust optimal designs, weighted over demand probabilities. The main idea is to cluster the demand data, solve a stochastic robust design optimization problem over the clusters and iteratively improve the designs to generate the minimum set of robust designs that will span the demand space. The paper is organized as follows. Following this introduction, in section 2, the detailed proposed framework is presented. In section 3 a case-study of a reactor design is presented to illustrated the relevance of the approach. Finally, section 4 summarizes the work and presents future work directions.

*Author to whom correspondence should be addressed. Email: marianth@sol.rutgers.edu. Tel: 732-445-2971. Fax: 732-445-2421

2 Module-Based Design Optimization Framework

The main target of the proposed algorithm is to expand the boundary of the design optimization problem by integrating the data-analysis stage, thus allowing more flexibility at the decision making process. The basic aim is, given a demand-map where each demand point has randomness, to determine an optimal portfolio of designs that jointly cover the entire demand space based on customer requirements. The proposed algorithm, consists of the following stages:

Step 1: The demand data are initially clustered in a minimal number of clusters using fuzzy-clustering.

Step 2: A stochastic robust design optimization problem is solved over each cluster to generate a design for the cluster-points.

Step 3: The "value" of a design is evaluated based on expected cost and feasibility criteria.

Step 4: The number of clusters are then increased by one and steps 1-3 are repeated for the new clusters.

Step 5: The new set of designs obtained at step 4 is compared with the previous set of designs on the basis of cost and flexibility criteria.

Step 6: If any "better" designs are obtained at step 4 then perform steps 4-5 again and compare the designs with the previous set of designs, until no new (smaller/cheaper) designs are obtained.

The following classifications should be made about the steps of the algorithm:

Step 1: It is not possible to know a-priori the minimal number of designs that can cover the entire demand space, thus an iterative procedure has to be utilized that starts from a small number of designs. Due to the probabilistic nature of the demand, a fuzzy clustering approach provides a systematic methodology to incorporate variance within a clustering framework. The clustering algorithm FANNY (Kaufman and Rousseeuw, 1990) is utilized to cluster the demand.

Step 2: The stochastic robust design optimization problem can be mathematically modeled in a discretized formulation as follows:

$$\min \sum_{s,k\in\delta} p_{s,k}\xi_{s,k} + \lambda \sum_{s,k\in\delta} p_{s,k}\Big(\xi_{s,k} - \sum_{s,k'\in\delta} p_{s,k'}\xi_{s,k'}\Big)^2 \quad (1)$$
$$+\omega \sum_{s,k\in\delta} p_{s,k}\Big(\sum_{\theta\in\Theta}(z^k_{\theta,m})^2\Big)$$
$$h(d, x_{s,k}, z_{s,k}, \theta_{s,k}) = 0$$
$$g(d, x_{s,k}, z_{s,k}, \theta_{s,k}) \leq 0$$
$$prod^k_{\theta_m} + z^k_{\theta_m} \geq dem^k_{\theta_m} \quad \forall\theta\in\Theta, k\in N$$
$$prod^{s,k}_\theta + Lz^{s,k}_\theta \geq dem^k_{\theta_m} + \sigma_i\Phi^{-1}(\alpha)\ \forall\theta\in\Theta, s,k\in\delta$$

where $\xi_{s,k} = C_{s,k}(d,x,z,\theta)$ represents the total operating and capital cost for scenario s for demand point k; d is the set of design variables; x, z are state and control variables and θ is the uncertain variables following a normal distribution $\mathrm{J}(\theta)$; h and g are the equality and inequality constraints respectively; $z^{s,k}_\theta$ represents the unmet demand for each point s_k and z^k_θ represents the unmet demand of the expected-mean θ^k_m for each point k; dem^s_θ and $prod^s_\theta$ are the required demand and production amount for demand θ in scenario s; $p_s = \mathrm{J}(\theta_s)$ is the probability of the demand and the above deterministic formulation is generated after the discretization of the uncertain-variables k into s samples, s_k.

The first term in the objective function is the expected cost of the design over all the demands in the cluster. The second term is the variance of the expected cost, weighted by parameter λ which penalizes a high variance in the operating cost of the optimal design due to a possible large demand span. The presence of outliers, may still drive the optimal solution toward an over-design, hence the third term is introduced to allow for a degree of infeasibility in the demand satisfaction, weighted by parameter ω.

The third constraint is the model robustness constraint that allows some infeasibility in the demand satisfaction using the means as the representative for each demand point distribution. The last constraint is the deterministic formulation of the chance constraint $P((prod_\theta \geq dem\theta) \geq \alpha))$ where α is the probability of satisfaction and Φ^{-1} is the inverse cumulative distribution function (Prekopa, 1995). Furthermore, an additional term, $Lz^{s,k}_\theta$ is added where L is a large positive constant to negate the effect of any demand-point that has been rejected by the robustness constraint as an outlier.

Steps 3: The "value" of a design is quantified based on two criteria. The first criterion is the expected total cost. If a demand point is feasible for two alternative designs, then the cheaper design is used to satisfy the demand. The second criteria is the number of feasible demand points for a design. For convex systems, to evaluate the operability limits of a design the simplicial-approximation approach proposed by Goyal and Ierapetritou, 2002 can be utilized.

Steps 4-6: In step 4, the number of clusters is increased by one and an alternative set of designs are obtained for each of the new clusters. The new set of designs are then compared with the previous sets on the basis of design "value". Since at each iteration, the demand-space for each design optimization is reduced due to increase in number of clusters, a better (smaller/cheaper) design is obtained for some of the demand-points. This process is repeated until, no new designs are obtained at the clustering stage.

Thus at the end of the algorithm, a set of design alternatives are developed that span the entire demand space. The various alternatives would result in an increased decision making flexibility due to better utilization of the available demand data. Specifically, great savings can be achieved since modular-based designs are obtained to cover the entire range of demand and better information is provided regarding different design alternatives involving design-costs, flexibility and robustness.

It should be noted that the existence of outlier points in the demand-space, may disturb an effective clustering and hence affect the outcome of the algorithm. However, the outliers can be identified if they persistently appear as unmet demands in the robust design-optimization problem. A case-study of a reactor-separator system is presented in the next section to illustrate the applicability of the proposed approach.

3 Case Study: Reactor-Separator System

The case study presented in this section is the design of a CSTR in series with an ideal separator. The aim is to minimize the cost to convert raw material A into two finished products B and E. Details and explanations of the model equations and kinetics are described in Rooney and Biegler (1999). The design variable is the CSTR volume V and the objective function consists of minimizing the expected capital cost. The demand plot for the products B and E were randomly generated assuming a bivariate normal distribution with a standard deviation $(\sigma_i^2) = 36$ and a correlation coefficient $(\rho_i) = 0.5$ and the expected means of the demands are as shown in Figure 1.

The objective of the problem is to determine the optimal set of designs (V, F_{ao}) to cover the demand space, where F_{ao} is the input flow-rate of A. The optimal set is defined as the minimum number of designs with minimum total cost that can be used to cover the entire demand space. The problem is modeled following the stochastic design optimization framework (problem 1). The variance of the cost of design is determined using the variability of the operating cost consisting of the cost for recycle. The robustness-parameters are fixed to, $\lambda = 0.1$ and $\omega = 1000$ for the case-study solved in this paper and each demand distribution is discretized into a set of 10 scenarios. The proposed approach is thus applied for the reactor system and the results are as summarized below.

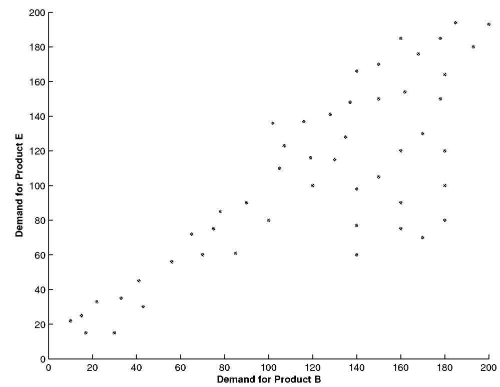

Figure 1: Demand Space

3.1 Computational Results

The application of the proposed approach is initialized by assuming three clusters. Steps 1-2 of the proposed approach are applied where the stochastic robust design optimization problem is solved considering the corresponding discretization for each of the three clusters. A design with V = $32.4m^3$, F_{ao} = 108.8 mol/hr and an expected cost = \$3344.5 is obtained for the cluster illustrated by "+" in Figure 2; a design with V = 65.3, F_{ao} = 259.9 and an expected cost of \$6785.9 is obtained for the cluster shown by $\triangle$ in Figure 2; whereas a design with V = 119.4, F_{ao} = 401.3 and an expected cost of \$12343.6 is obtained for the third cluster shown by "$\circ$" in Figure 2. Due to the non-convex nature of the constraints and the low dimensionality of uncertainty, feasible region approximation is performed using grid-search simulations and the results are as shown in Figure 2. The first design is feasible over 8 out of 14 demand points, the second over 11 out of 28 points and the third design is feasible over 10 out of 18 demand points in the cluster.

Steps 4-5 of the proposed approach are then applied by increasing the number of clusters by one. This results in one additional robust-designs with: $(V, F_{ao}$,capital cost,feasibility) = (100.7, 420, \$10487.6, 7 over 12 demands). The remaining three clusters yield no new designs. The new design configurations is "better" than the previous obtained designs with respect to cost and hence is accepted and another iteration is carried out. Thus, at the end of the second iteration, a total of four (three from previous iteration) robust designs have been determined with varied degree of flexibility that span the demand space.

The number of clusters is then increased to five, resulting in the following three new design configuration: (V,F_{ao},capital cost, flexibility) = (26, 80, \$2620, 5 over 8 demands); (V,F_{ao},capital cost, flexibility) = (58, 220, \$6020, 6 over 8 demands) and (V,F_{ao},capital cost, flexibility) = (96, 316.9, \$9916.9, 8 over 11 demands). Thus at the end of the 3^{rd} iteration a total of seven robust designs have been developed that cover the demand space.

The number of clusters is then increased to six, resulting in one new robust design with (V,F_{ao},capital cost, flexibility) = (136.1, 459.7, \$14067.8, 6 over 8 demands). Thus, another iteration of the procedure is carried out by increasing the number of clusters to seven but no new design configurations were obtained at this stage. Hence the iterative procedure is terminated at the 5^{th} iteration.

To summarize the results, at the end of the optimization process, a set of **eight modular** robust designs have been developed that span the demand-space. The set of robust designs obtained after the first iteration are also summarized in Figure 3, where grid-search simulations have also been performed to show the feasibility of the designs. For the points that are covered by more than one designs there are some interesting trade-offs that have to be considered in the decision making process since the more expensive designs have higher flexibility and thus higher profit can be anticipated if higher demand is to be realized in the future.

4 Summary and Future Directions

A novel framework is presented in this paper for the integration of data-analysis and robust design optimization stages that have been traditionally performed separately. The basic idea is to apply a fuzzy

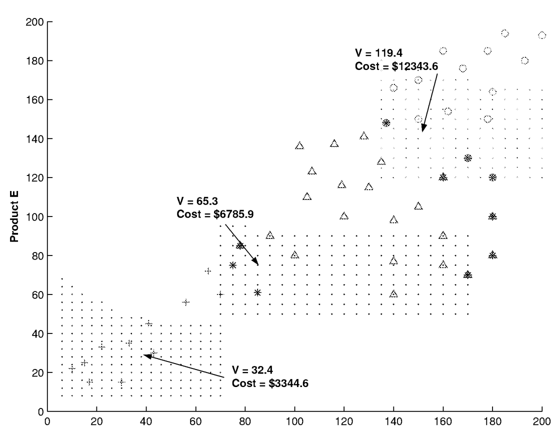

Figure 2: Feasible designs at the first step of the proposed approach

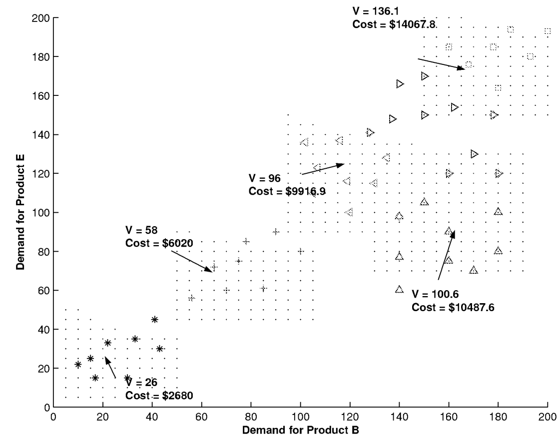

Figure 3: Set of robust design for the five clusters

clustering methodology and stochastic design optimization iteratively allowing re-partitioning of data based on search for an alternative "better" design with respect to the current clustering of data. Research is currently underway regarding a better understanding of the tradeoffs between the different objectives in the robust optimization framework.

Acknowledgments Financial support from the National Science Foundation under the NSF CAREER program CTS-9983406 is gratefully acknowledged

References

Goyal V., and M.G., Ierapetritou, 2002, Determination of Operability Limits using Simplicial Approximation, *AIChE J* , **48**, 2902.

Grossmann, I.E., and R.W.H. Sargent, 1978, Optimal Design of Chemical Plants Design with Uncertain Parameters, *AIChE J* , **24**, 1021.

Kaufman, L., and P.J. Rousseeuw, 1990, *Finding Groups in Data* , Wiley Series, NY.

Mulvey, J.M., R.J., Vanderbei, and S., Zenios, 1995, Robust Optimization of Large-Scale Systems, *Oper. Res.* , **43**, 264.

Pistikopoulou, E.N., and M.G. Ierapetritou, 1995, A Novel Approach for Optimal Process Design under Uncertainty, *Comp. Chem. Eng.* , **19**, 1089.

Prekopa, A., 1995 *Stochastic Programming* , Kluwer Publishers

Rooney, W.C., and L.T., Biegler, 1999, Incorporating Joint Confidence Regions into Design under Uncertainty, *Comp. Chem. Eng.* , **23**, 1563.

Pertsinidis, A., I.E., Grossmann, and G.J., McRae, 1998, Parametric optimization of MILP programs and a framework for the parametric optimization of MINLPs, *Comp. Chem. Eng.* , **22**, S205.

Suh Min-Ho and Lee Tai-Yong, 2001, Robust Optimization Method for the Economic Term in Chemical Process Design and Planning, *Ind. Eng. Chem. Res.* , **40**, 5950.