Chance Constrained Optimization with Robust and Sampling Approximation

by

Zhuangzhi Li

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Process Control

Department of Chemical & Materials Engineering

University of Alberta

Abstract

Uncertainty is pervasive in various process operations problems. Its appearance spans from the detailed process description of multi-site manufacturing. In practical applications, there is a need for making optimal and reliable decisions in the presence of uncertainties. Asking for constraint satisfaction at some level of probability is reasonable in many applications, which calls for the utilization of chance constraints. This thesis studies the approximation methods for solving the chance constrained optimization problems. Two approximation methods were considered: Robust optimization approximation and Sample average approximation. For the robust optimization approximation method, an optimal uncertainty set size identification algorithm is proposed, which can find the smallest possible uncertainty set size that leads to the least conservative robust optimization approximation. For the sample average approximation method, a new linear programming based scenario reduction method is proposed, which can reduce the number of samples used in the sample average approximation problem, thus lead to reduction of computational complexity. Furthermore, the proposed scenario reduction method is computationally more efficient than the existing methods. The effectiveness of the proposed methods is demonstrated by several case studies.

Acknowledgement

I would like to thank a number of people, who have helped me throughout the thesis.

First, I would like to thank Dr. Zukui Li, my supervisor in my research. Li provided valuable guidance which helps me a lot to address the problems I met in the research. He is also the key person responsible for the formation of this thesis.

I also would like to thank Miss Wei Ding and my parents, for their support which gives me courage and confidence to overcome the problems I met during the research.

Thanks also go to my colleagues, Dr. Chi Feng, and Mr. Bang Liu, for their suggestions during the writing stage of the thesis and technical support. Last, but not least, I would like to thank Natural Sciences and Engineering Resource Council (NSERC) of Canada Discovery Grant (RES0016150) for the financial support.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Optimization Problem under Uncertainty

Modern process industries face the increasing pressures for remaining competitive in the global marketplace. To reduce costs, inventories and environmental impact, as well as to maximize profits and responsiveness, a major goal of the process industry is to optimize the process operations in supply, manufacturing and distribution activities. Major activities of process operations include production planning, process scheduling and supply chain management. Those operational activities complement the role of process design and synthesis, and seek to improve existing operating process.

Uncertainty is pervasive in various process operations problems. Its appearance spans from the detailed process description of multi-site manufacturing. The source of uncertainty ranges from orders placed and equipment availability in scheduling problems, to uncertainties in prices and demands for large-scale supply chains. Since decisions made under deterministic assumptions can lead to suboptimal or even infeasible operations, a major interest of process industry is to generate agile and efficient process operations decisions that allow the producer to be more adaptive to uncertainties in manufacturing process and dynamics in the market.

The operation of chemical processes is inherently subject to uncertainty. For instance, the production planning problem needs to consider the availability and prices of raw materials, and the demand for products. If the optimization under uncertainty can be well handled, the efficiency of chemical process operations can be improved, which will lead to a higher profit. A lot of previous studies are focused on the chemical process operations under uncertainty. For example, Bitran et al.[1], Lenstra et al.[2] and Escudero et al.[3] studied the production planning problem under uncertainty; Liu and Sahinidis[4], Acevedo and Pistikopoulos[5] and Gupta and Maranas[6] investigated the process design optimization of chemical processing systems; Birge et al.[7] and Tayur et al.[8] studied the process scheduling problem under uncertainty.

Except process operations, many other realistic decision making problems also face data uncertainty in optimization. Among those problems, water reservoir management is a typical case. Water reservoirs are natural or artificial lake, storage pond, or impoundment of a dam which is used to store water. Water reservoirs are widely used for direct water supply, hydroelectricity, controlling watercourses, flow balancing, and recreation. However, water reservoirs are highly impacted by the environment, such as climate change, human impact, and demand uncertainties. Therefore, it is critical to conduct efficient water management for the poverty of water resources. The study of water management optimization under uncertainty can be found in Dupačová et al.[9], Ouarda et al.[10], Sreenivasan et al.[11], and Dhar and Datta[12].

Along with the development of economy, financial risk management has received more and more attention. This encourages the study of portfolio optimization, which is the process of choosing the proportions of various assets to be take place in a portfolio according to some criteria, such as the expected value of the portfolio's rate of return, and the possibility of it. However, portfolio optimization is easy to be affected by some uncertainty parameters, such as commodity prices, interest rates, exchange rates, and the random return. Hence, it is essential to conduct optimization for portfolio selection problems under uncertainty to guarantee the maximum return of the total investment with risk control. This kind of study can trace back to Markowitz[13]. Further study can be found in Ermoliev et al.[14], Bonami et al.[15], and Pagnoncelli et al.[16].

Uncertainty poses major challenges to decision making for the above different problems. The challenges lie in not only the modeling/representation of the uncertainty but also in the decision making under uncertainty. Those challenges include: 1) lack of efficient scenario generation and reduction technique for modeling uncertainty; 2) difficulty in solving large scale stochastic optimization problems; 3) complexity in accounting for correlations between uncertain parameters; and 4) lack of simple and intuitive modeling and solution platform for optimization under uncertainty. In practical applications, there is a need for making optimal and reliable decisions in the presence of uncertainties. In many applications, asking for constraint satisfaction at some level of probability is reasonable, which calls for the utilization of chance constraints (probabilistic constraints). Chance constraints can be used to model the degree of constraint violation tolerance, the level of satisfaction. Accordingly, chance constraints can be used for modeling restrictions on decision making with the appearance of uncertainty, which induces methods called chance constrained optimization (CCP).

1.2. Chance Constrained Optimization

Chance constraint (also called probabilistic constraint) is an important tool for modeling reliability on decision making in the presence of uncertainty. A general chance constrained optimization problem (CCP) takes the following form

$$\begin{aligned} &\max_{x \in X} f(x) \\ &s.t. \ \Pr\left\{h(x,\xi) \le 0\right\} \ge 1-\alpha \end{aligned} \tag{1}$$

where x represents the decision variables, ξ denotes the uncertain parameters, α is a reliability parameter representing the allowed constraint violation level ($0 < \alpha < 1$). The chance constraint $\Pr\left\{h(x,\xi) \le 0\right\} \ge 1-\alpha$ enforces that the constraint $h(x,\xi) \le 0$ is satisfied with probability $1-\alpha$ at least (or violated with probability α at most).

Chance constrained optimization problem was introduced in the work of Charnes et al.[17] and an extensive review can be found in Prékopa[18]. There are many challenging aspects of solving chance constrained optimization problem. It is very hard to evaluate the chance /probabilistic constraints when solving the CCP problems, for the requirement of a multi-dimensional integration. Therefore, Monte-Carlo simulation is the only way. Some breakthrough innovations have been achieved in recent years. For example, Alizadeh and Goldfarb[19] gave the theory and practice that the individual chance constraint can be transformed into a second order cone, Lagoa[20] proved that the individual chance constrained problem is convex under the uniform distribution over a convex symmetric set, and Calafiore and El Ghaoui[21] showed that the individual chance constraints can be converted to second-order cone constraints under radial distribution. Except a few specific probability distributions (e.g. normal distribution), it is difficult to formulate an equivalent deterministic constraint and the feasible region of chance constrained optimization problem is often nonconvex. To avoid the above difficulties, existing methods for solving chance constrained optimization problem largely rely on solving an approximation problem. Generally, there are two types of approximation methods in the literature to approximate a chance constraint: analytical approximation approach and sampling based approach.

Analytical approximation approach transforms the chance constraint into a deterministic counterpart constraint. Compared to the sampling based approximation, analytical approximation provides safe approximation and the size of the model is independent of the required solution reliability. There are several ways to perform the transformation from chance constraints to deterministic constraints.

Conditional value of risk (CVaR) is one of the transformation methods, which is a risk measurement. It is used in finance (and more specifically in the field of financial risk measurement) to evaluate the market risk or credit risk of a portfolio. CVaR theory have been explicitly described in the work of Stambaugh[22], Pritsker[23] and Philippe[24]. It is a substitute for value at risk (VaR) that is more sensitive to the shape of the loss distribution in the tail of the distribution. Based on the contribution of Artzner et al.[25], Ogryczak[26], RockFellar, and Uryasev[27, 28], conditional value-at-risk theory poses a potentially attractive alternative to the probabilistic constrained optimization framework and other uncertainty approaches. Nemirovski and Shapiro[29] give the convex approximation for probabilistic constraints by using CVaR. Verderame and Floudas[30] extend the work to the robust optimization problems, which have been widely used to solve chance constrained optimization problems. Other deterministic approximation of individual chance constraints include using Chebyshev's inequality[31], Bernstein inequality[32-34], and Hoeffding's inequality[35] as the bound of the probability of constraint violation.

Robust optimization (RO) provides another way for analytically approximating a chance constraint. Robust optimization often requires only a mild assumption on probability distributions, and it provides a tractable approach to obtain a solution that remains feasible in the chance constrained problem. Hence, robust optimization has been widely used to construct a safe approximation for chance constraint. One of the earliest papers on robust counterpart optimization is the work of Soyster[36]. Li et al.[37] studied the robust counterpart optimization techniques for linear optimization and mixed integer linear optimization problems. Even though robust optimization can provide a safe approximation to the chance constrained problem, the quality of the approximation has not received attention in the existing literature. A safe approximation can be unnecessarily conservative and lead to a solution that is of bad performance in practice. In this thesis, a two-step algorithm is proposed to address the optimal robust optimization approximation.

For the sampling based approach, random samples are drawn from the probability distribution of the uncertain parameters and they are further used to approximate the chance constraint. Scenario approximation and sample average approximation are two different ways of sampling based methods.

Sample average approximation (SAA) uses an empirical distribution associated with random samples to replace the actual distribution, which is further used to evaluate the chance constraint. By generating a set of samples $\xi^1, \xi^2, \ldots, \xi^K$ of the random parameters ξ and approximate the chance constraint with a new approximated constraint $\frac{1}{K}\sum_{i=1}^{K} 1_{(0,\infty)}\left(h\left(x, \xi^i\right)\right) \le \alpha$. Here, $1_{(0,\infty)}\left(x\right)$ is an indicator function. This kind of constrained problems has been investigated by Luedtke & Ahmed[38], Atlason et al.[39] and Pagnoncelli et al.[40]. An essential aspect of the sample average approximation is the choice of the generated sample size *K*. If the sample size is too big, the computational complexity will be very high, while if the sample size is too small, the reliability of the solution will be very low. Lately, Li and Floudas[41] proposed a novel method for scenario reduction. The scenario reduction is formulated in a mixed integer linear optimization problem. It aims at selecting a small number of scenarios to represent the entire set of possible scenarios. However, the complexity of this method will be very high when the size of the entire set of scenarios is very large. A new scenario reduction method, which is based on a linear programming (LP) problem, is proposed in this thesis.

Scenario approximation is another sampling based approximation method for solving chance constrained problems. The general idea is to generate a set of samples $\xi^1, \xi^2, \ldots, \xi^K$ of the random parameters ξ and approximate the chance constraint with a set of constraints $h\left(x, \xi^k\right) \le 0, k = 1, \ldots, K$. The scenario approximation itself is random and its solution may not satisfy the chance constraint. Research contributions in this direction have been made by Calafiore et al.[42], Luedtke et al.[38], and Nemirovski et al.[43].

1.3. Thesis Objective

The objective of this thesis is to solve the chance constrained optimization problems with robust optimization approximation method and sample average approximation method. Section 2 presents the optimal robust optimization approximation method, and demonstrates the effectiveness of the proposed method with three case studies: portfolio optimization problem, production planning problem, and process scheduling problem. A linear programming based scenario reduction method is proposed in Section 3. It aims at selecting a small number of scenarios to represent the original large number of scenarios. Although a mixed integer linear optimization (MILP) based scenario reduction method was proposed before, its computational time can be very large when the original number of scenarios is very large. The proposed LP based scenario reduction method can efficiently reduce the computational complexity of sample average approximation problem. In Section 4, sample average approximation is applied to solve the chance constrained problems with the proposed scenario reduction method. Five problems are considered including three linear problems and two nonlinear problems. Conclusion and future work are presented in Section 5.

2. Optimal Robust Optimization Approximation

2.1. Robust Optimization

2.1.1. Problem Formulation

In this thesis, linear constraint under uncertainty is investigated. Consider the following optimization problem with parameter uncertainty:

$$\begin{aligned} &\max_{x\in X} cx \\ &s.t. \quad \sum_j \tilde{a}_j x_j \le b \end{aligned} \tag{2}$$

where the constraint coefficients $\tilde{a}_j$ are subject to uncertainty. Define the uncertainty as $\tilde{a}_j = a_j + \xi \hat{a}_j, \forall j \in J$, where j is the index of uncertainty parameters, a_j represent the nominal value of the parameters, $\hat{a}_j$ represent positive constant perturbations, ξ_j represent independent random variables which are subject to uncertainty and J represents the index subset that contains the variables whose coefficients are subject to uncertainty. Constraint in (2) can be rewritten by grouping the deterministic part and the uncertain part as follows:

$$\sum_j a_j x_j + \sum_{j\in J} \xi_j \hat{a}_j x_j \le b \tag{3}$$

In the set induced robust optimization method, the aim is to find solutions that remain feasible for any ξ in the given uncertainty set U with size Δ so as to immunize against infeasibility. The corresponding robust optimization is

$$\begin{aligned} &\max_{x\in X} \ cx \\ &\text{s.t.} \quad \sum_j a_j x_j + \max_{\xi\in U(\Delta)} \left\{ \sum_{j\in J} \xi_j \hat{a}_j x_j \right\} \le b \end{aligned} \tag{4}$$

The formulation of the robust counterpart optimization problem is connected with the selection of the uncertainty set U. Based on the work of Li et al.[37], different robust counterpart optimization formulations can be developed depending on the type of uncertainty set. For example, the box uncertainty set $U_\infty = \left\{ \xi \;\middle|\; |\xi_j| \le \Psi,\ \forall j \in J \right\}$ induced robust counterpart optimization constraint is given by:

$$\sum_j a_j x_j + \Psi \sum_{j \in J} \hat{a}_j |x_j| \le b \tag{5}$$

And the ellipsoidal uncertainty set $U_2 = \left\{ \xi \middle| \sum_{j \in J} \xi_j^2 \le \Omega^2 \right\}$ induced robust counterpart optimization constraint is:

$$\sum_j a_j x_j + \Omega \sqrt{\sum_{j \in J} \hat{a}_j^2 x_j^2} \le b \tag{6}$$

where Ψ and Ω are the size of the box and ellipsoidal uncertainty set, respectively.

2.1.2. Relationship between Objective Value and Uncertain Set Size

For the same type of uncertainty set, as the set size increases, the optimal objective of the robust optimization problem (4) will decrease (for a maximizing objective) because the feasible region of the robust optimization problem (4) becomes smaller. This is shown in the following motivating example.

$$\begin{aligned}
&\max_{x \ge 0} \; c^T x \\
&\text{s.t.} \quad -\tilde{a}^T x \le 0.2 \\
&\qquad \sum_{i=1}^{5} x_i = 1
\end{aligned}$$

where $c = [0.00347 \quad -0.00126 \quad -0.00476 \quad 0.00094 \quad -0.0876]^T$. Parameter $\tilde{a}_i$ are subject to uncertainty and they are defined as: $\tilde{a}_i = a_i + \xi_i \hat{a}_i,\ i = 1,\ldots,5$, with $a_i = c_i,\ \hat{a}_i = 0.1 a_i$, and ξ_i are the random parameters.

For the uncertain constraint, the box type uncertainty set induced robust counterpart optimization constraint can be formulated as

$$-\sum_i a_i x_i + \Psi \sum_i \hat{a}_i x_i \le 0.2$$

where Ψ is the size of the uncertainty set. For each fixed Ψ, the resulting robust optimization problem is a deterministic linear optimization problem. The optimal objective value of the robust optimization problem is solved for different Ψ. The relationship between the uncertainty set size and the optimal objective value is plotted in Figure 1, which shows that the optimal objective value is a monotonically decreasing function of the uncertainty set size.

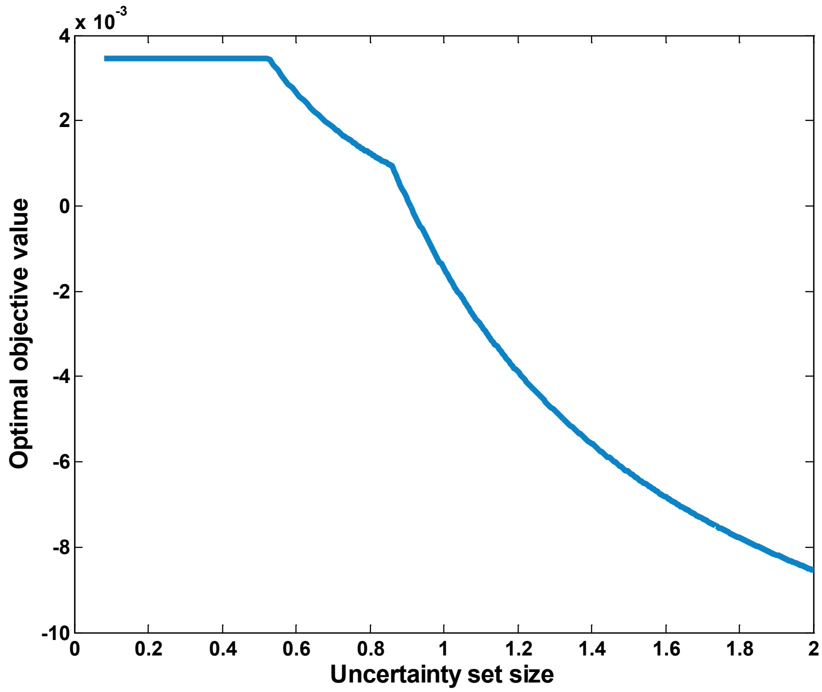


Figure 1. Relationship between optimal objective value and uncertainty set size

2.2. Robust Optimization Approximation to Chance Constraint

Chance constraint models the solution reliability in an optimization problem. For the uncertain constraint in problem (2), its chance constrained version can be formulated by setting a lower bound on the probability of constraint satisfaction:

$$\max_{x \in X} \; cx$$
$$\text{s.t.} \quad \Pr\left\{ \sum_{j} a_j x_j + \sum_{j \in J} \xi_j \hat{a}_j x_j \le b \right\} > 1 - \alpha \tag{7}$$

or equivalently as follows by setting an upper bound on constraint violation:

$$\max_{x \in X} \; cx$$
$$\text{s.t.} \quad \Pr\left\{ \sum_{j} a_j x_j + \sum_{j \in J} \xi_j \hat{a}_j x_j > b \right\} \le \alpha \tag{8}$$

where parameter α denotes the reliability level $(0 < \alpha < 1)$.

If the box type uncertainty set induced robust counterpart is applied, and the corresponding robust optimization problem can be formulated:

$$\max_{x \ge 0} \; c^T x$$
$$\text{s.t.} \quad -\sum_{j \in J} a_j x_j + \Psi \sum_{j \in J} \hat{a}_j x_j \le b \tag{9}$$

and if the ellipsoidal uncertainty set induced robust counterpart is applied, and the corresponding robust optimization problem can be formulated:

$$\max_{x \ge 0} \; c^T x$$
$$\text{s.t.} \quad -\sum_{j \in J} a_j x_j + \Omega \sqrt{\sum_{j \in J} \hat{a}_j^2 x_j^2} \le b \tag{10}$$

2.3. Reliability Quantification

2.3.1. *A Priori* Probability Bound Based Method

Under certain assumptions on the distributions of the uncertainty, the reliability of robust solution can be qualified by the so-called a *priori* probability bound[44], which is a function of the uncertainty set size and provides an upper bound on probability of constraint violation. If the uncertainty set's size Δ satisfies $p_{violation}^{prioriUB}\left(\Delta\right) \leq \alpha$, then:

$$\Pr\left\{\sum_{j} a_j x_j + \sum_{j \in J} \xi_j \hat{a}_j x_j > b\right\} \leq p_{violation}^{prioriUB}(\Delta) \leq \alpha \tag{11}$$

or a lower bound on constraint satisfaction:

$$\Pr\left\{\sum_{j} a_j x_j + \sum_{j \in J} \xi_j \hat{a}_j x_j \leq b\right\} > p_{satisfaction}^{prioriLB}(\Delta) = 1 - p_{violation}^{prioriUB}(\Delta) \tag{12}$$

For example, if ξ_j are independent and subject to bounded and symmetric probability distribution supported on [-1,1], then for the box and ellipsoidal uncertainty sets induced robust counterparts (5), (6), one valid *a priori* upper bound is given by Li et al.[44]:

$$p_{violation}^{prioriUB}(\Delta) = \exp\left(-\frac{\Delta^2}{2}\right) \tag{13}$$

where Δ represents the size of the uncertainty set.

While the *a priori* probability bound based robust optimization provides safe approximation to chance constraint, it is a conservative approximation since the feasible set of the robust optimization problem is always less than the feasible set of the original chance constrained problem as seen from equation (11). In other words, the optimal solution of robust optimization problem will be always less than the true optimum of the chance constrained problem (for a maximizing objective).

2.3.2. *A Posteriori* Probability Bound Based Method

To find less conservative robust optimization approximation that still satisfies the desired probability of constraint satisfaction, a tighter probability upper bound on constraint violation can be used.

With a given solution x^* to the robust optimization problem, the corresponding probability of constraint violation can be quantified by the *a posteriori* probability bound[44]. If the probability distribution information of the uncertain parameters is known, then the following relationship holds:

$$\Pr\left\{\sum_{j} a_j x_j + \sum_{j\in J} \xi_j \hat{a}_j x_j > b\right\} \leq p_{violation}^{posterioriUB}(x^*) \tag{14}$$

or

$$\Pr\left\{\sum_{j} a_j x_j + \sum_{j\in J} \xi_j \hat{a}_j x_j \leq b\right\} \geq p_{violation}^{posterioriLB}(x^*) = 1 - p_{violation}^{posterioriUB}(x^*) \tag{15}$$

where

$$p_{violation}^{posterioriUB}(x^*) = \min_{\theta>0} \ \exp\left(-\theta(b - \sum\nolimits_j a_j x_j^*) + \sum_{j\in J} \ln E[e^{\theta \xi_j \hat{a}_j x_j^*}]\right) \tag{16}$$

The above *a posteriori* probability bound is in general tighter than the *a priori* probability bound. The illustration will be shown later.

2.3.3. Iterative Method

In order to improve the quality of the solution when robust optimization approximation is used to approximation chance constraints, Li and Floudas[45] provided an iterative method which compromise the aforementioned two methods. This method combined the *a priori* probability bound and the *a posteriori* probability bound. The initial size of the uncertainty set is determined by the *a priori* probability bound, and then use the *a posteriori* probability bound to adjust the size of the uncertainty set. This iterative method provides a heuristic way to improve the robust solution quality. Specifically, if the probability calculated by the *a posteriori* probability bound is larger than the desired level, the set size should be decreased; if the probability is smaller than the desired level, the set size should be increased. The adjustment of the set size can lead to an improved robust solution from the set induced robust optimization problem, and the solution feasibility is guaranteed for the original chance constrained problem. Illustration of this method will be shown in the next section.

2.3.4. Application of Reliability Quantification Methods

The motivating example in section 2.1.2 is reconsidered here with a chance constrained version of the uncertain constraint with $\alpha = 0.5$. The corresponding chance constrained model can be written as follows:

$$\begin{aligned}
&\max_{x\geq 0}\ c^T x \\
&\text{s.t.}\quad \text{Prob}\left\{-\tilde{a}^T x > 0.2\right\} < 0.5 \\
&\qquad\ \sum_{i=1}^{5} x_i = 1
\end{aligned}$$

Assume that all the random parameters ξ_i are uniformly distributed in [-1, 1]. For this type of uncertainty distribution, it is hard to obtain an equivalent deterministic formulation of the chance constraint. Robust optimization approximation is used here to solve the chance constrained problem. The original chance constraint is replaced with the box type uncertainty set induced robust counterpart, and the corresponding robust optimization problem can be formulated:

$$\begin{aligned}
&\max_{x\geq 0}\ c^T x \\
&\text{s.t.}\quad -\sum_{i=1}^{5} a_i x_i + \Psi \sum_{i=1}^{5} \hat{a}_i x_i \leq 0.2 \\
&\qquad\ \sum_{i=1}^{5} x_i = 1
\end{aligned}$$

Following the aforementioned approximation method in section 2.3.1, the *a priori* probability bound based method, the size of the uncertainty set is determined by the *a priori* probability bound $\exp(-\Psi^2/2) \leq 0.5$ and the result is $\Psi \geq 1.1774$. To make the approximation less conservative, we choose $\Psi = 1.1774$ and solve the robust optimization problem. The optimal objective value of the robust optimization problem is 3.66×10^{-3}. To verify the solution reliability, Monte Carlo simulation is used to estimate the true probability of constraint violation with the obtained robust solution. A total number of N=100000 samples are generated based on the uncertainty distribution. The constraint is then evaluated and the number of times that the constraint is violated is recorded as *V*. The probability of constraint violation is estimated as *V*/*N*. For the above robust solution, the estimated probability of constraint violation is 0.0662. It is seen that this value is much less than the target 0.5, which means that the approximation is relatively conservative.

The above example illustrates a limitation of the traditional *a priori* probability bound based robust optimization approximation to chance constraint: the approximation can be very conservative such that the true probability of constraint violation is much smaller than the target.

As for the *a posteriori* probability bound based method, for the motivating example, the *a posteriori* probability lower bound given by (15) and (16) and the *a priori* probability lower bound given by (12) and (13) on constraint satisfaction are plotted against uncertainty set size in Figure 2, which shows that the *a posteriori* bound is tighter.

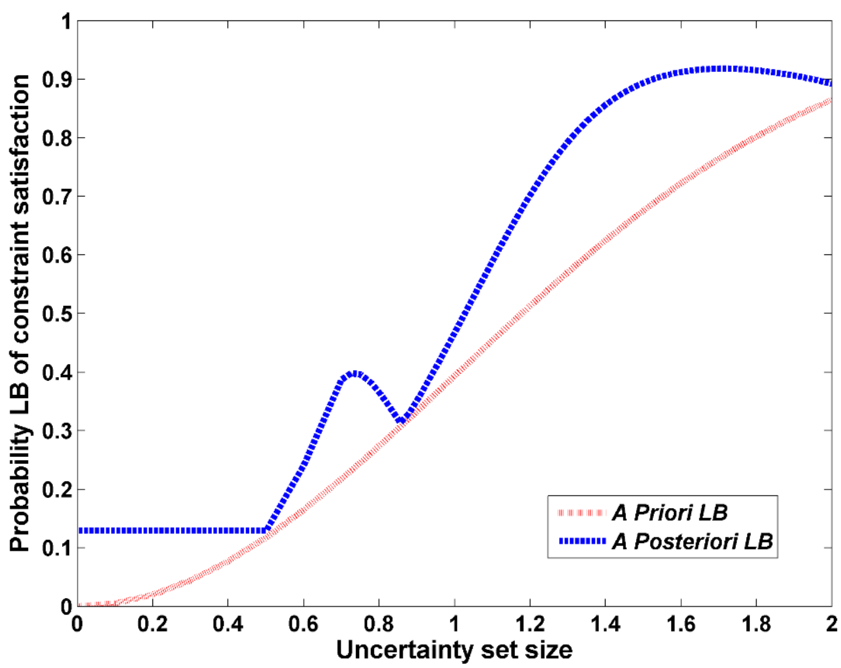


Figure 2. A posteriori and a priori probability lower bound of constraint satisfaction

Based on this observation, the size of the uncertainty set can be adjusted and the robust solution can be improved in an iterative process, until the *a posteriori* probability bound of the constraint satisfaction is close to the target level in the chance constraint, which is the iterative method[45].

Based on the iterative method proposed in Li and Floudas[45], the following solution can be obtained as shown in Table 1. The results show that the solution quality is improved (i.e. the objective value is increased from -3.66×10^{-3} to 3.01×10^{-3}) and the solution reliability still satisfies the requirement (i.e., the solution has a probability of constraint violation less than 0.4961 and the target is less than 0.5).

Table 1. Solution procedure of iterative method[45]

k	Ω_k	$\exp(-\Omega^2/2)$	Obj*	$prob_{violation}^{posterioriUB}$
1	1.1774	0.5	-3.66×10^{-3}	2.26×10^{-6}
2	0.5887	0.8409	2.77×10^{-3}	0.4117
3	0.2944	0.9576	3.47×10^{-3}	0.6325
4	0.4415	0.9071	3.47×10^{-3}	0.6325
5	0.5151	0.8758	3.47×10^{-3}	0.6325
6	0.5519	0.8587	3.17×10^{-3}	0.5464
7	0.5703	0.8499	2.96×10^{-3}	0.4791
8	0.5611	0.8543	3.06×10^{-3}	0.5130
9	0.5657	0.8521	3.01×10^{-3}	0.4961

While the *a posteriori* probability bound can be used to improve the robust solution quality, it can also be used to extend the application of robust optimization approximation to general uncertainty distributions. Notice that the *a priori* probability bounds are obtained based on certain assumptions on the uncertainty distribution (e.g., bounded and symmetric).

For the motivating example, it can be assumed that the uncertain parameters $\tilde{\mathbf{a}} = \left[\tilde{a}_1, \tilde{a}_2, \tilde{a}_3, \tilde{a}_4, \tilde{a}_5\right]^{\mathrm{T}}$ are subject to independent normal distribution (i.e., it is not a bounded distribution) with mean vector $\boldsymbol{\mu} = \left[0.00347,\ -0.00126,\ -0.00476,\ 0.00094,\ -0.0876\right]^T$ and variance $\boldsymbol{\sigma}^2 = \left[0.1494,\ 0.0818,\ 0.0923,\ 0.0546,\ 0.0086\right]^T$. Although the *a priori* probability bound is not applicable to this distribution, the iterative algorithm can still be applied and it leads to a solution that satisfies the desired reliability, as shown in Table 2.

Table 2. Solution procedure of the iterative method

k	Ω_k	$\exp(-\Omega^2/2)$	Obj*	$prob_{violation}^{posterioriUB}$
1	1.1774	0.5	-3.66×10^{-3}	0.3224
2	0.5887	0.8409	2.77×10^{-3}	0.7798
3	0.8831	0.6771	4.89×10^{-4}	0.6670
4	1.0302	0.5882	-1.90×10^{-3}	0.4965

2.4. Optimal Robust Optimization Approximation

2.4.1. Relationship between Uncertainty Set Size and Robust Solution Quality

While the applicability of robust optimization can be extended and the quality of robust solution can be improved with the *a posteriori* probability bound as shown in section 2.3, a natural question on the robust optimization approximation for chance constrained optimization problem can be raised: What is the best possible robust optimization approximation?

Considering a solution obtained from robust optimization approximation, while the solution reliability satisfies the original chance constraint, the optimal objective value of the robust approximation problem should be as close to the true optimum of the chance constrained optimization problem as possible. This also means that the uncertainty set should be designed as small as possible while the reliability of the solution is satisfied.

In the development of an algorithm for finding the best uncertainty set size, the following issues need to be considered:

1) For a robust optimization approximation problem with the same type of uncertainty set, is the solution more reliable when the set size is larger?
2) Can the robust solution's reliability reach the desired level by adjusting the size of the uncertainty set? If not, what is the maximum possible reliability that the robust solution can reach?
3) If the robust solution's reliability can reach the desired level, what will be the set size that leads to the best robust solution (i.e., best objective value)?

The above issues will be investigated through the motivating example.

To study the relationship between solution and the size of uncertainty set, a study is made between the solution optimality (i.e., the optimal objective value), the solution reliability (i.e., the probability of constraint satisfaction) and the uncertainty set size. Here, it is assumed that the uncertainty is following the normal distribution introduced in section 2.3.4. The rest studies on the motivating example will be also based on this assumption.

By varying the uncertainty set size, the *a priori* probability lower bound, the *a posteriori* probability lower bound, and the true probability (estimated by simulation) of constraint satisfaction are plotted together in Figure 3. The following observations can be made. While both the *a priori* and the *a posteriori* probability bounds underestimate the true probability of constraint satisfaction, the *a priori* probability lower bound is more conservative. The true probability of constraint satisfaction is not a monotonically increasing function of the uncertainty set size. That is, a larger uncertainty set may not necessarily lead to a robust solution with higher reliability.

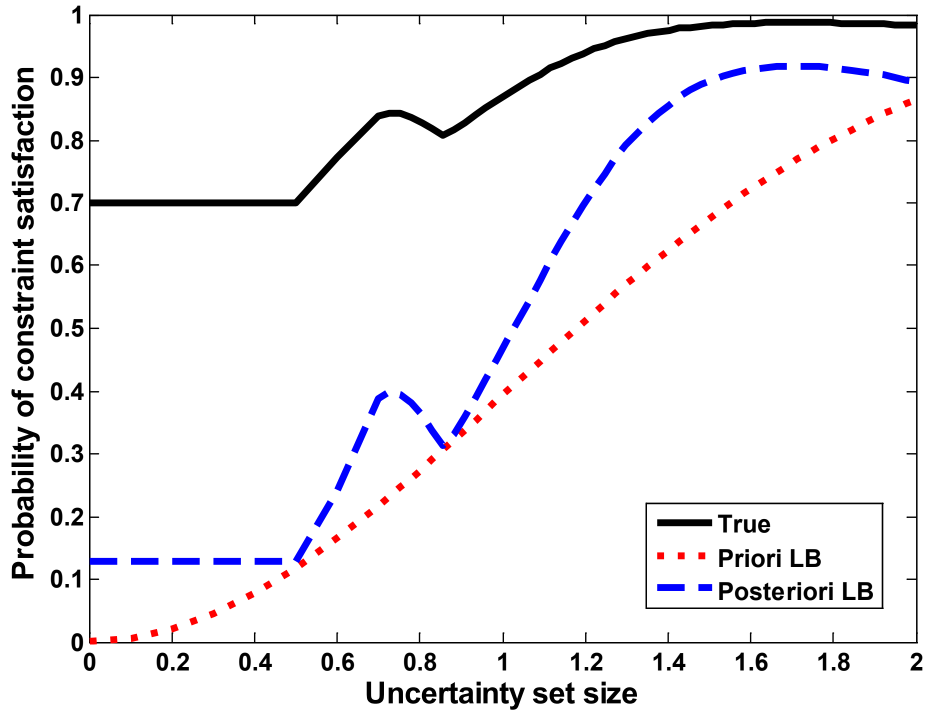


Figure 3. Comparison of probabilities (bounds) of constraint violation

Figure 3 shows that the true probability of constraint satisfaction and the *a posteriori* probability lower bound demonstrate strong (piecewise) nonlinear behavior as a function of the uncertainty set size. This can be explained from the parametric programming point of view. Since the uncertainty set size appears as the left-hand-side constraint coefficient (e.g., Ψ in (5)), the optimal solution of the robust optimization problem is a nonlinear parametric function of the set size parameter and the function form varies in different critical regions (e.g., different intervals of Ψ). Furthermore, since both the true probability and the *a posteriori* probability bound are a function of the optimal solution, so both of them demonstrate piecewise nonlinear behavior of the uncertainty set size parameter.

Next, the true probability of constraint satisfaction is plotted against the optimal objective value of the robust solution. As shown in Figure 4, the black curve represents the solution from box type uncertainty set induced robust optimization approximation problem. To show that the robust optimization approximation is conservative approximation, the true optimal objective of the chance constrained problem and the corresponding probability of constraint satisfaction is also plotted in green curve. The optimal objective of chance constrained problem is obtained by replacing the chance constraint with its deterministic equivalent (which can be explicitly derived since the uncertainty distribution is multivariate normal). It is seen that the black curve representing robust solution is always under the solution curve for the chance constraint problem, which shows that the robust optimization provides a conservative approximation to the original chance constrained problem.

If a target reliability level 0.82 (i.e., $\alpha = 0.18$) is set for the chance constraint, as shown by the blue dash line in Figure 4, the following observations can be made. All the black points, above or on the blue line, represent the robust solutions that are feasible to the original chance constraint problem. There are three robust solutions on the blue line, that is, they exactly match the desired reliability level. Among all the feasible solutions, the best possible robust solution is the one with the largest objective value (i.e., point "1" in Figure 4). So, point "1" corresponds to the optimal robust optimization approximation problem.

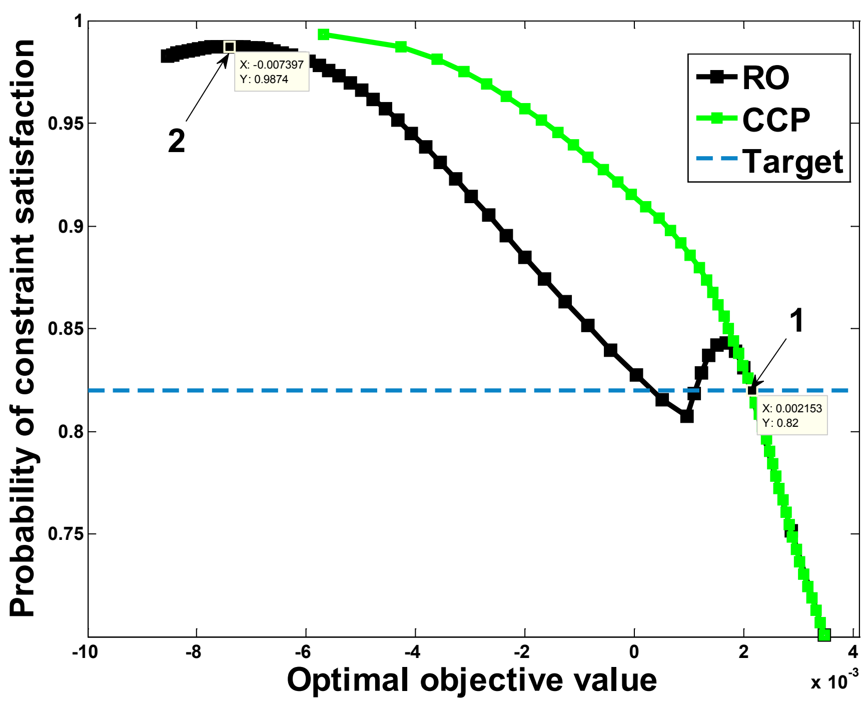


Figure 4. Comparison of solution reliability versus optimal objective value

If the target reliability level in the chance constraint is set as 0.995 (i.e., $\alpha = 0.005$), then a conclusion can be made from Figure 4 that it is impossible to obtain a robust solution that satisfies the target. The reason is that the largest possible probability of constraint satisfaction from a robust solution is 0.9874 (i.e., point "2" in Figure 4). In this case, the best option is to return this best possible robust solution and report its reliability.

Based on the above observations, algorithms are developed in the next subsection for identifying the optimal robust optimization approximation (i.e., optimal uncertainty set's size) for chance constrained optimization problem.

2.4.2. Algorithm for Optimal Robust Optimization Approximation

Based on the analysis in previous sections, we formally post the optimal robust optimization approximation problem as follows:

Case 1: If the robust optimization approximation based solution can lead to the desired probability of constraint satisfaction, then the minimum possible set size should be identified (e.g., point "1" in Figure 4 or point "a" in Figure 5 with the target reliability level 0.82). The corresponding optimal robust optimization approximation problem can be stated as the following optimization problem

$$\begin{aligned} &\min \ \Delta \\ &s.t. \quad p_{satisfaction}^{true}(\Delta) \geq 1-\alpha \end{aligned} \tag{17}$$

which is also equivalent to identifying the first root of equation $prob_{satisfaction}^{true}(\Delta)-(1-\alpha)=0$.

Case 2: If the robust optimization approximation based solution can not lead to the desired probability of constraint satisfaction, then the maximum probability of constraint satisfaction and the corresponding set size should be identified (e.g., point "2" in Figure 4 or point "b" in Figure 5 with the target reliability level 0.995). The corresponding optimal robust optimization approximation problem can be stated as follows:

$$\max_{\Delta} \ p_{satisfaction}^{true}(\Delta) \tag{18}$$

In the above problem definition, $obj(\Delta)$ is the optimal objective of the robust optimization problem when the size of the uncertainty set is Δ, and $p_{satisfaction}^{true}(\Delta)$ is the probability of constraint satisfaction with the robust solution obtained from the robust optimization approximation problem (the probability of constraint violation is $p_{violation}^{true} = 1 - p_{satisfaction}^{true}$). The probability value $p_{satisfaction}^{true}(\Delta)$ is estimated as follows. Given a solution x^* for problem (1), we would like to quantify the probability of constraint violation

$$p_{violation}^{true} = \Pr\left\{h(x^*, \xi) > 0\right\} \tag{19}$$

The above probability is estimated with Monte Carlo sampling technique by testing feasibility of N samples on the constraint, and the estimation to $p_{violation}^{true}$ is given by

$$\tilde{p}_{violation} = \frac{1}{N}\sum_{i=1}^{N} 1_{(0,\infty)}\left(h(x^*, \xi)\right) \tag{20}$$

Where $1_{(0,\infty)}(t) = \begin{cases} 1, & \text{if } t > 0 \\ 0, & \text{if } t \le 0 \end{cases}$.

Since the evaluation procedure is based on the known value of x, the simulation can be performed with a relatively large sample size so as to get a reliable estimation.

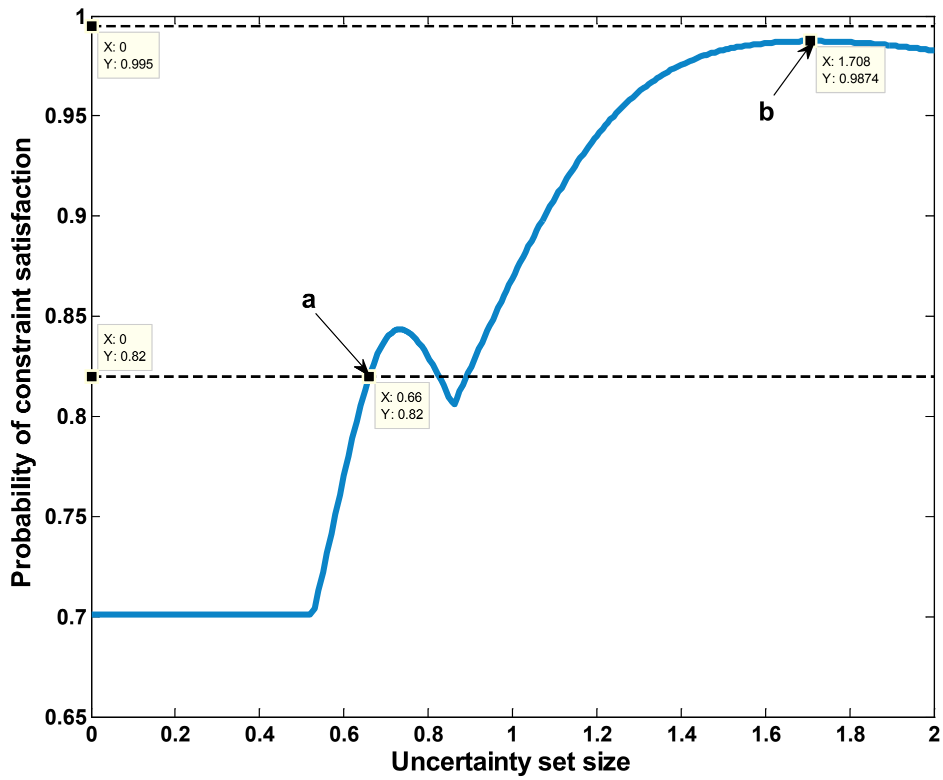
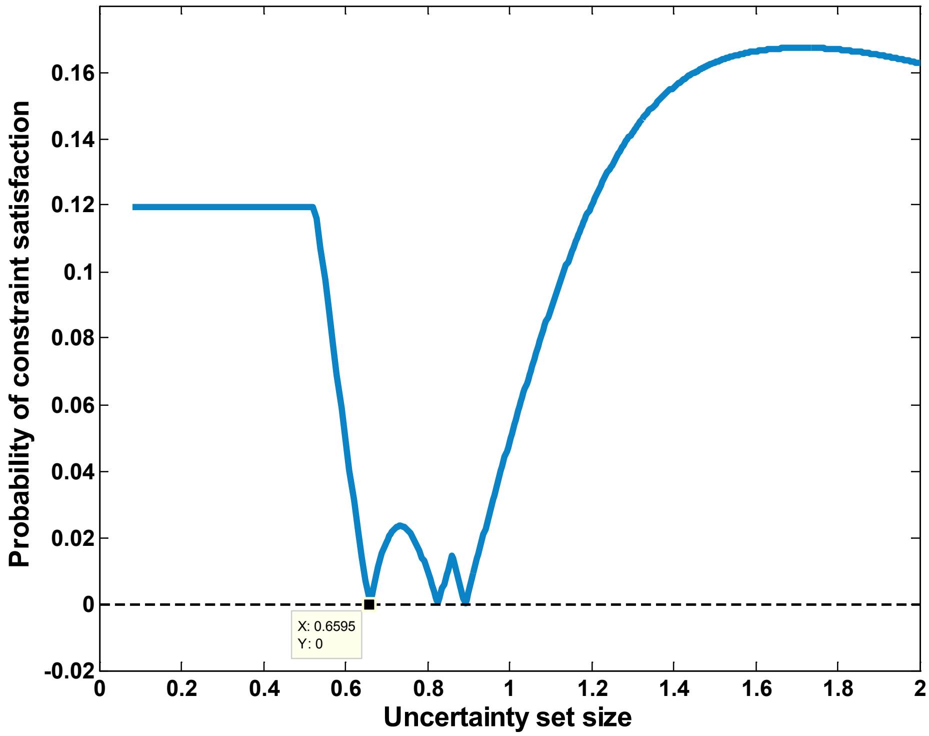


Figure 5. Solution reliability as a function of the uncertainty set size

For both case defined above for the optimal robust optimization approximation, finding the optimal set size is equivalent to finding the first global minimum of the absolute difference between the solution reliability and desired reliability:

$$\min_{\Delta} \left| prob_{satisfaction}^{true}(\Delta) - (1-\alpha) \right| \tag{21}$$

For the motivating example, the above objective function is shown in Figure 6. If the first global minimum objective value is 0, then it means the robust solution can reach the desired reliability (case 1). Otherwise, it means the robust optimization approximation cannot reach the desired reliability (case 2), and the first global minimum solution corresponds to the maximum possible reliability from the robust optimization approximation.


Probability of constraint satisfaction
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
X: 0.6595
Y: 0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Uncertainty set size

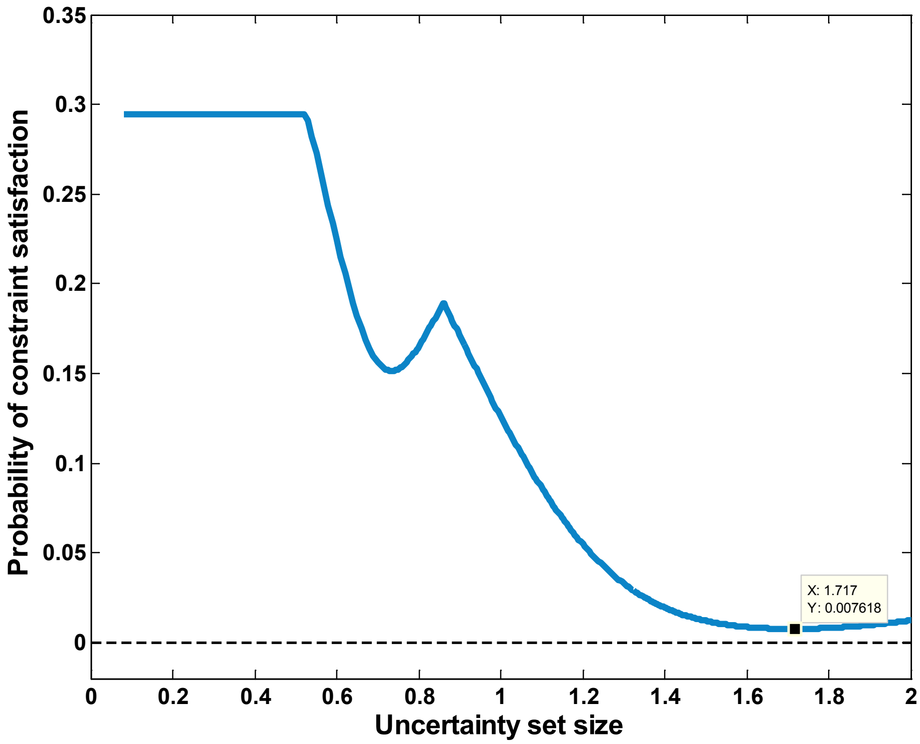

Probability of constraint satisfaction
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
X: 1.717
Y: 0.007618
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Uncertainty set size

Figure 6. Equivalent objective function for optimal set size identification

(Top: α=0.18; Bottom: α=0.005)

To identify the optimal robust optimization approximation, a novel two-step algorithm is proposed in this section, where an upper bound of the optimal set size is identified first, and the optimal set size is identified next.

First, if the uncertainty set size for a robust optimization problem is too large, then the robust optimization problem can be infeasible. So it is meaningful to identify the maximum set size that makes the robust optimization problem feasible. The proposed algorithm for the maximum set size identification is based on checking the feasibility of the robust optimization problem followed by bound contraction. Starting from a feasible set size (lower bound) and an infeasible set size (upper bound), the algorithm will gradually reduce the interval until the predefined tolerance is satisfied. The algorithm is summarized in Table 3.

Table 3. Identify a feasible upper bound of the optimal set size

Algorithm: Feasible upper bound identification
1. Initialization $[lb, ub]$
where lb is a set size that makes the RO feasible (e.g., 0)
where ub is a set size that makes the RO infeasible
2. While $\lvert ub - lb \rvert > \varepsilon$
3. solve RO with set size $\Delta = (lb + ub)/2$
4. if RO problem is feasible
5. $lb = \Delta$
6. else (RO problem is infeasible)
7. $ub = \Delta$
8. end
9. End
10. Return lb as maximum feasible set size

For the motivating example, the following solution procedure shown in Table 4 is used to obtain the maximum feasible set size.

Table 4. Solution procedure for the numerical example

Iteration	$[lb, ub]$	RO feasibility with $\Delta = (lb + ub)/2$
1	[0, 4]	feasible
2	[2, 4]	feasible
3	[2, 3]	infeasible
4	[2, 2.5]	infeasible
5	[2, 2.25]	infeasible
6	[2, 2.125]	infeasible
7	[2.0625, 2.125]	feasible
8	[2.0625, 2.0938]	infeasible
9	[2.0625, 2.0781]	infeasible
10	[2.0625, 2.0703]	infeasible
11	[2.0625, 2.0664]	infeasible
12	[2.0645, 2.0664]	feasible
13	[2.0654, 2.0664]	feasible

In the above procedure, the final largest value of the set size that makes the robust optimization problem feasible is 2.0654. This will be a valid upper bound for optimal set size identification.

Once the maximum feasible set size is identified, we can search for the optimal set size. The proposed algorithm is based on the branch and bound idea. The following basic principle is used to reduce the search region: whenever a solution with zero objective value (within tolerance ε) is found for problem (21), all the regions on the right hand side of the solution point can be removed; otherwise, the current search region will be branched at the solution point. If a search region width is smaller than tolerance σ, its solution is recorded and the corresponding region is skipped. The optimal set size identification algorithm is summarized in Table 5.

Table 5. Optimal set size identification algorithm

Algorithm: Optimal set size identification
1. Initialization $[lb, ub] = [lb_0, ub_0]$, set tolerance parameter ε, σ

2. Initial $X^{local} = \{\}, F^{local} = \{\}$

3. While interval set L is not empty

4. Take out the first interval L_1 from set $L : L = L - \{L_1\}$

5. Find the local minimum (x^*, f^*) of f in the first interval L_1

6. If $f^* \le \varepsilon$

7. $x^{global} = x^*$, $f^{global} = f^*$

8. Empty L

9. $L = L + \{[lb, (lb + x^*)/2]\}$, $L = L + \{[(lb + x^*)/2, x^*]\}$

10. Else

11. If $ub - lb > \sigma$

12. $L = L + \{[lb, (lb + ub)/2]\}$, $L = L + \{[(lb + ub)/2, ub]\}$

13. Else

14. $X^{local} = X^{local} + \{x^*\}, F^{local} = F^{local} + \{f^*\}$

15. End

16. End

17. Order the intervals in L by ascending value of lower/upper bounds

18. End

19. If x^{global} is empty

20. Find the minimum solution from the set X^{local}, F^{local} $x^{global} = \arg\min F^{local}$

21. End

22. Return x^{global}, f^{global}

The motivating example is used to compare the three methods for robust optimization approximation to chance constraint. First, the target probability of constraint satisfaction is set as 0.5 (i.e., $\alpha = 0.5$) in the chance constraint. By applying the traditional method with *a priori* probability bound (traditional), the iterative method with *a posteriori* probability bound and the method proposed in this work with the optimal set size to the numerical example, the following results are obtained as shown in Table 6.

Table 6. Comparing the different solutions for the motivating example (α=0.5)

	Traditional	Iterative method	Optimal approximation
Obj^*	-3.656×10^{-3}	-1.901×10^{-3}	3.469×10^{-3}
$p_{violation}^{true}$	0.0662	0.1183	0.2993
CPU time (s)	0.56	2.16	950

In Table 6, "Traditional" method corresponds to the *a priori* probability bound based method introduced in section 2.3.1, "Iterative method" corresponds to the method introduced in section 2.3.3, "Optimal approximation" corresponds to the proposed method in section 2.4.2. All the three methods lead to feasible solutions for the original chance constrained problem, and the result of the proposed optimal approximation method leads to the best objective value. Compared to the traditional approximation method, the robust solution has been improved from -3.656×10^{-3} to 3.469×10^{-3} with the proposed method. For the proposed optimal approximation method, the solution leads to a probability of violation 0.2993, which is closest to the target 0.5. Notice that there is a gap to the desired value 0.5 because the minimum possible probability of constraint satisfaction is 0.7, as shown in Figure 5.

Next, the reliability for chance constraint satisfaction is changed to 0.82 (i.e., $\alpha = 0.18$), and the different methods are applied again to this case. The corresponding results are shown in Table 7.

Table 7. Comparing the different solutions for the motivating example (α=0.18)

	Traditional	Iterative method	Optimal approximation
Obj^*	-7.994×10^{-3}	-5.052×10^{-3}	2.143×10^{-3}
$p_{violation}^{true}$	0.0139	0.033	0.1801
CPU time (s)	0.56	3.14	326

In this situation, the true probability of violation is 0.1801 for the proposed optimal approximation method, which is very close to the target value 0.18. Comparing the results of those three methods, it is seen that by identifying the optimal set size, the quality of the solution is improved while still ensures the desired degree of constraint satisfaction. The robust solution has been improved from -7.994×10^{-3} to 2.143×10^{-3}. As a trade-off, the computation time is increased, since a branch and bound method is used to find the global optimal set size.

2.5. Quantify the Quality of Robust Solution

In the previous section, algorithms for optimal robust optimization approximation have been introduced. While the solution of the approximation problem will be used as a candidate solution for the original chance constrained optimization problem, its quality should be quantified with a certain confidence interval. In this section, both the feasibility and the optimality of the obtained robust solution will be studied. Specifically, the feasibility is quantified using an upper bound on the constraint violation probability with predefined confidence level, and the optimality is quantified using upper bound (for a maximization problem) on the optimal objective with predefined confidence level.

2.5.1. Quantify the Feasibility

With a given solution to problem (1), a (1-δ)-confidence upper bound on $p_{violation}^{true}$ in equation (19) can be evaluated by the procedure shown in Table 8[29].

Table 8. Quantify feasibility of robust solution

Procedure: Feasibility bound evaluation

1. Set confidence level δ
2. Generate N samples
3. Evaluate the constraint for the N samples
4. Count the number of times that the constraint is violated V
5. Evaluate an upper bound on the constraint violation probability

   $$\hat{\alpha} = \max_{\gamma\in[0,1]}\left\{\gamma : B(V;\gamma,N) = \sum_{i=0}^{V}\binom{N}{i}\gamma^i(1-\gamma)^{N-i} \geq \delta\right\}$$

   $B(V;\gamma,N)$ is the cumulative distribution function of binomial distribution
6. Then with probability of at least $1-\delta$, the quantity $\hat{\alpha}$ is an upper bound for the true probability of constraint violation

The above method is applied to the motivating example and the following results are obtained as shown in Table 9.

Table 9. Results for the reliability quantify

α	$\tilde{p}_{violation}(x^*)$	$\hat{\alpha}$
0.5	0.2993	0.3026
0.18	0.1801	0.1819

In Table 9, α is the reliability parameter in the chance constraint, $\tilde{p}_{violation}(x^*)$ is the estimated probability of constraint violation, $\hat{\alpha}$ denotes the upper bound. It can be observed that when the probability of constraint violation is set as 0.5, the estimated probability of constraint violation is 0.2993, and the 90% confidence level upper bound on the probability of constraint violation is 0.3026. As the target probability of constraint violation is set as 0.18, the estimated probability is also close to the 90% upper bound 0.1819.

2.5.2. Quantify the Optimality

To quantify the optimality of a solution to the original chance constrained problem, an optimality upper bound can be evaluated from the following scenario optimization problem

$$\begin{aligned} &\max_x \; f(x) \\ &\text{s.t.} \;\; h(x, \xi^{(s)}) \le 0, \quad \forall s = 1, \dots, N \end{aligned} \tag{22}$$

The procedure for evaluating the optimality upper bound is reported in Luedtke & Ahmed[38], and summarized in Table 10.

Table 10. Optimality bound quantification

Procedure : Optimality bound evaluation
1. Set parameters N : sample size for the Scenario Optimization problem δ : confidence level 2. Determine parameter M using the following formula, this is the number of Scenario

Optimization problems to be solved

$$M \geq \ln\left(\frac{1}{\delta}\right)\frac{1-(1-\alpha)^N}{(1-\alpha)^N}$$

3. Solve M times the scenario optimization problem (22)
4. Obtain the optimal objective value obj_N^m, m=1,…,M

   For infeasible problem set $obj_N = -\infty$, for unbounded problem set $obj_N = \infty$
5. Find the maximum objective value $obj_N^{(\max)}$
6. Then with probability of at least $1-\delta$, the quantity $obj_N^{(\max)}$ is an upper bound for the true optimal value of the chance constrained optimization problem.

2.6. Case Study

To evaluate the proposed optimal robust optimization approximation method for solving chance constrained optimization problem, three case studies are investigated in this section. The first case is a portfolio optimization problem which enforces investment risk by chance constraint, the second case is a production planning problem, and the last case is a process scheduling problem.

2.6.1. Portfolio Optimization Problem

Consider the following portfolio optimization problem:

$$\begin{aligned}
&\max_{x\in X}\ \bar{r}^T x \\
&\text{s.t.}\quad Prob\left\{r^T x \geq v\right\} \geq 1-\alpha \\
&\qquad\ \ X = \left\{x \in R^n : e^T x = 1, x \geq 0\right\}
\end{aligned} \tag{23}$$

where x represents the percentage of a capital invested in each of the available assets, r denotes the vector of random returns of the assets, and $\bar{r}$ is the expected returns of the assets. Historical stock data from Yahoo Finance is obtained for the 100 assets in S&P100. We assume the data follow multivariate lognormal distribution and the distribution parameters are estimated using monthly stock price data from January 2003 to December 2013.

The proposed optimal robust optimization approximation algorithm is applied to solve the problem under different target reliability levels, and solution is summarized in Table 10. The allocation plan is shown in Figure 7-9. When α=0.01, the total investment goes to 11 stocks, AAPL, AIG, BAC, C, EBAY, EXC, F, FCX, MO, MS, and UNH. There are just three same stocks are invested for both α=0.07 and α=0.15, and the difference lies in the percentage for each stock. When α=0.07, among the total investment, 26 percent goes to the stock AIG, 51 percent goes to the stock BAC, and 23 percent goes to the stock C, while the corresponding allocation plan is 61 percent, 15 percent and 24 percent, respectively, for the case when α=0.15.

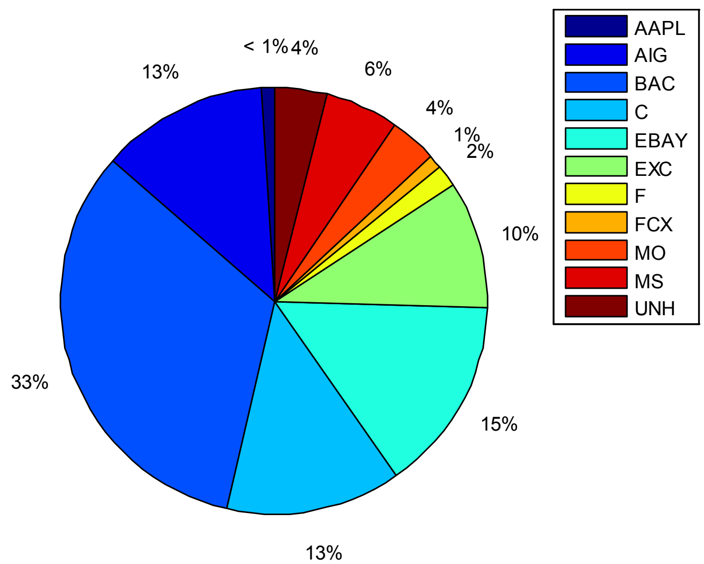


Figure 7. Allocation plan for portfolio problem when α=0.01

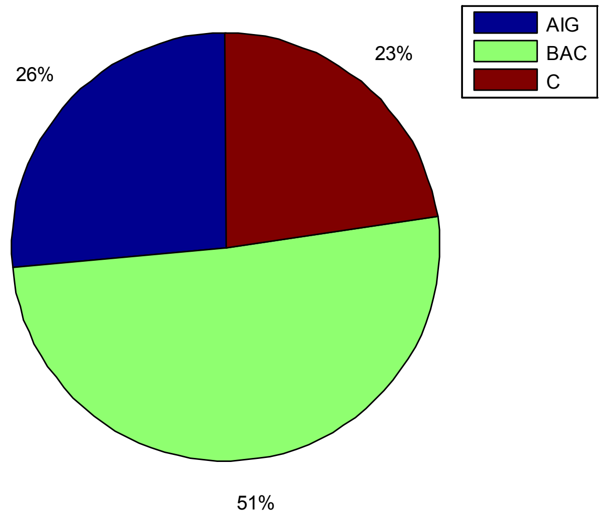


Figure 8. Allocation plan for portfolio problem when α=0.07

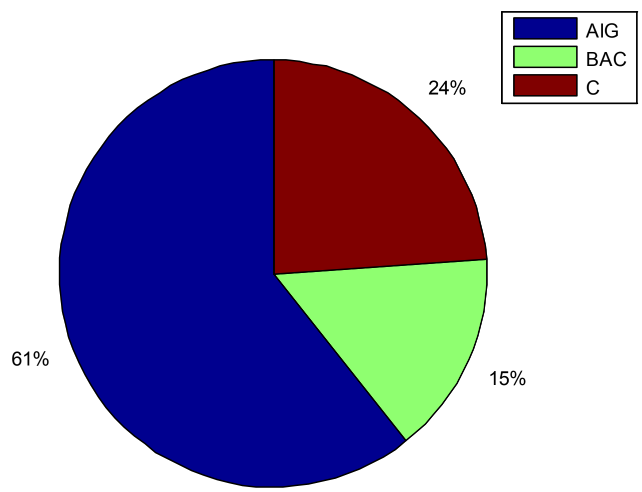


Figure 9. Allocation plan for portfolio problem when α=0.15

To evaluate the upper bound on optimal objective and upper bound on constraint violation probability, the confidence level is set as 0.9. For optimality upper bound evaluations, the sample size is chosen as 100 for scenario optimization.

Table 11. Results for the portfolio optimization problem

α	$\hat{\alpha}$	Δ^*	obj^*	obj^{UB}
0.01	0.011505	0.582340	0.021128	0.029177
0.02	0.022308	0.502012	0.023801	0.035687
0.03	0.032013	0.460838	0.025516	0.037676
0.04	0.043243	0.426199	0.027207	0.040642
0.05	0.052948	0.392795	0.029040	0.043930
0.06	0.061981	0.367536	0.030538	0.042449
0.07	0.071503	0.346112	0.031920	0.039054
0.08	0.082855	0.310249	0.034269	0.041876
0.09	0.092926	0.289499	0.035679	0.041164
0.10	0.101532	0.269111	0.037170	0.043759
0.11	0.113190	0.251315	0.038597	0.043293
0.12	0.123138	0.237170	0.039843	0.042398
0.13	0.134308	0.223249	0.041194	0.047350
0.14	0.143890	0.213045	0.042280	0.042609
0.15	0.154816	0.202955	0.043448	0.050325

In Table 11, $\hat{\alpha}$ is the constraint violation probability upper bound with 90% confidence level, Δ^* is the optimal uncertainty set size obtained from the proposed algorithm, obj^* is the objective value of robust optimization problem, and obj^{UB} is the upper bound of the chance constrained optimization problem's optimal objective with 90% confidence level.

From Table 11, it can be found that when the desired reliability level decrease (i.e., α increases), the optimal set size is smaller, and the objective value of the robust optimization problem becomes larger. For each value of α, the corresponding value of $\hat{\alpha}$ is very close to α, which means that the solution reliability is very close to the desired target (notice that $\hat{\alpha}$ is an upper bound on violation probability with 90% confidence level). Furthermore, the objective value of the robust optimization problem is not far from the optimality upper bound, either. The small gap between them means that the solution of optimal robust optimization approximation is close to the true solution of the chance constrained problem.

2.6.2. Production Planning Problem

A company needs to make a production plan for the coming year, divided into six periods, to maximize the sales with a given cost budget. The production cost includes the cost of raw material, labor, machine time, etc. The product manufactured during a period can be sold in the same period, or stored and sold later on. Operations begin in period 1 with an initial stock of 500 tons of the product in storage, and the company would like to end up with the same amount of the product in storage at the end of period 6. A linear optimization formulation of this problem can be formulated as below:

$$\max \sum_j P_j z_j \tag{24a}$$

$$\text{s.t.} \quad \Pr\left\{\sum_j \tilde{C}_j x_j + \sum_j V_j y_j \le 400,000\right\} \ge 1-\alpha \tag{24b}$$

$$500 + x_1 - (y_1 + z_1) = 0 \tag{24c}$$

$$y_{j-1} + x_j - (y_j + z_j) = 0 \qquad \forall j = 2,...,6 \tag{24d}$$

$$y_6 = 500 \tag{24e}$$

$$x_j \le U_j \qquad \forall j = 1,...,6 \tag{24f}$$

$$z_j \le D_j \qquad \forall j = 1,...,6 \tag{24g}$$

$$x_j, y_j, z_j \ge 0 \quad \forall j = 1,...,6 \tag{24h}$$

In the above model, decision variables x_j represent the production amount during period *j*, y_j represent the amount of product left in storage (tons) at the end of period *j* and z_j represent the amount of product sold to market during period *j*. The objective function (24a) maximizes the total sales. The first constraint (24b) is a chance constraint, which enforces that the total cost does not exceed a given budget with certain probability level α. Constraints (24c) and (24d) represent the inventory material balances. Constraint (24e) requires that the final inventory meet the desired level (i.e., 500 tons). Constraints (24f) and (24g) represent the production capacity limitations and demand upper bounds, respectively. Detailed data for the above LP problem are in Table 11, where j denotes period, P_j represents the selling Price ($/Ton), C_j is the production cost ($/Ton), V_j is the storage cost ($/Ton), U_j is the production capacity (Tons) and D_j is the demand (Tons). Assume the production costs $\tilde{C}_j$ are subject to uniform probability distribution and that there is a maximum of 50% perturbation around their nominal values as listed in Table 12

Table 12. Problem data for the production planning problem

j	P_j	C_j	V_j	U_j	D_j
1	180	20	2	1500	1100
2	180	25	2	2000	1500
3	250	30	2	2200	1800
4	270	40	2	3000	1600
5	300	50	2	2700	2300
6	320	60	2	2500	2500

For this chance constrained optimization problem, the proposed optimization robust approximation algorithm is applied under different reliability targets. The solution reliability and optimality is evaluated under 90% confidence level. The solution is reported in Table 13.

Table 13. Results for the production planning problem

α	$\hat{\alpha}$	Δ^*	obj^*	obj^{UB}
0.01	0.012848	1.238295	2501776.7	2514535.30
0.02	0.024689	1.112829	2532496.5	2568517.00
0.03	0.032501	1.051526	2548290.4	2660844.90
0.04	0.045807	0.971924	2569575.8	2586843.60
0.05	0.054352	0.931861	2580221.3	2599373.50
0.06	0.063995	0.889808	2591638.2	2632869.30
0.07	0.076630	0.840582	2605347.5	2627209.50
0.08	0.083344	0.816420	2612131.0	2651092.40
0.09	0.093719	0.776477	2623375.0	2661518.50

0.10	0.104950	0.747998	2631422.1	2656777.00
0.11	0.117767	0.712319	2641546.7	2661465.60
0.12	0.130524	0.676653	2651725.6	2724279.40
0.13	0.135712	0.663140	2655597.7	2684241.50
0.14	0.148773	0.631832	2664611.5	2752402.70
0.15	0.157501	0.610250	2670859.4	2687806.40

From Table 13, it can also be observed that the robust solution's reliability is close to the target, and the robust solution's objective values are close to the chance constrained problem's solution. For example, as the desired constraint satisfaction probability is 0.95 (i.e., α=0.05), the optimal set size identified by the proposed algorithm is 0.931861, which leads to a robust solution with objective 2580221.3. The 90% confidence level upper bound on constraint violation probability is 0.054352, which means the robust solution will satisfy the chance constraint with a probability larger than 0.945648 under 90% confidence. The absolute gap between the robust solution's objective and the true solution of chance constrained problem will be less than |2580221.3-2599373.5| =19152.2 under 90% confidence, which corresponds to a relative gap less than 0.74%.

The *a priori* probability bound based method (traditional), the *a posteriori* probability bound based method, the iterative method and the optimal robust optimization approximation method are applied to solve the chance constrained problem, and the corresponding results are summarized in Table 14.

Table 14. Comparison of different methods for the production planning problem

	A priori	*A posteriori*	Iterative	Optimal
obj^*	2350437	2569024	2563742	2667047
$P_{violation}^{true}$	0.0186	0.0672	0.0665	0.1479
time(s)	3.7	1000	1.8	266

From Table 14, it can be found that while all the solutions satisfy the probabilistic requirement, the solution of the optimal robust approximation is the best. It has the largest objective value and the largest probability of violation of the constraint, while still less than the target level 0.15. The solution obtained from the *a posteriori* probability based method is a little better than the one got from the iterative method. However, solve time of the iterative method is much less. By comparing the traditional method and the iterative method, it can be noticed that the quality of the solution has been greatly improved (i.e., from 2,350,437 to 2,563,724, a 9% increase). After comparing those four methods, it can be concluded that the iterative method and optimal approximation method are good options for solving the original chance constrained problem.

2.6.3. Process Scheduling Problem

In this section, a problem about the scheduling of a batch chemical process related to the production of one chemical product using three raw materials is involved[37]. The state-task-network representation of the example is shown in Figure 10.

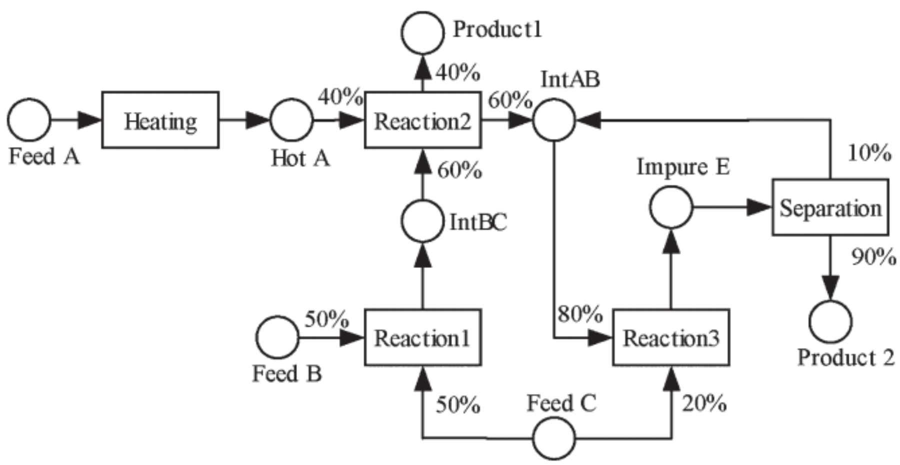


Figure 10. State Task Network representation of the batch chemical process

The mixed integer linear optimization model is formulated as follows.

$$\max \ profit$$

$$s.t. \quad profit = \sum_{s \in S_p, n} price_s d_{s,n} - \sum_{s \in S_r} price_s \left(STI_s - STF_s \right) \tag{25a}$$

$$\sum_{i \in I_j} wv_{i,j,n} \le 1 \ \forall i \in I \tag{25b}$$

$$st_{s,n} = st_{s,n-1} - d_{s,n} - \sum_{i \in I_s} \rho^C_{s,i} \sum_{j \in J_i} b_{i,j,n} + \sum_{i \in I_s} \rho^P_{s,i} \sum_{j \in J_i} b_{i,j,n-1}, \forall s \in S, n \in N \qquad (25c)$$

$$st_{s,n} \le st_s^{\max}, \forall s \in S, n \in N \qquad (25d)$$

$$v_{i,j}^{\min} wv_{i,j,n} \le b_{i,j,n} \le v_{i,j}^{\max} wv_{i,j,n}, \forall i \in I, j \in J_i, n \in N \qquad (25e)$$

$$\sum_n d_{s,n} \ge \tilde{r}_s, \forall s \in S_p \qquad (25f)$$

$$Tf_{i,j,n} \ge Ts_{i,j,n} + \alpha_{i,j} wv_{i,j,n} + \beta_{i,j} b_{i,j,n}, \forall i \in I, j \in J_i, n \in N \qquad (25g)$$

$$Ts_{i,j,n+1} \ge Tf_{i,j,n} - H\left(1 - wv_{i,j,n}\right), \forall i \in I, j \in J_i, n \in N \qquad (25h)$$

$$Ts_{i,j,n+1} \ge Tf_{i',j,n} - H\left(1 - wv_{i',j,n}\right), \forall i, i' \in I, j \in J_i, n \in N \qquad (25i)$$

$$Ts_{i,j,n+1} \ge Tf_{i',j',n} - H\left(1 - wv_{i',j',n}\right), \forall i, i' \in I, j, j' \in J_i, n \in N \qquad (25j)$$

$$Ts_{i,j,n+1} \ge Ts_{i,j,n}, \forall i \in I, j \in J_i, n \in N \qquad (25k)$$

$$Tf_{i,j,n+1} \ge Tf_{i,j,n}, \forall i \in I, j \in J_i, n \in N \qquad (25l)$$

$$Ts_{i,j,n} \le H, \forall i \in I, j \in J_i, n \in N \qquad (25m)$$

$$Tf_{i,j,n} \le H, \forall i \in I, j \in J_i, n \in N \qquad (25n)$$

The objective function (25a) maximizes the profit; (25b) state that only one of the tasks can be performed in each unit at an event point (n); (25c) represent the material balances for each state (s) expressing that at each event point (n) the amount is equal to that at event point (n-1), adjusted by any amounts produced and consumed between event points (n-1) and (n), and delivered to the market at event point (n); (25d) and (25e) express the storage and capacity limitations of production units; (25f) are written to satisfy the demands of final products; (25g) to (25n) represent time limitations due to task duration and sequence requirements in the same or different production units.

In this problem, the demand constraint $\sum_n d_{s,n} \geq \tilde{r}_s, s = P_1$ is affected for the uncertainty of demand P_1. There is a 20% perturbation of the nominal demand data ($r_{P_1} = 50$): $\tilde{r}_s = r_s + 0.2\xi_s r_s, s = P_1$ and a uniform distribution on the product is assumed, i.e., ξ_s is uniformly distributed in [-1,1]. Set the desired minimum probability of constraint satisfaction to 0.5. The corresponding chance constraint is

$$\Pr\left\{\sum_n d_{s,n} - \tilde{r}_s \geq 0\right\} \geq 0.5 \quad s = P_1 \tag{26}$$

Based on Li et al.[37], the box uncertainty set induced robust optimization constraint is used to approximate the chance constraint. The *a priori* probability bound based method (traditional), the *a posteriori* probability bound based method, the iterative method and the optimal robust optimization approximation method are applied to solve the chance constrained problem, and the corresponding results are summarized in Table 15. The corresponding robust schedules for those three methods are shown in Figure 11-14, respectively.

Table 15. Results of the process scheduling problem: case 1

	A priori	*A posteriori*	Iterative	Optimal
obj^*	1015.06	1071.20	1070.06	1088.75
$P_{violation}^{true}$	0	0.1927	0.1832	0.3979
time (s)	1.1	1000	5.2	2953

If the desired minimum constraint satisfaction is set to 0.9, case 2, which means the maximum constraint violation is set to 0.1, the obtained results are shown in Table 16 as follows:

Table 16. Results for process scheduling problem: case 2

	A priori	*A posteriori*	Iterative	Optimal
obj^*	Infeasible	1050.79	1025.55	1054.14
$P_{violation}^{true}$		0.0247	9.4e-14	0.1
time (s)		1000	3.4	1492

It can be seen from Table 15 and Table 16 that the objective value of the traditional method is the worst, the probability of the violation of its solution is also the smallest, and it is even infeasible for case 2. The objective value of the iterative method is smaller than the *a posteriori* probability bound based method, because its violation of constraint is smaller. However, the computational time of the iterative method is much less. The objective value of the optimal robust optimization method is the largest, and the violation of the constraint is smaller than the desired largest violation probability level and is the closest one to it. Notice that true probability of violation of the optimal solution is still no larger than the desired level. The reason for that case 1 has a true probability of violation less than 0.5 is that the minimum probability of satisfaction of the robust solution is 0.6021 (i.e., the upper bound of the probability of violation is 0.3979). The results demonstrated the trade-off between solution quality and computation time. While computational time is not a practical restriction, optimal robust approximation will be the best method since it leads to least conservative robust solution. Otherwise, the iterative method will be a good option since it leads to good quality solution with relatively small computation time.

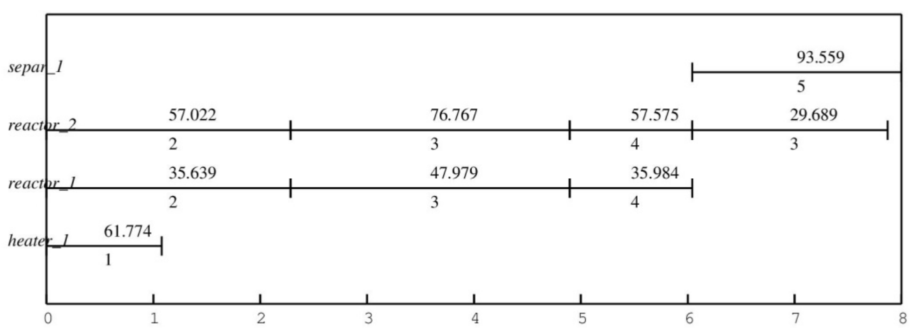


Figure 11. Schedule obtained from the traditional method

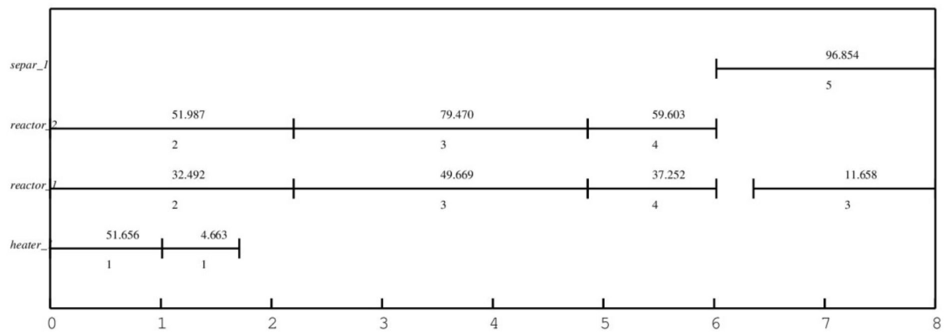


Figure 12. Schedule obtained from the a posteriori probability bound based method

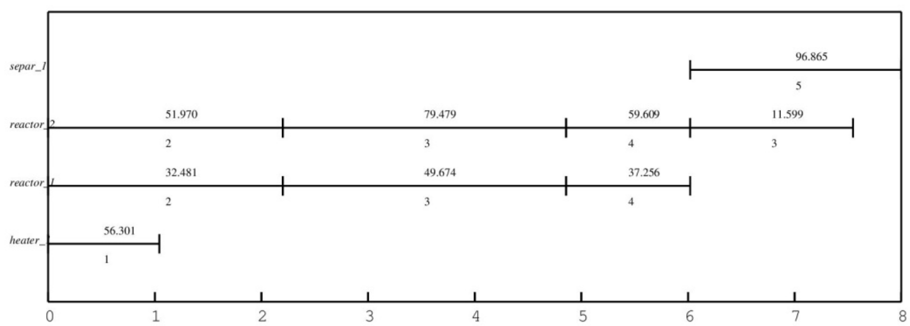


Figure 13. Schedule obtained from the iterative method

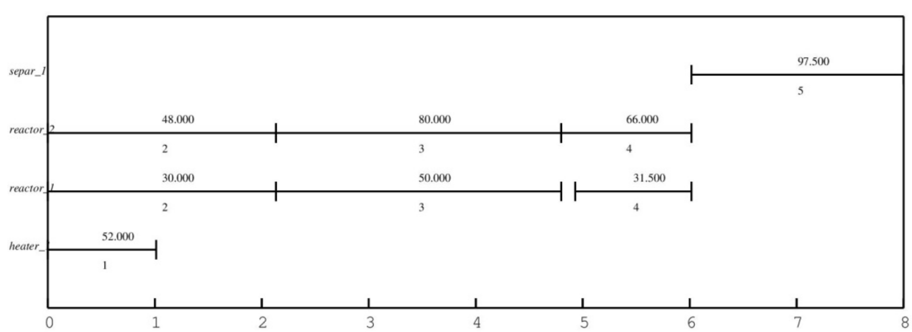


Figure 14. Schedule obtained from the optimal robust optimization approximation

2.7. Summary

In the traditional robust optimization framework, the uncertainty set's size is determined using the a priori probability bound[44] under the assumption that the uncertainty is subject to certain distribution (e.g., bounded and symmetric). However, the solution can be very conservative. The iterative method in Li & Floudas[45] combining the a priori probability bound and a posteriori probability bound can improve the solution quality, and extend the applicability to general distributions. In this section, a novel optimal robust optimization approximation method is proposed to identify the best possible robust solution that solves the chance constrained problem. The proposed method in this section does not rely on the a priori probability bound, so the uncertainty distribution is not limited. By finding the optimal uncertainty set size, the optimality of the robust solution is greatly improved while the desired reliability level is still satisfied. The effectiveness of the proposed method is demonstrated by a numerical example, as well as applications in the portfolio optimization problem, the production planning problem, and the process scheduling problem. Finally, it is worth mentioning that only individual chance constraint is investigated in this section. Joint chance constraint and its solution method based on robust optimization approximation will be studied in future work.

3. Linear Programming Based Scenario Reduction

3.1. Motivation

The solution of the sampling approach based approximation method is generated by considering all the selected scenarios. There is a direct correlation between the size of the approximation problem and the number of scenarios. In general, samples with large size can guarantee the reliability of the solution, but lead to high computational complexity. An appropriate choice of samples can balance the reliability of the solution and the computational complexity. There are several methods for scenario generation are proposed, e.g., Luedtke and Ahmed[38], Ahmed and Shapiro[40], and Wang and Ahmed[46]. However, in reality, it is very often that a huge number of scenarios need to be considered due to the large number of uncertainty parameters, which may result in numerically intractable problems because of the limitation of the computational recourses. Therefore, demand for finding subset of scenarios which can best approximate the original large number of scenarios is generated. Scenario reduction, another important topic of scenario based decision making, is induced. It aims at selecting a few representative scenarios among the original large number of scenarios, and new probabilities will be assigned to each representative scenario.

Scenred2[47, 48] is a tool for the reduction of scenarios modeling the random data processes, which is available in GAMS. However, the reduced scenarios are not the optimal ones. Li and Floudas[41] proposed a mixed integer linear problem (MILP) based scenario reduction method, which can give the optimal reduced subset to represent the initial super set, so as to reduce the computational complexity for solving the sampling based approximation problems while still guarantee the solution reliability. This method considers the performance of both input and output, while making a decision.

Although the MILP based scenario method can provide the optimal reduced subset of scenarios, its computational complexity can be very high when the size of the initial set of scenarios is very large (e.g. 5000). This may be unacceptable for solving a sampling based approximation problem. So a question is naturally raised: Is there a new scenario reduction method to balance the performance of the selected subset of scenarios and the computational complexity? In this thesis, a linear programming (LP) based scenario reduction method is proposed.

3.2. Mixed Integer Linear Programming Based Scenario Reduction

The mixed integer linear programming based scenario reduction method was proposed by Li & Floudas[41] recently. This method minimizes not only the probabilistic distance between the original and reduced input scenario distribution, but also minimizes the difference between the best, worst and expected performances of the output measure of the original and the reduced scenario distributions. It leads to reduced distribution not only closer to the original distribution in terms of the transportation distance, but also captures the performance of the output. However, in order to do the comparison with the scenario reduction method proposed in this thesis, a similar mixed integer programming based scenario reduction is considered here. It is formulated as follows[41]:

$$\min_{y_i, v_{i,i'}, d_i, p_{i'}^{new}} \sum_{i \in I} p_i^{orig} d_i \tag{27a}$$

$$s.t. \quad \sum_{i \in I} y_i = N \tag{27b}$$

$$\sum_{i' \in I} v_{i,i'} \ge y_i, \forall i \in I \tag{27c}$$

$$0 \le v_{i,i'} \le 1 - y_{i'}, \forall i, i' \in I \tag{27d}$$

$$\sum_{i' \in I} v_{i,i'} \le 1, \forall i \in I \tag{27e}$$

$$d_i = \sum_{i' \in I} c_{i,i'} v_{i,i'}, \forall i \in I \tag{27f}$$

$$0 \le d_i \le y_i c_{\max}, \forall i \in I \tag{27g}$$

$$p_{i'}^{new} = \left(1 - y_{i'}\right) p_{i'}^{orig} + \sum_{i \in I} v_{i,i'} p_i^{orig}, \forall i' \in I \tag{27h}$$

$$y_i \in \{0,1\}, \forall i \in I \tag{27i}$$

where i and i' represent scenarios, I is the set of all scenarios, p_i^{orig} represent probability of scenario i in original discrete distribution, d_i represent the minimum distance from all the selected scenarios to a removed scenario i, binary variables y_i denote whether a scenario is removed $(y_i = 1)$ or not $(y_i = 0)$, continuous variables $v_{i,i'}$ denote whether a scenario i is removed and assigned to scenario i' ($v_{i,i'} > 0$) or not ($v_{i,i'} = 0$), continuous variables $p_{i'}^{new}$ denote the new probabilities of the scenarios ($p_{i'}^{new} = 0$ means scenario i' is removed), $c_{i,i'}$ is the distance between two scenarios which can be evaluated by $c_{i,i'} = \sum_{t=1}^{T} \left| \theta_t^i - \theta_t^{i'} \right|$, θ^i is a realization of uncertain parameters in scenario i, $\theta^i = \{\theta_1^i, \theta_2^i, \ldots, \theta_T^i\}$, $c_{\max}$ is the maximum distance between two scenarios. The objective function (27a) minimizes the total transportation cost; (27b) requires the total number of the binary variables equal to the size of the original scenario set; (27c) to (27e) are written to determine the transportation plan; (27f) and (27g) express the distance between removed scenario i and preserved scenario i'; (27h) state the way to calculate the new probability of a selected scenario.

The optimal objective value of the mixed integer problem above is obtained when the new probability of the selected scenario is equal to the sum of its former probability and of all probabilities of removed scenarios that are closest to it[47] (i.e., $p_{i'}^{new} = (1 - y_{i'}) p_{i'}^{orig} + \sum_{i \in I} v_{i,i'} p_i^{orig}, \forall i' \in I$), as illustrated in Figure 15[41].

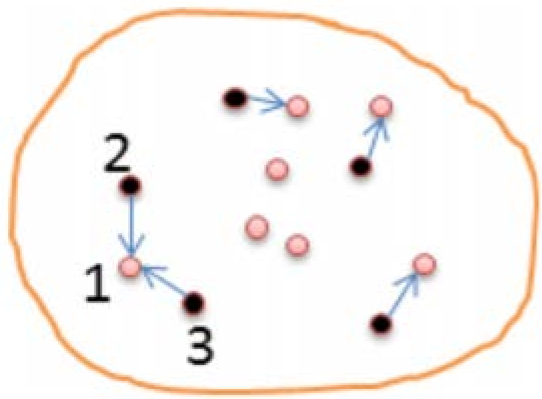


Figure 15. Illustration of computing new probability in reduced distribution

As shown in Figure 15, black dots represent the preserved scenarios and the red dots represent the preserved scenarios. Then, scenario 2 and scenario 3 will be assigned to scenario 1 since the scenario 1 is closest to them compared to the other preserved scenarios. Thus, the new probability of scenario 1 will be equal to the sum of its original probability and the probabilities of scenarios 2 and 3.

3.3. Linear Programming Based Scenario Reduction: Fixed Subset Size

To quantify the difference between the original superset of scenarios and the selected subset, the criteria can be classified based on the input parameter space and the system's output space. Transportation distance is a mathematical measurement for quantifying the cost of the movement from one location to another. Kantorovich distance is one type of them, which was studied by Kantorovich[49]. In a transportation problem, we intend to minimize the transportation cost (i.e., the Kantorovich distance). As for the scenario reduction problem, we can take it as a transportation problem, from the super sample set I to the preserved sample set S. The Kantorovich distance between I and S can be used to quantify the difference between those two sets. If there is less difference between I and S, the corresponding Kantorovich distance will be smaller.

Data and distribution are two distinctive features of a sample set. If the sub sample set is given, then, by calculating the Kantorovich distance between I and S, the new distribution of the sub sample set can be determined. Kantorovich distance between two sets can be calculated by solving the following linear minimization problem:

$$\begin{aligned}
&\min_{\eta_{i,i'}} \sum_{i\in I}\sum_{i'\in S}\eta_{i,i'}c_{i,i'} \\
&s.t.\ \sum_{i\in I}\eta_{i,i'} = p_{i'}^{new}, \forall i' \in S \\
&\qquad \sum_{i'\in S}\eta_{i,i'} = p_i^{orig}, \forall i \in I \\
&\qquad \eta_{i,i'} \geq 0, \forall i \in I, \forall i' \in S
\end{aligned} \tag{28}$$

where i represent scenarios in the super sample set *I*; i' represent scenarios in the sub sample set *S*; p_i^{orig} represent the probability of scenario i in the original distribution; $p_{i'}^{new}$ represent the probability of scenario i' in the new distribution; $\eta_{i,i'}$ is the decision variables represent the transportation plan; $c_{i,i'}$ is the distance between two scenarios, which can be evaluated by $c_{i,i'} = \sum_{t=1}^{T}\left|\theta_t^i - \theta_t^{i'}\right|$; θ^i is a realization of uncertain parameters in scenario i , $\theta^i = \left\{\theta_1^i, \theta_2^i, \ldots, \theta_T^i\right\}$.

According to Dupačová et al.[47], the optimal solution for (28) is obtained at:

$$p_{i'}^{new} = p_{i'}^{orig} + \sum_{i\in I(i')} p_i^{orig}, \forall i' \in S \tag{29}$$

The reduced scenarios will be assigned to the closest preserved scenario[41]. Each preserved scenario with the reduced scenarios which are assigned to it can be considered as a set. The central sample of this set is the preserved scenario. Then, the super sample set *I* is divided into several subsets. In order to reduce the Kantorovich distance, it can be achieved by reducing the cost of each subset. As shown in Figure 16, the black dots are reduced scenarios and the red dots are preserved scenarios.

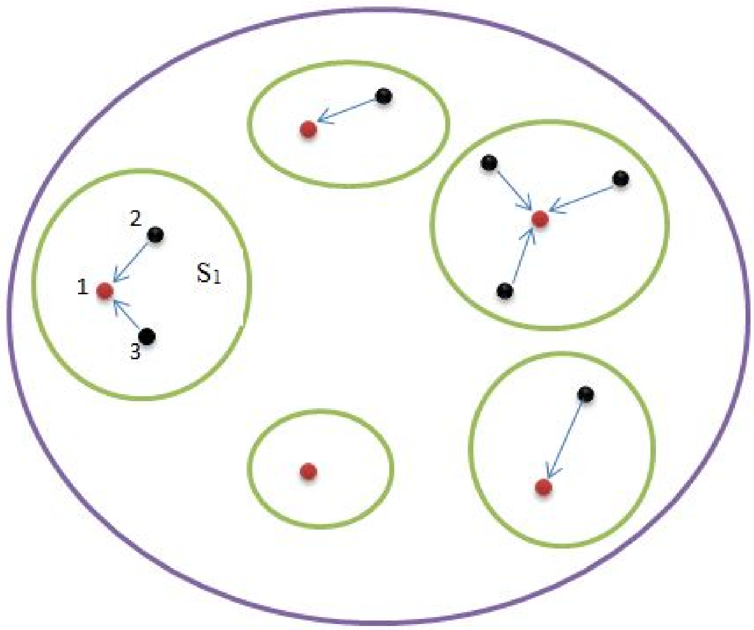


Figure 16. Illustration of the algorithm

The subset S_1 has three scenarios, scenario 1, 2 and 3. Scenario 2 and scenario 3 will be assigned to scenario 1. The central scenario is scenario 1. Next, consider scenario 2 and scenario 3 as the central scenario, respectively, and calculate the corresponding cost under the assumption that the other scenarios in the same subset are assigned to it. The total cost for one central scenario is the summation of the cost of transportation from the rest scenarios in the same subset to it. For example, the total cost for scenario 1 is the summation of the cost of transportation from scenario 2 to it and the cost corresponds to scenario 3. Update the central sample of S_1 to the one with the minimum cost among scenario 1, 2 and 3, so as to the rest subsets. The cost between two scenarios can be calculated by $cost_{i,i'} = p_i^{orig} c_{i,i'}$, where scenario i' is the preserved scenario (i.e., the central scenario). All the updated central samples make up the new preserved scenario set S, which is more representative. The algorithm is summarized in Table 17.

Table 17. Linear programming based scenario reduction algorithm (fixed subset size)

Algorithm 1: Linear programming based scenario reduction (fixed subset size)
1. Initialization the number of samples for the preserved scenario set, K
2. Initialization S with K samples, $\{S_1, S_2, \ldots, S_K\}$, from the super scenario

set I

3. Initialization tolerance ε
4. While relative step error > ε
5. Solve problem (28) with I and S
6. Obtain the transportation plan and Kantorovich distance
7. Find the K subsets
8. For each subset, calculate the cost for each scenario
9. For each subset, find the scenario with minimum cost as the new central scenario
10. Update the K central scenarios
11. Update S with the new K central scenarios, $\{S_1', S_2', \ldots, S_K'\}$
12. Calculate the relative step error
13. End
14. Return S as the final preserved scenario set

After the transportation plan is obtained, there will be a vector $\eta_{i,i'}, \forall i \in I$ for each i' (i.e. subset). In step n ($n \geq 2$), the relative step error is calculated by:

$$nth\ relative\ error = \left| \frac{KanDist(n) - KanDist(n-1)}{KanDist(n-1)} \right| \tag{30}$$

where $KanDist(n)$ represents the Kantorovich distance between super scenario set I and the preserved scenario set S obtained in the *n-th* step.

Property 1: The Kantorovich distance obtained by Algorithm 1 in each step is monotonically decreasing.

Proof: In each step, the central scenarios are updated to the one with the minimum cost. The preserved scenario set is formed by the updated central scenarios. So, the Kantorovich distance will be no larger than the former step.

3.4. Illustration of the Algorithm

In this section, small examples will be used to illustrate the LP based scenario reduction algorithm proposed in Table 17. The comparison between the proposed method and the MILP based scenario reduction method will be conducted. All of the calculations in Section 3 have been made on a Windows 8 system with an Intel Core i7 (2.40 GHz) and 16 GB of RAM, using MATLAB and GAMS with CPLEX optimizer.

3.4.1. Normal Distribution

Consider a normal distribution with mean 0 and variance 1. Generate a super set *I* with 100 scenarios, each scenario has two elements (i.e., *I* is a 100×2 matrix), and set the preserved scenario set *S* to have a size of 5 (i.e., *K*=5). Initialize *S* with 5 scenarios selected from *I*. By applying Algorithm 1 and set the tolerance $\varepsilon = 0.01$, the updated *S* and the corresponding distribution can be obtained.

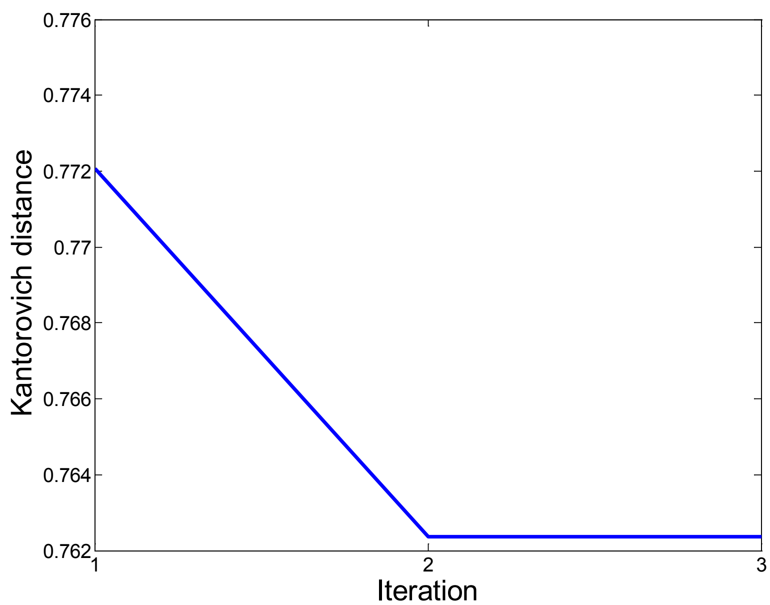


Figure 17. Kantorovich distance between *I* and *S* under normal distribution

The iteration stops at step 3, as shown in Figure 17, which means the Kantorovich distance will not significantly decrease and it is the same in this case. There is no need to continue updating *S*.

Plot all the scenarios in each step, as shown in Figure 18 and 19. The scenarios of the same color belong to the same subset. They are used to update the central scenario (i.e., the preserved scenario). The area of each central scenario represents the probability of this scenario in the distribution of *S*. The final Kantorovich distance is 0.7624. There are two central scenarios updated. For example, scenario 1 is the central scenario for the subset with color of green in Figure 18. In Figure 19, the central scenario is updated to scenario 2.

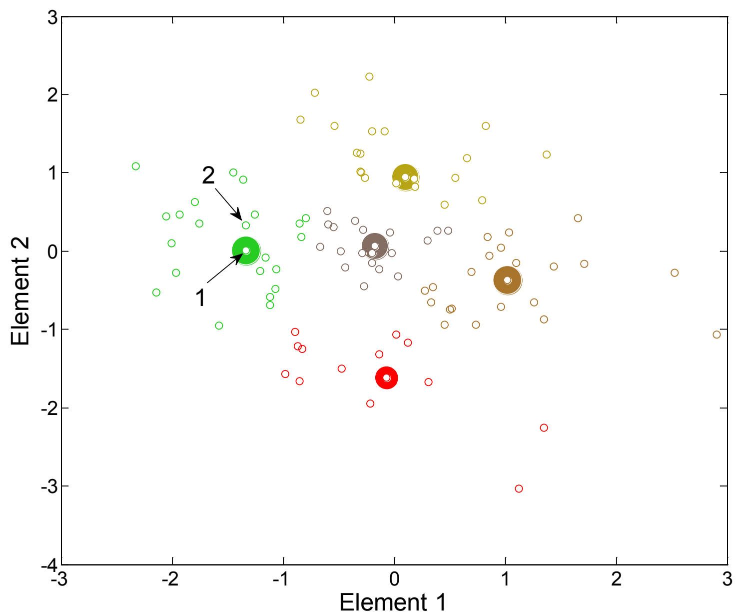


Figure 18. The first step of the procedure under normal distribution

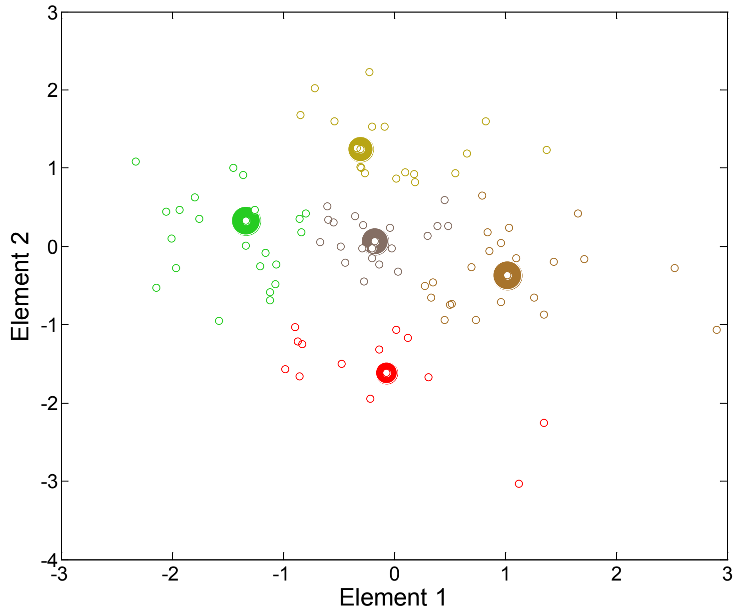


Figure 19. The second step of the procedure under normal distribution

For the same super set *I*, apply the MILP based scenario reduction method. The obtained final Kantorovich distance is 0.7560, and the final distribution of the preserved scenario set is shown in Figure 20.

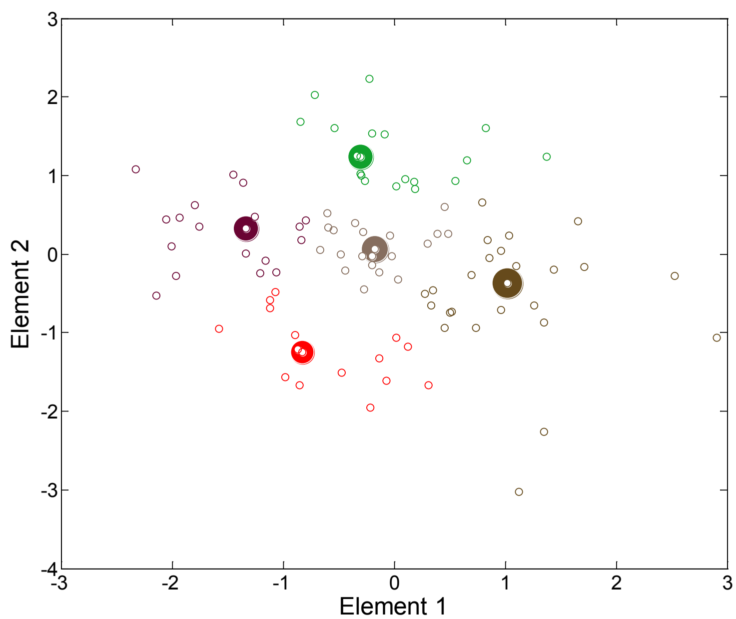


Figure 20. Final distribution obtained by MILP method under normal distribution

By comparing the final distribution of the preserved scenario set (i.e., Figure 19 and Figure 20), it can be found that for the five preserved scenarios, only one scenario of those obtained by the LP based scenario reduction method is different from those selected by the MILP based scenario reduction method. The difference comes from different transportation plans determined by MILP based scenario reduction and LP based scenario reduction. The absolute gap between the Kantorovich distances obtained by those two methods is $|0.7624 - 0.7560| = 0.0064$, which corresponds to a relative gap of 0.85%.

3.4.2. Uniform Distribution

Generate 100 scenarios from a uniform distribution in $[-1,1]$ as the super scenario set I (a 100×2 matrix). Select 5 scenarios to form the initial preserved scenario set S. The tolerance is set to $\varepsilon = 0.01$. By applying Algorithm 1, the Kantorovich distance can be obtained, as shown in Figure 21. The procedure is shown in Figure 22 and 23. The preserved scenario set S is updated three times until there is no big difference between the Kantorovich distances, same in this case, and the final Kantorovich distance is 0.4044. As for the results obtained by the MILP scenario reduction method, its final Kantorovich distance is 0.4042, and its final distribution of the preserved scenario set is shown in Figure 24. There is only one scenario that is not the same in the final preserved scenario sets obtained by those two methods, due to the difference between the determined transportation plans. The absolute gap between two Kantorovich distance is $|0.4044 - 0.4042| = 0.0002$, which leads to a relative gap of 0.05%.

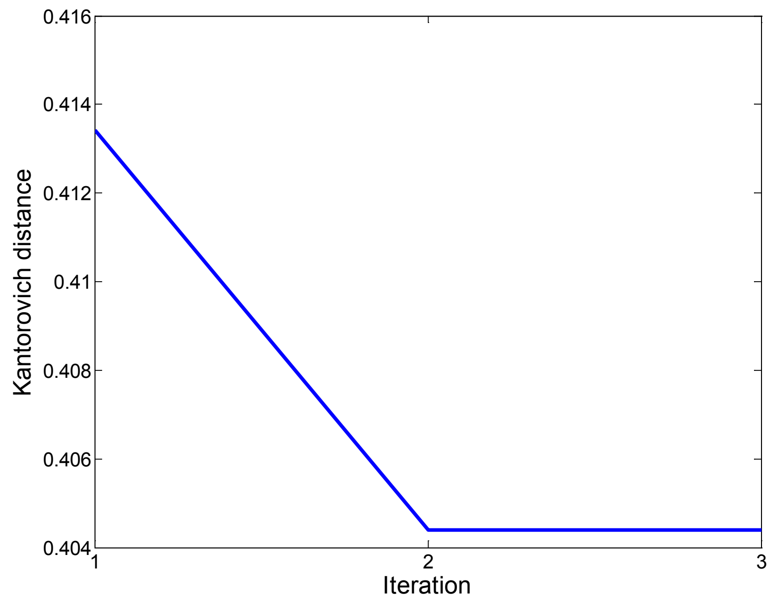


Figure 21. Kantorovich distance between *I* and *S* under uniform distribution

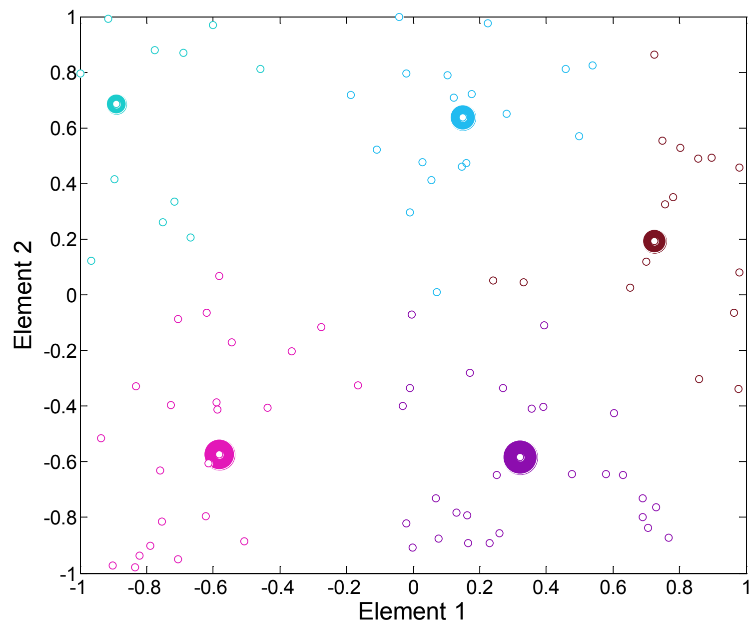


Figure 22. The first step of the procedure under uniform distribution

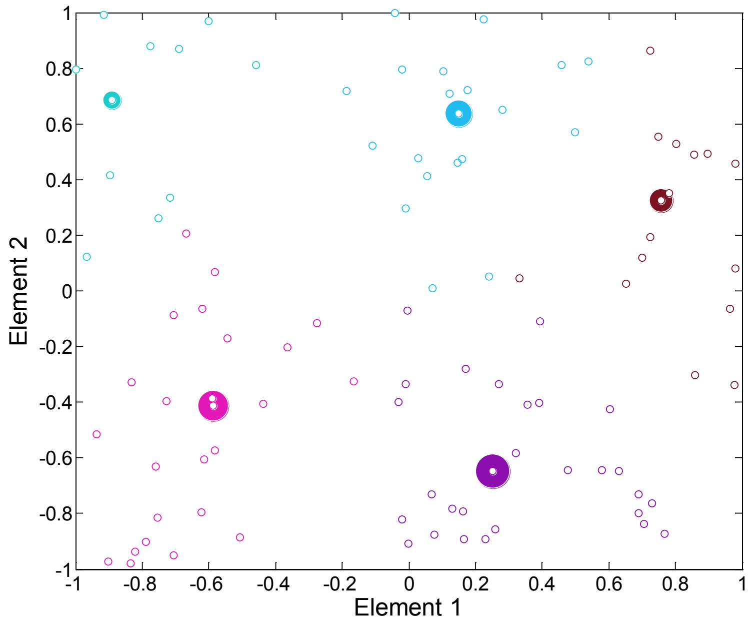


Figure 23. The second step of the procedure under uniform distribution

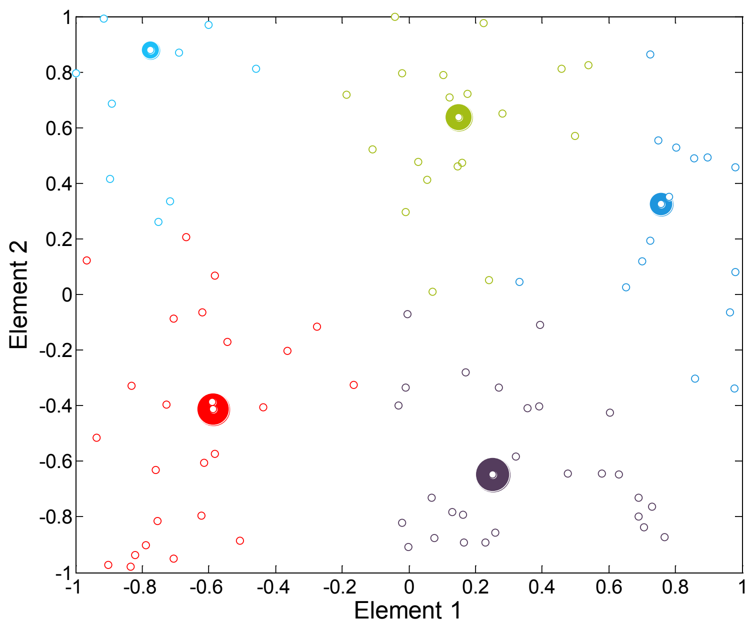


Figure 24. Final distribution obtained by MILP method under uniform distribution

3.5. Selection of the Initial Preserved Scenario Set

In order to find a more representative sub scenario set, the proposed LP based scenario reduction needs to be provided an initial scenario set with a size of the desired number of scenarios. There are several methods to conduct the selection. The effectiveness of different choice will not be the same.

K-means clustering is a method of vector quantization. It originally comes from signal processing, which is popular for cluster analysis. It aims to partition the total *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster[50]. The initial preserved scenario set can be the combination of the most representative scenario in each cluster, and the most representative scenario can be obtained by select the one closest to the center of the corresponding cluster.

To demonstrate the effectiveness of different initial selection, a two dimensional super scenario set with 1000 scenarios is generated. The desired number of selected scenarios is set to 50 (i.e., K=50 in Algorithm 1). Three initial sub scenario sets are generated using k-means clustering, and the other two sets are randomly selected from the original super scenario set. The tolerance is set to ε=0.01. The relationship between Kantorovich distance and different initial selections is plotted in Figure 25 and it can be observed that k-means clustering is a better method to provide the initial preserved scenario set. A better initial selection will result in a smaller final Kantorovich distance (i.e., a more representative final sub scenario set), and less effort in the computation.

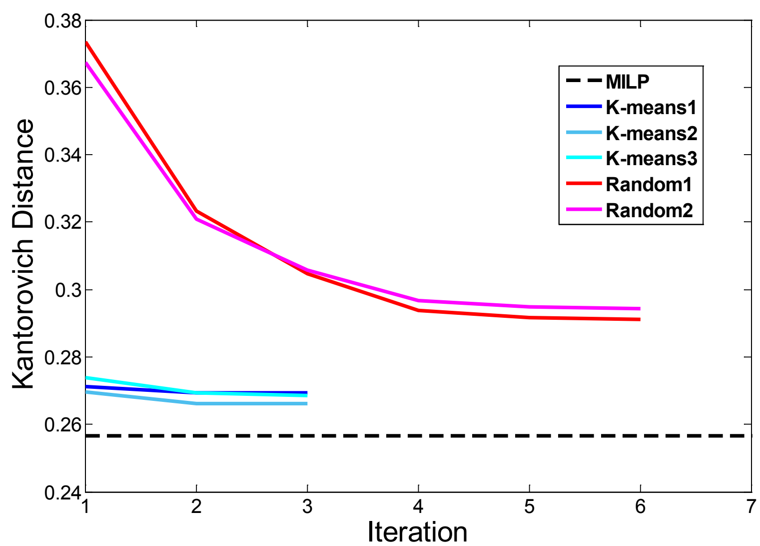


Figure 25. Kantorovich distance for different initial selection of sub scenario set

3.6. LP based scenario reduction V.S. Scenred2

In this section, the proposed method and the GAMS scenario reduction routine Scenred2 are both applied to perform a comparison study.

In order to do the comparison, a 2-dimensional super set of scenarios of size 1000, 5000, 10000, 15000, 20000, 30000 and 40000 is generated, respectively, under the normal distribution with mean 0 and variance 1. The desired number of scenarios to be preserved is set to 50, i.e., reduce the super set of scenarios to a sub set with 50 scenarios. For the proposed method, the tolerance is set to 0.01. The results obtained by applying those two methods are shown in Table 18.

Table 18. Results of comparison between Scenred2 and Algorithm 1

	LP Fixed		Scenred2	
	Kantorovich Distance	CPU time(s)	Kantorovich Distance	CPU time(s)
1000	0.2736	6.6	0.2790	1.3
5000	0.2888	13	0.2955	21
10000	0.2963	27	0.3026	87
15000	0.3000	42	0.3045	214
20000	0.2986	71	0.3070	443
30000	0.3025	163	0.3087	1735
40000	0.2993	313	0.3067	3664

From the results, it can be observed that the transportation distance (i.e., Kantorovich distance) of the proposed algorithm 1, LP based scenario reduction, is consistently smaller than that of the Scenred2 tool. Along with the size of the original scenario set getting larger, the CPU time for the reduction of the proposed method is much smaller than that of the Scenred2 tool. For the proposed method, it can provide a more representative sub scenario set while the computational complexity is lower, compared with the Scenred2 tool. The MILP based scenario reduction method introduced in section 3.2 is not applicable here, because it cannot conduct scenario reduction when the original number of scenario is very large. For example, it will be intractable for the case of 40000 scenarios.

In order to see the difference between the final transportation plans obtained by both methods, taking the case 40000 as an example, the obtained results are shown in Figure 26 and 28. Scenarios of the preserved scenario set are shown in Figure 27 and Figure 29. As it can be observed, there is no big difference, which can be also seen from Table 18. The Kantorovich distance just has a relative error of $\left|\frac{0.3067-0.2993}{0.2993}\right| = 2.5\%$ between two methods.

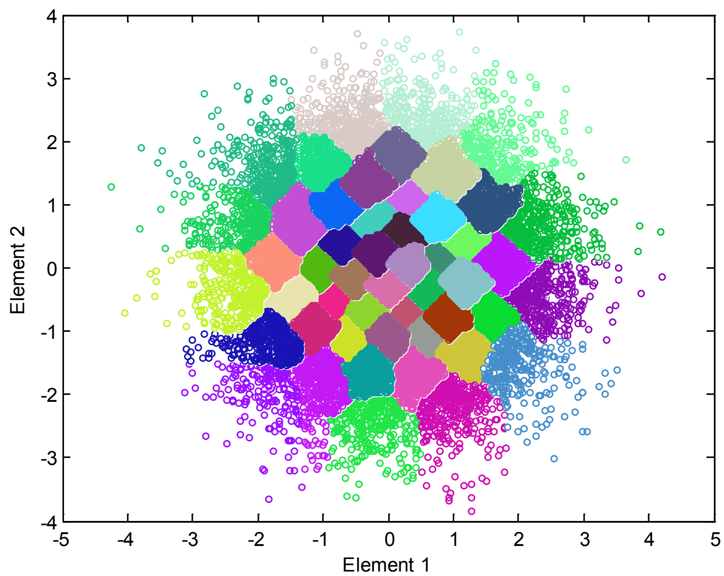


Figure 26. Final transportation plan for Algorithm 1

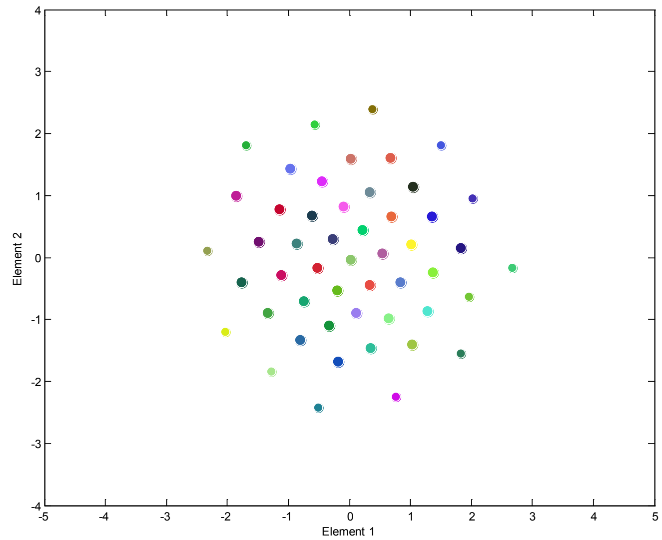


Figure 27. Final preserved scenarios for Algorithm 1

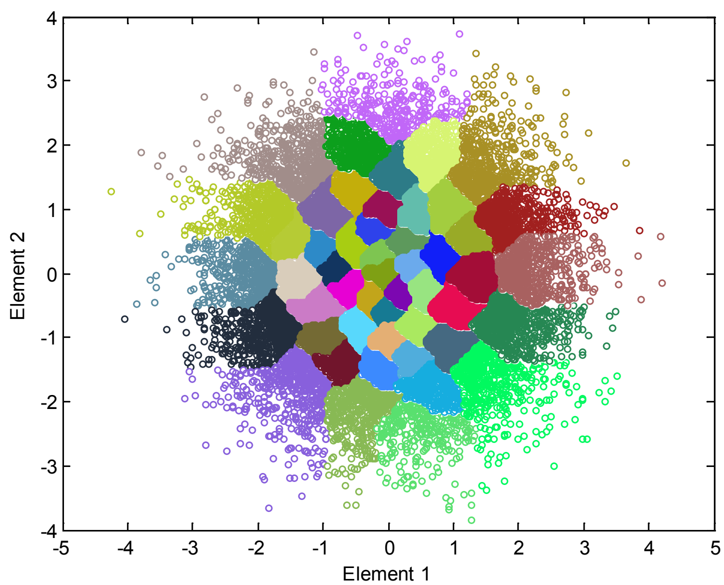


Figure 28. Final transportation plan for Scenred2

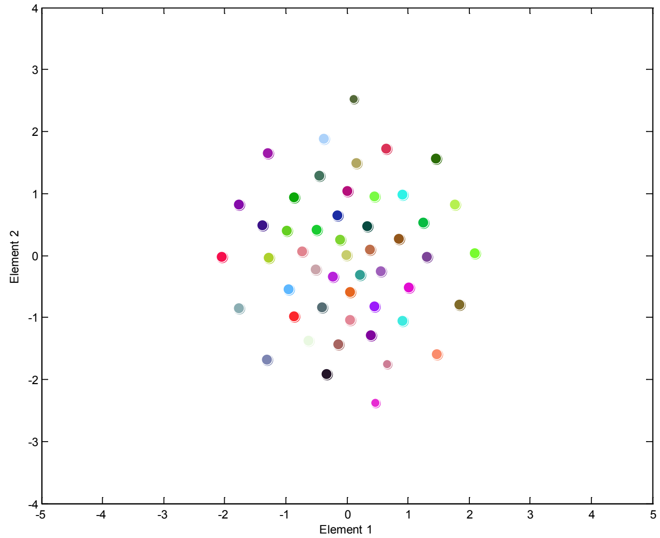


Figure 29. Final preserved scenarios for Scenred2

If change the way of calculating the distance between two scenarios to squared Euclidean distance: $c_{i,i'} = \sum_{t=1}^{T} \left(\theta_t^i - \theta_t^{i'}\right)^2$, the final transportation plans for both methods are shown in Figure 30 and 32, and the final preserved scenarios are shown in Figure 31 and 33.

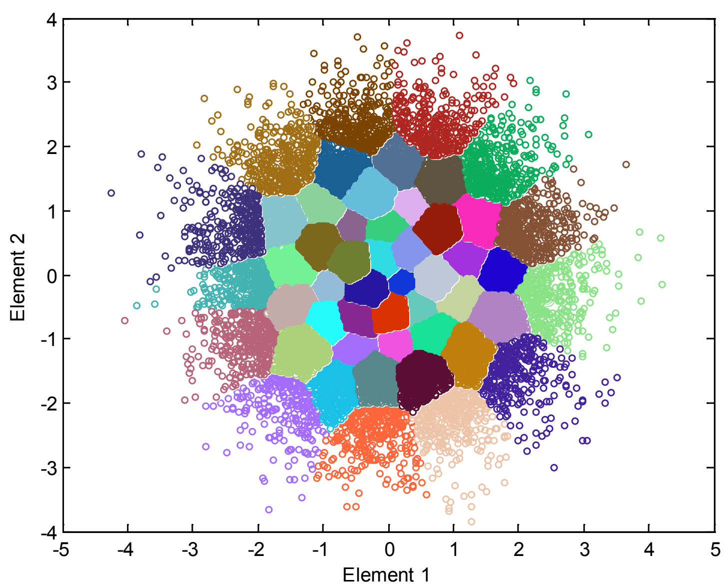


Figure 30. Final transportation plan for Algorithm 1 with squared Euclidean distance

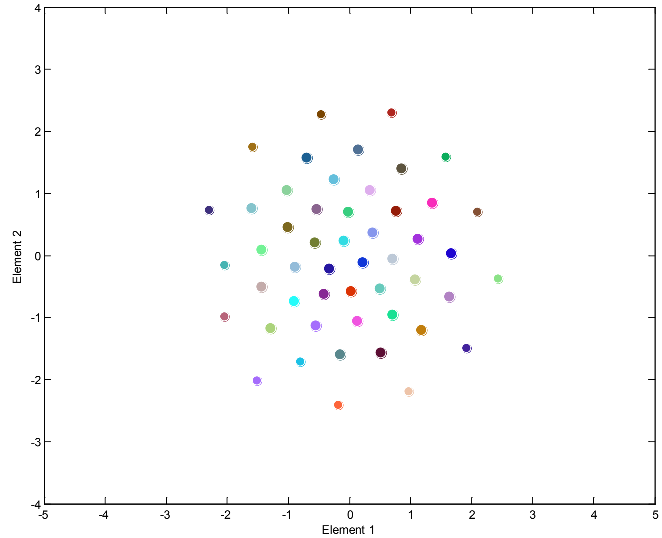


Figure 31. Final preserved scenarios for Algorithm with squared Euclidean distance

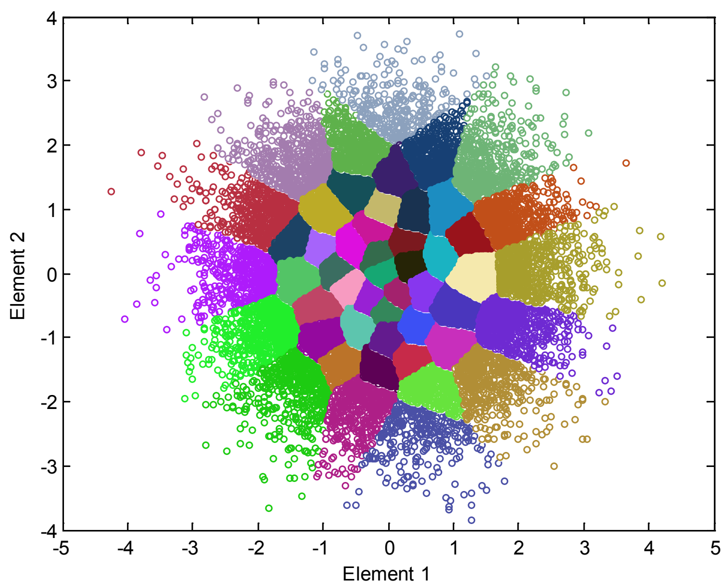


Figure 32. Final transportation plan for Scenred2 with squared Euclidean distance

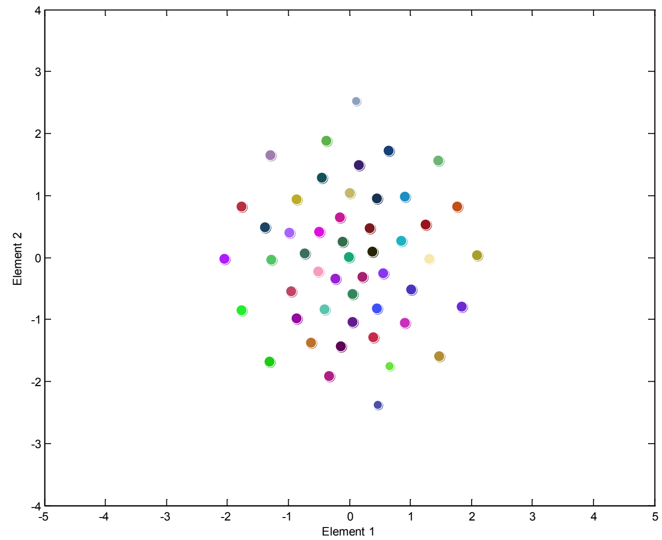


Figure 33. Final preserved scenarios for Scenred2 with squared Euclidean distance

By comparing the results obtained by both two methods with different distance calculation, it can be found that the final transportation plan and the final preserved scenarios are not the same for each method.

If still use the Manhattan distance and generate the 2-dimensional super scenario set based on the uniform distribution whose support is [-1, 1], the final transportation plan and the final preserved scenarios obtained by applying Algorithm 1 are shown in Figure 34 and 35.

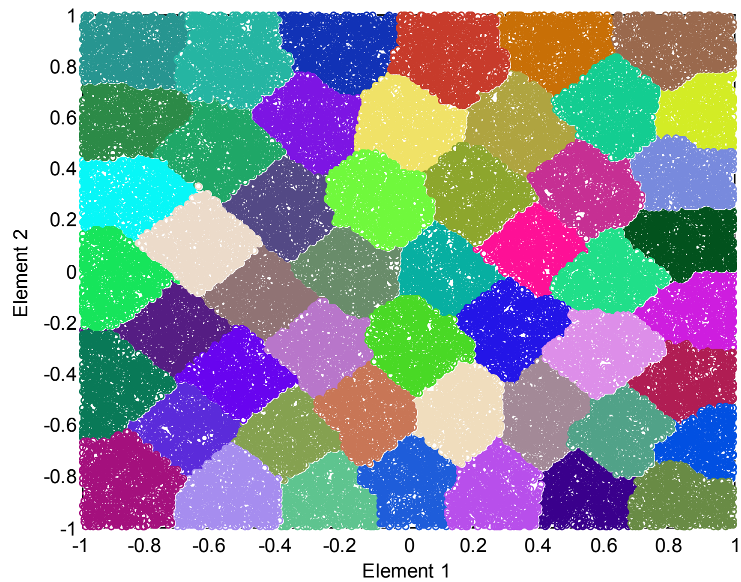


Figure 34. Final transportation plan for Algorithm 1 under uniform distribution

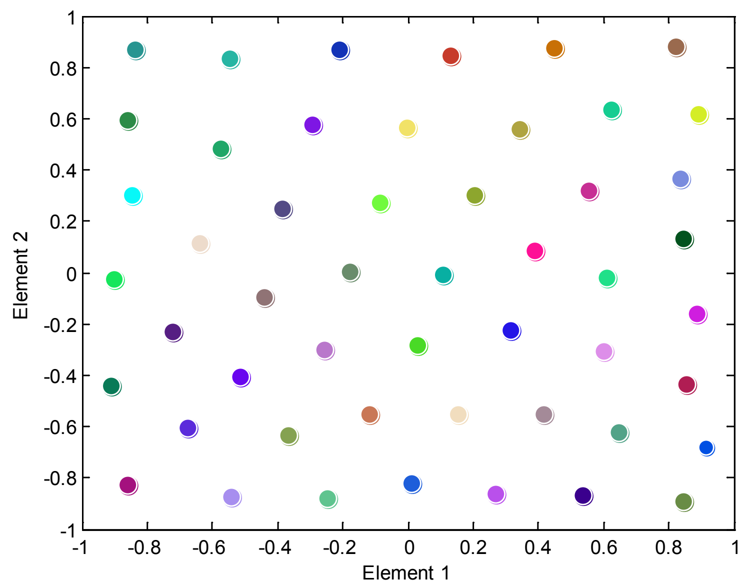


Figure 35. Final preserved scenarios for Algorithm 1 under uniform distribution

If the final preserved scenarios are taken as the seeds for the Voronoi diagram, then the corresponding Voronoi cells can be generated. The result is shown in Figure 36.

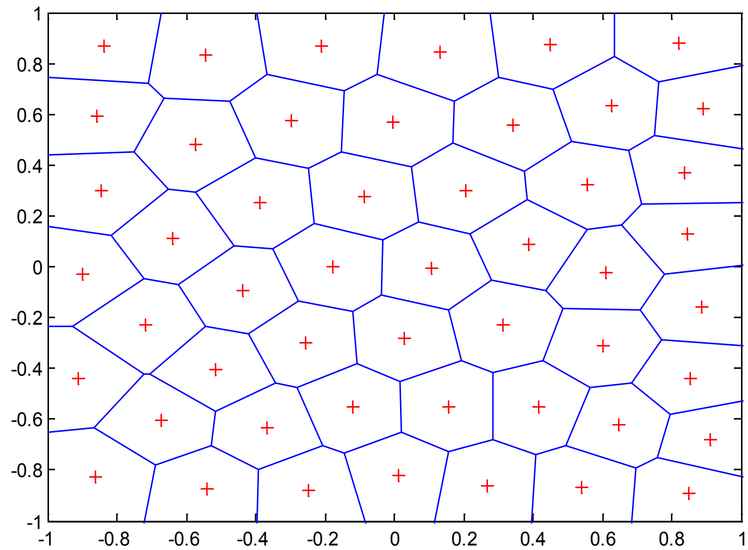


Figure 36. Voronoi diagram

By comparing Figure 34, Figure 35 and Figure 36, it can be found that the final transportation plan by applying Algorithm 1 is the same as the Voronoi diagram.

3.7. LP Based Scenario Reduction: Incremental Subset Size

While the initial selection of the sub scenario set is very important, a new method for generating the start sub scenarios based on the following algorithm can be considered. The algorithm is summarized in Table 19.

Table 19. Linear programming based scenario reduction method (incremental subset size)

Algorithm 2: LP scenario reduction method (incremental subset size)
1. Initialization the desired number of samples, K;
2. Initialization a step size d and the start number of samples, K_1;
3. Initialization S with K_1 samples from the super scenario set I, $\left\{S_1, S_2, \ldots, S_{K_1}\right\}$;
4. Initialization tolerance ε ;
5. While the size of $S<K$

6. Randomly choose d samples from I and add to S;
7. While relative step error < ε
8. Solve problem (28) with I and S
9. Obtain the transportation plan and Kantorovich distance
10. Find the K_1 sub sets;
11. For each sub set, calculate the cost for each scenario;
12. For each sub set, find the scenario with minimum cost as the new central scenario;
13. Update the K_1 central scenarios, $\{S'_1, S'_2, \ldots, S'_{K_1}\}$;
14. Update S with the new K_1 central scenarios;
15. Calculate the relative step error
16. End
17. End
18. Return S as the final sub scenario set

In Algorithm 2, a start sub scenario set S with a size (i.e., K_1) smaller than the desired number of scenarios (i.e., K), is generated first. A number of d scenarios are randomly chosen and add to S'. Then, Algorithm 1 is used to find the final preserved scenario set S'. It's the same way to find the K_1 subsets and calculate the relative step error as Algorithm 1. The size of S' is updated to K_1+d after 1 iteration. While the size of S' is less than K, keep conducting the procedure of applying Algorithm 1 and adding new scenarios. The sub scenario set S' can be taken as an initial preserved scenario set when its size is updated to K. The corresponding final preserved scenario set can be obtained by applying S' to Algorithm 1.

In order to find the best step size d, a 2-dimensional super set of 1000 samples is generated. The size of the subset is set to 50. K is set to 1. The step size d is set to 1, 5, and 10, respectively. For this case, if the MILP based scenario reduction method is applied, the obtained optimal Kantorovich distance is 0.2565. Apply Algorithm 2 with different d, and the obtained results are shown in Figure 37. It can be seen that a smaller d is better.

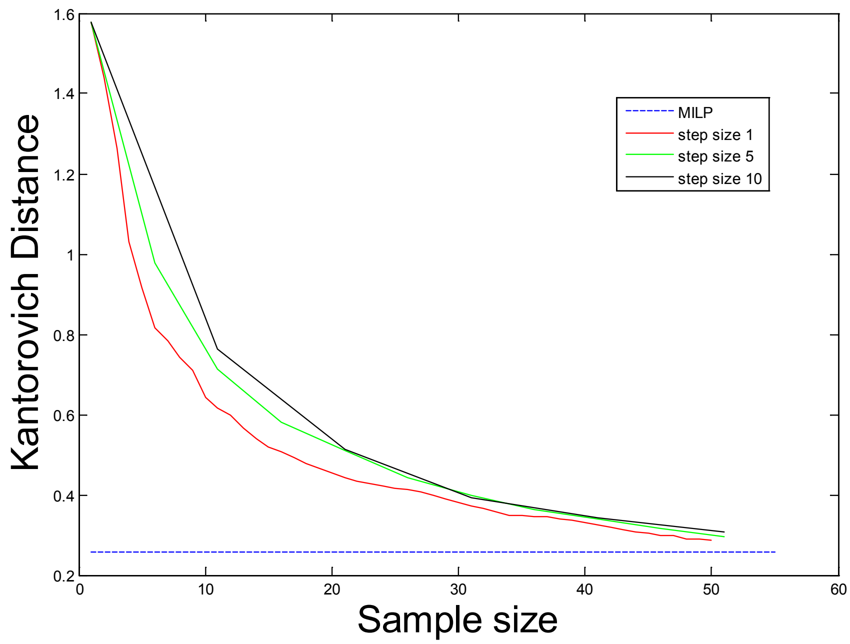


Figure 37. Results for the different step size

In order to compare Algorithm 1 and Algorithm 2, two 2-dimensional super scenario sets are generated based on the normal distribution with mean 0 and variance 1. The size of them is 1000 and 10000, respectively. The desired number of scenarios is 50 (i.e., K=50). For Algorithm 2, set K_1=1 and d=1, and the initial selection of sub scenario set for Algorithm 1 is conducted by k-means clustering. The tolerance ε is set to 0.01. The obtained results are shown in Table 20.

Table 20. Comparison of two LP based scenario reduction methods

	1000		10000	
	LP Fixed	LP Incremental	LP Fixed	LP Incremental
Kantorovich Distance	0.2809	0.2910	0.2955	0.2966
CPU time (s)	3	108	28	1203

It can be seen form Table 20 that the final Kantorovich distance of Algorithm 2 is larger than the one obtained by Algorithm 1, the LP based scenario reduction with fixed subset size, for both cases, which means the initial preserved scenario set generated by k-means clustering is better. The computational time of the LP based scenario reduction method (fixed subset size) is smaller for both cases.

3.8. Extension of LP Based Scenario Reduction

The MILP based scenario reduction method OSCAR proposed in Li & Floudas[41] minimized both the probabilistic distance and the differences between the base, worst and expected performance of the output, while the LP based scenario reduction method proposed in this thesis just considered minimizing the probabilistic distance. However, it is possible to extend our method. The difference of expected performance is calculated by the expected value of the difference between the objective value obtained by the original scenario set and the selected scenario set. This procedure is a linear programming. As for the difference of the best performance, it is obtained by the largest objective value of both the original scenarios and the selected scenarios. In order to minimizing this difference, it needs to be guaranteed that no less than one of the largest objective values obtained by the original scenarios should be included in the ones obtained by the selected scenarios, i.e., the scenario corresponding to the largest objective value should be selected. Binary variables should be introduced. However, only a very small number of binary variables should be generated. Similarly, the scenario corresponding to the smallest objective value should be included in the selected scenarios to realize minimizing the difference of the worst performance of the output of both the original scenario set and the selected scenario set. Both the worst and best performance can be considered in the proposed LP based scenario reduction. Although the introduced binary variables make the problem a MILP problem, there are only a very small number of them. If just consider the expected performance, which still makes the optimization problem an LP problem, this optimization problem can be formulated as follows:

$$
\begin{aligned}
&\min_{\eta_{i,i'},fexp_{err}} \sum_{i\in I}\sum_{i'\in S}\eta_{i,i'}c_{i,i'} + fexp_{err} \\
&s.t.\ \sum_{i\in I}\eta_{i,i'} = p_{i'}^{new}, \forall i' \in S \\
&\quad \sum_{i'\in S}\eta_{i,i'} = p_{i'}^{orig}, \forall i \in I \\
&\quad \eta_{i,i'} \ge 0, \forall i \in I, \forall i' \in S \\
&\quad fexp_{err} \ge \sum_{i'\in S} p_{i'}^{new} fobj_{i'} - \sum_{i\in I} p_{i}^{orig} fobj_{i} \\
&\quad fexp_{err} \ge -\sum_{i'\in S} p_{i'}^{new} fobj_{i'} + \sum_{i\in I} p_{i}^{orig} fobj_{i}
\end{aligned} \tag{31}
$$

where $fexp_{err}$ is the error between the expected objective value of the original scenario set and the preserved scenario set, $fobj_i$ is the objective value obtained by scenario i. $fexp_{err} = \left| fexp_{orig} - fexp_{new} \right|$, where $fexp_{orig}$ is the expected value of the objective value of the original scenario set, and $fexp_{new}$ is the one for the preserved scenario set. Note that $fexp_{orig} = \sum_{i\in I} p_i^{orig} fobj_i$ and $fexp_{new} = \sum_{i'\in S} p_{i'}^{new} fobj_{i'}$, then the last two constraints of problem (31) can be added.

3.9. Summary

In this section, a new scenario reduction algorithm is proposed based on a linear optimization model. The proposed method selects a representative subset of scenarios and assigns new probabilities to them for a given superset of scenarios. The MILP based scenario reduction method proposed by Li and Floudas[41] can give the optimal transportation plan to make the preserved sub scenario set have a performance closest to the original scenario set. However, its computational complexity can be very high when the size of the original scenario set is very large. The proposed LP based scenario reduction method (fixed subset size) can efficiently reduce the computational complexity, while the selected subset still has a good performance. The difference between the solutions obtained by the proposed method and the MILP based scenario reduction method is very small. The representative of the determined preserved scenario set is acceptable, and close to the optimal one obtained by the MILP based scenario reduction. The difference lies in the determined transportation plans. The effectiveness of the proposed method will be demonstrated in the next section. The proposed Algorithm 1, LP based scenario reduction method (fixed subset size), is better than the incremental subset size LP based scenario reduction method, Algorithm 2, due to the better selection of the initial preserved scenario set, so the demonstration of the effectiveness will use Algorithm 1, the fixed subset size LP scenario reduction.

4. Sample Average Approximation for CCP with Scenario Reduction

4.1. Introduction

Sample average Approximation is a sampling based approximation method for solving chance constrained optimization problems. It seeks safe or conservative approximations which can be solved efficiently. The proposed approximation problem is convex and yields feasible solutions. Such an applicable approximation is attractive because it allows efficient generation of feasible solutions. The risk level in the sample average approximation problem can be positive, i.e., all sampled constrains to be satisfied is not required. It can be chosen optimally which constraints will be satisfied. In order to conduct the approximation, the actual distribution of the random vector ξ is replaced by the empirical distribution obtained from the random sample. Samples need to be generated to obtain the empirical measure to approximate the chance constraint, and the choice of its size is an essential work for conducting the approximation. The optimal solution of the approximated problem will converge exponentially fast to their true counterparts in the original chance constrained problem as the sample size increases. However, due to the limitation of the computational resources, it is always impossible to involve such a large number of samples, and the computational complexity will be very high. As a result, a demand of using fewer samples to replace the original large number of samples, while still guarantee the reliability and optimality, is raised. A MILP based optimal scenario reduction method OSCAR was proposed by Li and Floudas[41] recently. It can select the optimal sub scenario set to represent the original scenario set. However, its computational complexity will be very high when the size of the original scenario set is very large. The fixed subset size LP based scenario reduction method proposed in Section 3 is another one for selecting the representative subset.

4.2. Sample Average Approximation Method for CCP

The chance constraint in problem (1) can be rewritten in the following equivalent formulation:

$$P\{h(x,\xi)>0\}\le\alpha \tag{32}$$

Let $P(x):=P\{h(x,\xi)>0\}$. An indicator function of $(0,\infty)$ is defined as follows:

$$1_{(0,\infty)}(x)=\begin{cases}0, & if\ x\le 0\\ 1, & if\ x>0\end{cases} \tag{33}$$

Then, the following relationship holds:

$$P(x)=\mathbf{E}\left[1_{(0,\infty)}\left(h(x,\xi)\right)\right] \tag{34}$$

Generate N realizations of the random parameters ξ, $\{\xi^1,\xi^2,\ldots,\xi^N\}$. The sample average approximation $P_N(x)$ of function $P(x)$ can be obtained by replacing the actual distribution P by the empirical measure P_N [40], i.e.,

$$P_N(x)=\mathbf{E}_{P_N}\left[1_{(0,\infty)}\left(h(x,\xi)\right)\right]=\sum_{t=1}^{N} prob(t)1_{(0,\infty)}\left(h\left(x,\xi^t\right)\right) \tag{35}$$

where $prob(t)$ represents the probability for each realization. If all the realizations have equal probability, then it has a value of $1/N$ for each realization.

The sample average approximation of problem (1), which is associated with the generated sample $\{\xi^1,\xi^2,\ldots,\xi^N\}$, is

$$\begin{aligned}&\max_{x\in X} f(x)\\ &s.t.\ \sum_{t=1}^{N} prob(t)1_{(0,\infty)}\left(h\left(x,\xi^t\right)\right)\le\gamma\end{aligned} \tag{36}$$

Here, the significance level γ can be different from the significance level α of the original chance constrained problem[38]. The convergence analysis of the SAA problem (36) can be found in Luedtke & Ahmed[38], and a complementary study is provided by Pagnoncelli et al.[40]. With increase of the sample size N, an optimal solution of the SAA problem will approach an optimal solution of the original CCP problem (1) with the significance level γ.

4.3. Sample Size Selection

For a feasible solution of the SAA problem $\overline{x} \in X$, we can have $P_N\left(\overline{x}\right) \le \gamma$. This means that there are no more than $\gamma \cdot N$ times the satisfaction of $h\left(\overline{x}, \xi^t\right) > 0$ is achieved during N trails[40]. Since $P\left(\overline{x}\right)$ is the probability of $h\left(\overline{x}, \xi^t\right) > 0$ is satisfied, it follows that:

$$\Pr\left\{P_N\left(\overline{x}\right) \le \gamma\right\} = B\left(m, P\left(\overline{x}\right), N\right) \tag{37}$$

where m denotes the integer part of $\gamma \cdot N$, $B\left(m, P\left(\overline{x}\right), N\right) = \sum_{i=0}^{m} \binom{n}{i} P\left(\overline{x}\right)^i \left(1 - P\left(\overline{x}\right)\right)^{N-i}$ is the binomial distribution. Due to the fact that the binomial distribution is a monotonically increasing function about $P\left(\overline{x}\right)$ and $P\left(\overline{x}\right) \le \alpha$, the following relationship holds:

$$\Pr\left\{P_N\left(\overline{x}\right) \le \gamma\right\} = B\left(m, P\left(\overline{x}\right), N\right) \le B\left(m, \alpha, N\right) \tag{38}$$

Given $\delta \in \left(0,1\right)$ and $\alpha \in \left(0,1\right)$, N^* which satisfies $B\left(m, \alpha, N^*\right) \le \delta$ is a sample size that yields a solution $\hat{x}$ making $P_N\left(\hat{x}\right) \le \gamma$ have a probability no less than $1-\delta$ [51]. Here, m is a nonnegative integer no larger than N^*.

According to the study of Alamo et al.[51], under the assumption that X is a convex and closed set, the function $h\left(x, \xi\right)$ is convex in X, and the sample average approximation problem (36) is always feasible and attains a unique optimal solution for all possible multi-sample extractions $\left\{\xi^1, \xi^2, \ldots, \xi^N\right\}$, given $\alpha \in \left(0,1\right)$ and $\delta \in \left(0,1\right)$, if

$$N \ge \inf_{b>1} \frac{1}{\alpha} \frac{b}{b-1} \left(\ln \frac{1}{\delta} + n \ln b \right) \tag{39}$$

then, the optimal solution $\hat{x}_N$ to the optimization problem (36) satisfies the inequality $P_N(\hat{x}_N) \le \gamma$ with probability no less than $1-\delta$. n is the number of variables. A suboptimal sample size bound can be obtained by taking b equal to the Euler constant:

$$N \ge \frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta}+n\right) \tag{40}$$

However, in reality, the problems can have a very large number of variables, which will make the sample size very large. In order to reduce the computational complexity, a smaller set size can be considered to replace the original set size, i.e., select a subset from the original scenario set. The appropriate set size can be selected by the following algorithm. Starting from a small reduced set size (lower bound), as shown below in (41), and the original set size calculated by (40) (upper bound), the algorithm will gradually reduce the set size until the predefined tolerance is satisfied. The algorithm is summarized in Table 21.

$$N \ge \frac{1}{\alpha}\frac{e}{e-1}\ln\frac{1}{\delta} \tag{41}$$

Table 21. Selection of the reduced set size

Algorithm 3: Reduced set size selection
1. Initialization $[lb, ub]$, and δ, ε' where $lb = 0$ where $ub = n$ 2. Solve SAA problem with set size $\Delta = \frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta}+lb\right)$, and scenarios of size Δ are selected using scenario reduction method from the original scenario set of size which is the lower bound of N obtained by $N = \frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta}+n\right)$ 3. If $P(\hat{x}_N) \le \alpha$

4. return $\frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta}\right)$ as the set size

5. else While $|ub-lb|>\varepsilon'$

6. solve SAA problem with set size

$$\Delta=\frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta}+(lb+ub)/2\right)$$

(scenarios are selected from the original scenarios)

7. if $P\left(\hat{x}_N\right)\le\alpha$

8. $ub=(lb+ub)/2$

9. else if $P\left(\hat{x}_N\right)>\alpha$

10. $lb=(lb+ub)/2$

11. end

12. End

13. End

14. Return $\Delta=\frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta}+lb+1\right)$ as the reduced set size

4.4. Linear SAA Problems with Scenario Reduction

In this section, the proposed Algorithm 3 are applied to several linear SAA problems with the MILP based scenario reduction algorithm OSCAR and the LP based scenario reduction Algorithm 1 proposed in Section 3, respectively. The tolerance is set to $\varepsilon=0.001$ for all the application of Algorithm 1. All of the calculations in Section 4 have been made on a Windows 8 system with an Intel Core i7 (2.40 GHz) and 16 GB of RAM, using MATLAB and GAMS with CPLEX optimizer.

4.4.1. Portfolio Optimization Problem

The portfolio optimization problem is formulated as below:

$$
\begin{aligned}
&\max_{x\in X}\ \bar{r}^T x \\
&\text{s.t.}\quad Prob\left\{r^T x \ge v\right\} \ge 1-\alpha \qquad\qquad (42)\\
&\qquad\quad X = \left\{x \in R^n : e^T x = 1, x \ge 0\right\}
\end{aligned}
$$

where x represents the percentage of a capital invested in each of the available assets, r denotes the vector of random returns of the assets, and $\bar{r}$ is the expected returns of the assets. Historical stock data from Yahoo Finance is obtained for the 29 stocks, MMM, AA, AXP, T, BAC, BA, CAT, CVX, KO CSCO, DIS, DD, XOM, GE, HPQ, HD, IBM, INTC, JNJ, JPM, MCD, MRK, MSFT, PFE, PG, TRV, UTX, VZ, and WMT. We assume the data follow multivariate lognormal distribution and the distribution parameters are estimated using monthly stock price data from January 2003 to December 2013.

By applying the sample average approximation optimization method, problem (42) can be reformulated in a mixed integer linear problem[40]:

$$
\begin{aligned}
&\max_{x\in X}\ \bar{r}^T x \\
&\text{s.t.}\quad r_i^T x + v z_i \ge v \\
&\qquad\quad \sum_{i=1}^{N} z_i \le N\gamma \qquad\qquad (43)\\
&\qquad\quad z \in \left\{0,1\right\}^N
\end{aligned}
$$

In this problem, $\alpha = 0.05$, $n = 29$ and $\delta = 0.001$, so, $lb = 0$ and $ub = 29$, then the upper bound of the set size is 1136, and the lower bound of the set size is 218. Let ε'=1 and $\delta = 0.001$. γ is set to 0, 0.025, 0.05, respectively. If apply the MILP scenario reduction method OSCAR directly, select 218 scenarios from the super set, and run 50 times for each γ. The results are shown in Figure 38, 39 and 40.

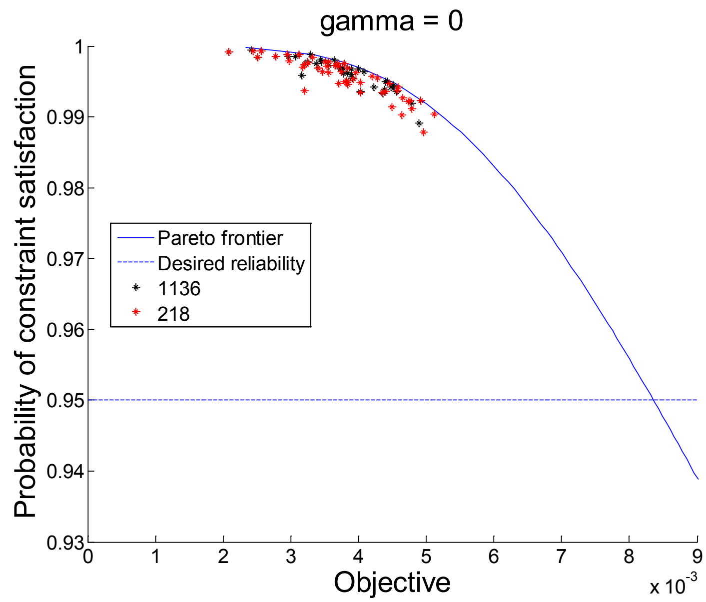


Figure 38. Results for portfolio with OSCAR when γ=0

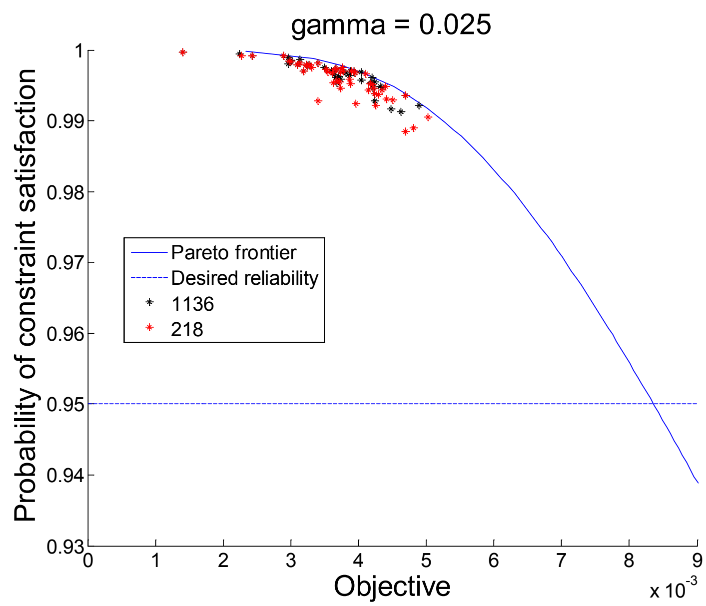


Figure 39. Results for portfolio problem with OSCAR when γ=0.025

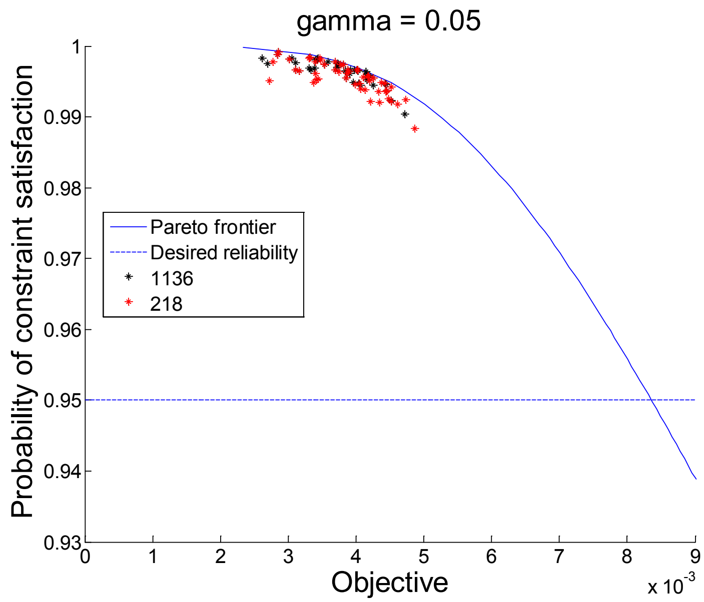


Figure 40. Results for portfolio problem with OSCAR when γ=0.05

From Figure 38-40, it can be seen that, in the first step, where lb is 0, the reduced scenario set satisfies the desired reliability level for all the 50 runs. The Pareto frontier is obtained by transforming the original chance constrained problem into a deterministic formulation, due to the fact that it follows multivariate lognormal distribution[19]. The average objective value and the average probability of satisfaction are shown in Table 22.

Table 22. Average objective value and probability of satisfaction for portfolio problem

		N			
		1136		218	
		obj_{ave}	p_{ave}	obj_{ave}	p_{ave}
γ	0	0.003702	0.99636	0.003797	0.99549
	0.025	0.003643	0.99657	0.003717	0.99582
	0.05	0.003747	0.99643	0.003847	0.99551

In Table 22, obj_{ave} is the average objective value for 50 runs, and p_{ave} is the average probability of satisfaction for 50 runs. The average objective value for the reduced scenario set is larger while the average probability of satisfaction is smaller than the original scenario set.

Algorithm 3 is applied for problem (43) with OSCAR and Algorithm 1, respectively. The obtained results are shown in Table 23.

Table 23. Results for portfolio optimization problem

		N							
		1136		218					
		SAA without scenario reduction		SAA with OSCAR			SAA with LP Fixed		
		Objective	$p_{satisfy}$	Objective	$p_{satisfy}$	CPU time(s)	Objective	$p_{satisfy}$	CPU time(s)
γ	0	0.003826	0.9963	0.003829	0.9951	15.5	0.003988	0.9964	13.9
	0.025	0.003567	0.9974	0.003850	0.9959	15.6	0.004939	0.9903	13.6
	0.05	0.004403	0.9934	0.004495	0.9915	15.5	0.005044	0.9900	13.3

From the above results, it can be observed that the lower bound set size corresponding to lb is the appropriate reduced set size. By applying Algorithm 3 with MILP and LP based scenario reduction, respectively, all the solutions satisfy the desired reliability level for both scenario reduction methods. However, the solution obtained by the MILP based scenario reduction method OSCAR is closer to the original solution when using 1136 scenarios. CPU time represents the time for conducting scenario reduction. Although the objective value of the LP based scenario reduction method is larger than the one obtained by the MILP based scenario reduction method OSCAR, and the corresponding probability of satisfaction $p_{satisfy}$ is less, the results obtained by the LP based scenario reduction still meet the desired reliability and its CPU time for conducting the scenario reduction is less.

4.4.2. Weighted Distribution Problem

A weighted distribution problem is considered here. A company sells n products with m machines. The objective is to minimize of the net cost, that is, the difference between the total cost and the total revenue in the time period. The constraint is the machine availability. The formulation of the problem is shown below[52]:

$$
\begin{aligned}
&\min \quad l \\
&\text{s.t.} \quad \Pr\left\{ l \ge \sum_{j=1}^{m}\sum_{k=1}^{n} c_{jk}x_{jk} + \sum_{k=1}^{n} h_k \left(\sum_{j=1}^{m} p_{jk}x_{jk} - d_k \right)_{+} - \sum_{k=1}^{n} u_k \min\left(\sum_{j=1}^{m} p_{jk}x_{jk}, d_k \right) \right\} \ge 1-\alpha \\
&\qquad \sum_{k=1}^{n} x_{jk} \le a_j \quad \forall j \\
&\qquad x_{jk} \ge 0 \qquad \forall j, \forall k
\end{aligned}
\tag{44}
$$

where $k = 1,\ldots,n$ is the index of products, $j = 1,\ldots,m$ is the index of machines, d_k represents the demand of product k over a given period, p_{jk} is the capacity parameter, i.e., quantity of product k that is produced in a time unit when machine j is allocated to that product, a_j is the limited amount of time that can be used for machine j, c_{jk} is the cost incurred when machine j is allocated to product k for a time unit, u_k is the revenue from selling a unitary of product k, h_k is the inventory holding cost for a unitary quantity of product k. x_{jk} are the decision variables, which represent the amount of time that machine j is allocated to product k.

In this problem, take $m = 5$ and $n = 10$. The corresponding parameters are given as below:

$$[c_{jk}] = \begin{bmatrix} 1.8 & 2.2 & 1.5 & 2.2 & 2.6 & 2.1 & 2.2 & 1.7 & 2.8 & 1.9 \\ 1.6 & 1.9 & 1.3 & 1.9 & 2.3 & 1.9 & 2.0 & 1.5 & 2.5 & 1.7 \\ 1.2 & 1.5 & 1.0 & 1.5 & 1.9 & 1.4 & 1.6 & 1.1 & 2.0 & 1.3 \\ 1.3 & 1.6 & 1.1 & 1.6 & 2.0 & 1.5 & 1.7 & 1.2 & 2.2 & 1.4 \\ 1.2 & 1.5 & 1.0 & 1.6 & 1.9 & 1.5 & 1.6 & 1.1 & 2.1 & 1.3 \end{bmatrix}$$

$$[a_j] = [10 \quad 13 \quad 22 \quad 19 \quad 21]$$

$$[h_k] = [1.3 \quad 1.3 \quad 1.3 \quad 1.3 \quad 1.3 \quad 1.3 \quad 1.3 \quad 1.3 \quad 1.3 \quad 1.3]$$

$$[u_k] = [1.5 \quad 1.8 \quad 1.2 \quad 1.9 \quad 2.2 \quad 1.8 \quad 1.9 \quad 1.4 \quad 2.4 \quad 1.6]$$

Demand d_k follows Dirichlet distribution $380 \times Dir(25,38,18,39,60,35,41,22,74,30)$, and capacity p_{jk} follows independent uniform distribution with a variation of $\pm 5\%$ around the following nominal valus

$$[p_{jk}] = \begin{bmatrix} 5.0 & 7.6 & 3.6 & 7.8 & 12.0 & 7.0 & 8.2 & 4.4 & 14.8 & 6.0 \\ 3.8 & 5.8 & 2.8 & 6.0 & 9.2 & 5.4 & 6.3 & 3.4 & 11.4 & 4.6 \\ 2.3 & 3.5 & 1.6 & 3.5 & 5.5 & 3.2 & 3.7 & 2.0 & 6.7 & 2.7 \\ 2.6 & 4.0 & 1.9 & 4.1 & 6.3 & 3.7 & 4.3 & 2.3 & 7.8 & 3.2 \\ 2.4 & 3.6 & 1.7 & 3.7 & 5.7 & 3.3 & 3.9 & 2.1 & 7.0 & 2.9 \end{bmatrix}$$

Applying SAA method to problem (44), it can be reformulated as below:

$$\begin{aligned} \min \quad & l \\ \text{s.t.} \quad & l - \sum_{j=1}^{m}\sum_{k=1}^{n} c_{jk} x_{jk} - \sum_{k=1}^{n} h_k \phi_k + \sum_{k=1}^{n} u_k \varphi_k \geq 0 \\ & \phi_k - \sum_{j=1}^{m} p_{jk}^{i} x_{jk} + d_k^i \geq u_i \quad \forall k, i \\ & \sum_{j=1}^{m} p_{jk}^{i} x_{jk} - \varphi_k \geq u_i \quad \forall k, i \\ & d_k^i - \varphi_k \geq u_i \quad \forall k, i \\ & u_i + M z_i \geq 0 \\ & \sum_{i=1}^{N} prob(i) \cdot z_i \leq \gamma \\ & \sum_{k=1}^{n} x_{jk} \leq a_j \quad \forall j \\ & x_{jk} \geq 0 \qquad \forall j, \forall k \\ & \phi_k \geq 0 \quad \forall k \end{aligned} \tag{45}$$

where $M = 1000$. Set the desired reliability level as 0.05, and γ is set to 0, 0.025, and 0.05, respectively. δ is set to be 0.001. Use the MILP based scenario reduction method OSCAR to select 218 scenarios from the original 1800 scenarios and run 50 times for each γ. The results are shown in Figure 41-43.

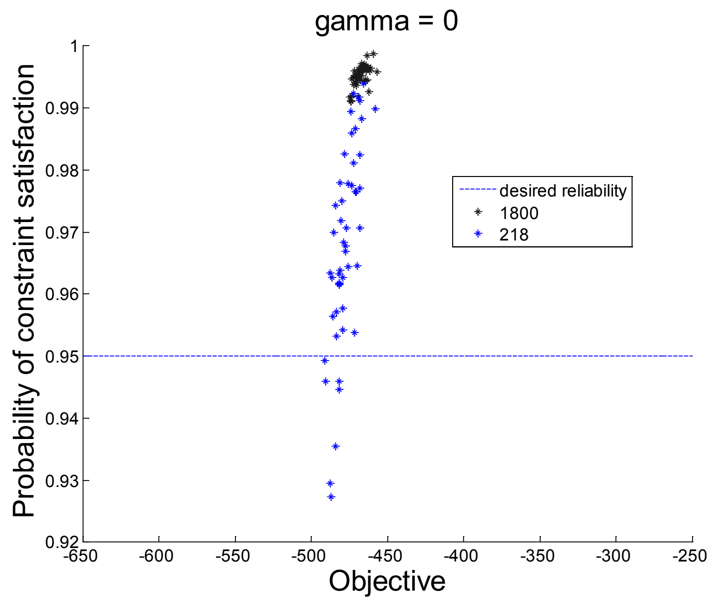


Figure 41. Results for weighted distribution problem with OSCAR for γ=0

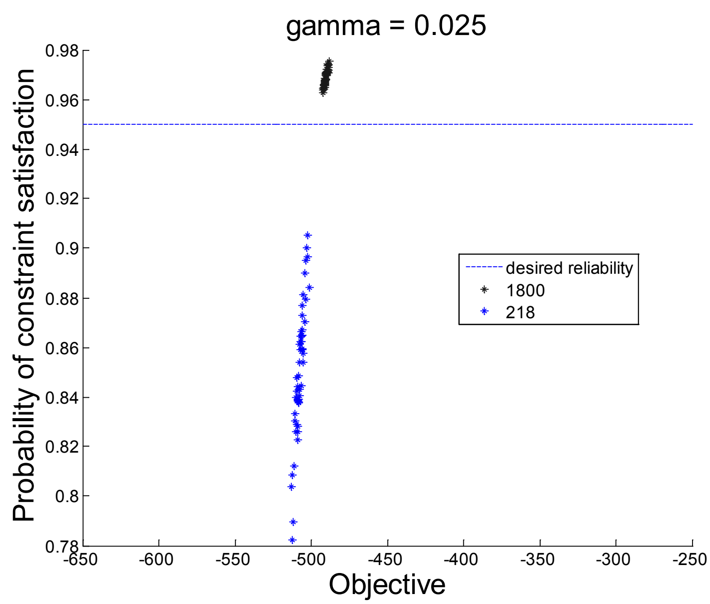


Figure 42. Results for weighted distribution problem with OSCAR when γ=0.025

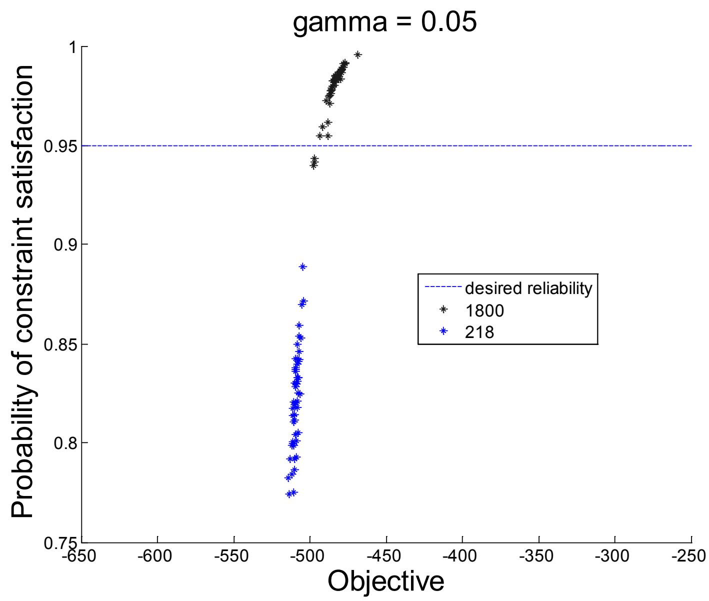


Figure 43. Results for weighted distribution problem with OSCAR when γ=0.05

From Figure 41-43, it can be seen that if set the size of the preserved scenario set equal to the one when $lb = 0$, the selected scenarios can give infeasible solution. The average objective value and the average probability of satisfaction are shown in Table 24.

Table 24. Average objective value and probability of satisfaction for weighted distribution problem

		N			
		1800		218	
		obj_{ave}	p_{ave}	obj_{ave}	p_{ave}
γ	0	-467.6	0.9952	-477.6	0.9669
	0.025	-490.4	0.9690	-507.4	0.8488
	0.05	-483.9	0.9790	-509.1	0.8205

From Table 24, it can be found that the average objective value of the reduced scenario set is smaller. However, the average of probability of satisfaction is smaller and even less than the desired reliability level 0.95.

In order to avoid this problem, Algorithm 3 can be applied. The desired reliability level is $\alpha = 0.015$, and set $\varepsilon' = 1$, $\delta = 0.001$. There are 50 variables, so $lb = 0$ and $ub = 50$. γ is set to 0, 0.0075, and 0.015, respectively. The lower bound of the set size is 728, and the upper bound is 6001. Different desired reliability level is chosen for demonstrating the effectiveness of

Algorithm 1 when the initial number of scenarios is very large. The obtained results are shown below:

Table 25. Results for weighted distribution problem

		γ		
		0	0.0075	0.015
SAA without scenario reduction	N	6001	6001	6001
	Objective	-453.8	-475.0	-473.3
	$p_{satisfy}$	0.9989	0.9940	0.9946
	CPU time(s)	398	1002	1023
SAA with LP Fixed	N	728	1150	1466
	Objective	-460.8	-482.2	-481.8
	$p_{satisfy}$	0.9956	0.9857	0.9867
	CPU time(s)	236	327	388

From Table 25, it can be found that the objective value obtained by Algorithm 3 with LP scenario reduction method is smaller than that of solving the SAA problem directly using the original super set of scenarios. However, the probability of satisfaction is smaller. All the solutions satisfy the desired reliability level, 0.985. The CPU time means the time for solving the SAA problem. It can be seen that the solving time of the reduced scenario set is less. For the MILP based scenario reduction method OSCAR, it is unable to conduct the scenario reduction due to the large number of the original scenario set. The CPU time for conducting the scenario reduction with the LP based scenario reduction method is 452, 16686, and 23687, respectively, for γ=0, 0.0075, and 0.015, respectively.

The solution procedure of Algorithm 3 with LP based scenario reduction method is shown in Table 26.

Table 26. Solution procedure for weighted distribution problem with Algorithm 1

γ					
0		0.025		0.05	
$[lb,ub]$	$[Sizelb,Sizeub]$	$[lb,ub]$	$[Sizelb,Sizeub]$	$[lb,ub]$	$[Sizelb,Sizeub]$
$[0,50]$	$[728,6001]$	$[0,50]$	$[728,6001]$	$[0,50]$	$[728,6001]$

		$[0,25]$	$[728,3365]$	$[0,25]$	$[728,3365]$
		$[0,12]$	$[728,1994]$	$[0,12]$	$[728,1994]$
		$[0,6]$	$[728,1361]$	$[6,12]$	$[1361,1994]$
		$[3,6]$	$[1044,1361]$	$[6,9]$	$[1361,1677]$
		$[3,4]$	$[1044,1150]$	$[6,7]$	$[1361,1466]$

For Algorithm 3 with LP based scenario reduction method, the final lb for different gamma is 0, 3, and 6, respectively, which lead to the returned set size 728, 1150, and 1466, respectively.

4.4.3. Blending Problem

A blending problem is considered here[40]. A farmer wants to use fertilizers to increase the production. The recommended plan is 7*g* of nutrient *A* and 4*g* of nutrient *B*. There are two kinds of fertilizers available: the first has $\omega_1 g$ of nutrient *A* and $\omega_2 g$ of nutrient *B* per kilogram. The second has 1*g* of each nutrient per kilogram. ω_1 and ω_2 are assumed to follow independent continuous uniform in the intervals [1,4] and [1/3, 1], respectively. The problem can be formulated as below[40]:

$$\begin{aligned} &\min_{x_1 \geq 0, x_2 \geq 0} x_1 + x_2 \\ &s.t. \quad \Pr\left\{\begin{matrix} \omega_1 x_1 + x_2 \geq 7 \\ \omega_2 x_1 + x_2 \geq 4 \end{matrix}\right\} \geq 1 - \alpha \end{aligned} \tag{46}$$

The SAA formulation of problem (46) is[40]:

$$\begin{aligned} &\min_{x_1 \geq 0, x_2 \geq 0} x_1 + x_2 \\ &s.t. \quad u_i \leq \omega_{1,i} x_1 + x_2 - 7, i = 1, \ldots, N \\ &\qquad u_i \leq \omega_{2,i} x_1 + x_2 - 4, i = 1, \ldots, N \\ &\qquad u_i + K z_i \geq 0, i = 1, \ldots, N \\ &\qquad \sum_{i=1}^{N} z_i \leq N\gamma \\ &\qquad z_i \in \{0,1\}^N \end{aligned} \tag{47}$$

where N is the number of samples, $\omega_{1,i}$ and $\omega_{2,i}$ are samples from ω_1 and ω_2, $\gamma \in (0,1)$ and K is a positive constant greater or equal than 7. In this case, $\alpha = 0.05$, $\varepsilon' = 1$, and $\delta = 0.001$. There are 2 variables, so $lb = 0$ and $ub = 2$. The upper bound of the set size is 281, and the lower bound of the set size is 218. γ is set to 0, 0.025, and 0.05, respectively. If apply the MILP based scenario reduction method OSCAR directly, select 218 scenarios from the original 281 scenarios, and run 50 times for each γ. The following results can be obtained, as shown in Figure 44-46.

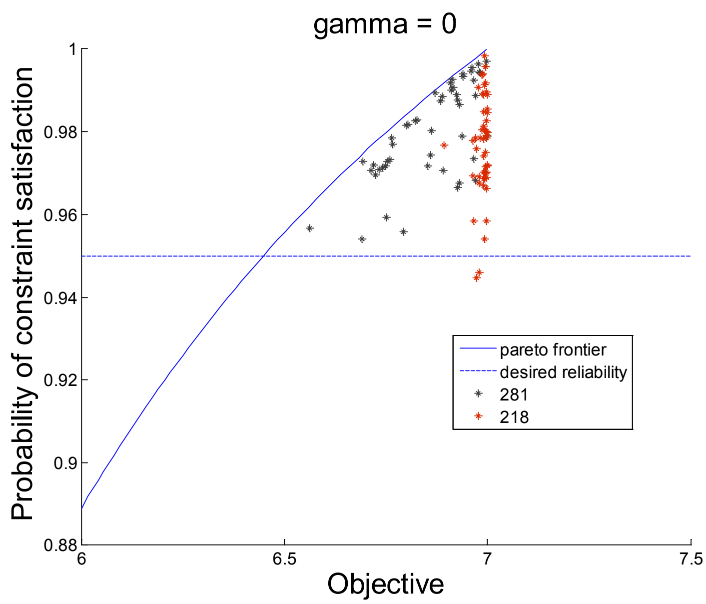


Figure 44. Results for blending problem with OSCAR when γ=0

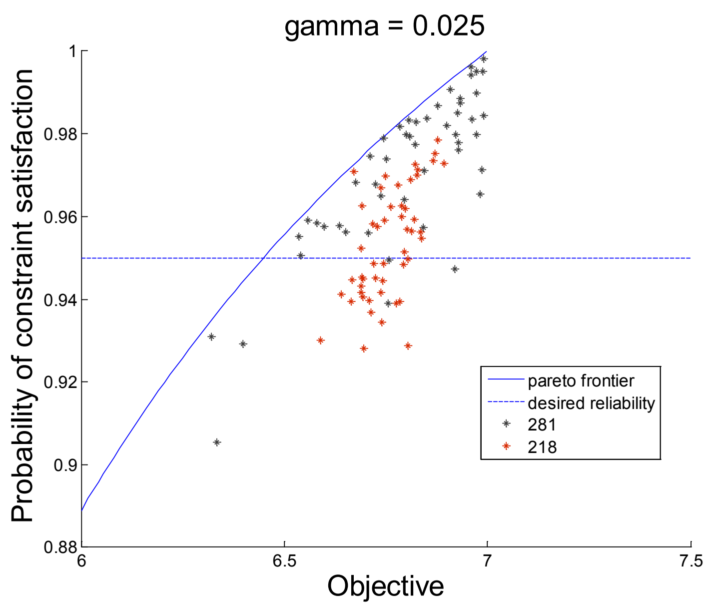


Figure 45. Results for blending problem with OSCAR when γ=0.025

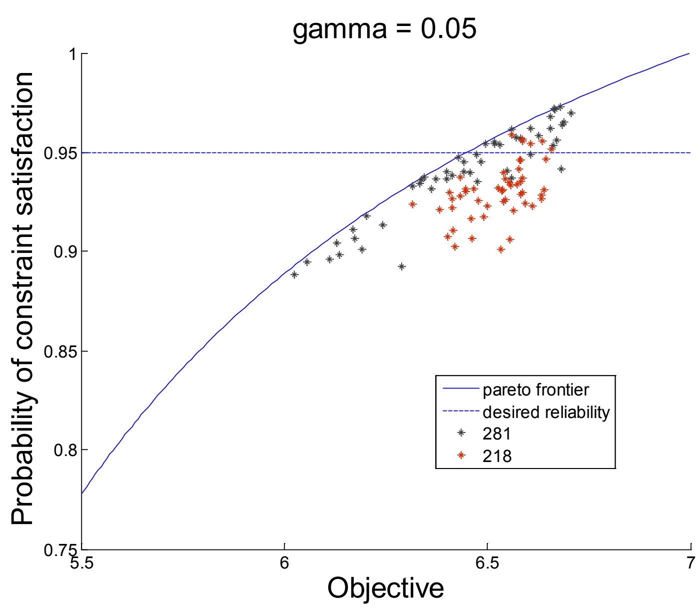


Figure 46. Results for blending problem with OSCAR when γ=0.05

From Figure 44-46, it can be seen that the result can be infeasible if use the OSCAR directly. The Pareto frontier is the actual solution for this problem. It can be explicitly solved for all values of α[40]. For $\alpha \in [1/2, 1]$:

$$x_1^* = \frac{18}{9+8(1-\alpha)} \qquad x_2^* = \frac{2(9+28(1-\alpha))}{9+8(1-\alpha)} \qquad v^* = \frac{4(9+14(1-\alpha))}{9+8(1-\alpha)}$$

For $\alpha \in [0,1/2]$:

$$x_1^* = \frac{9}{11-9(1-\alpha)} \qquad x_2^* = \frac{41-36(1-\alpha)}{11-9(1-\alpha)} \qquad v^* = \frac{2(25-18(1-\alpha))}{11-9(1-\alpha)}$$

where x_1^* and x_2^* are decision variables and v^* is the objective value. The average objective value and average probability of satisfaction for the total 50 runs are shown in Table 27.

Table 27. Average objective value and probability of satisfaction for blending problem

		N			
		281		218	
		obj_{ave}	p_{ave}	obj_{ave}	p_{ave}
γ	0	6.9190	0.9882	6.9190	0.9879
	0.025	6.7267	0.9685	6.7407	0.9694
	0.05	6.4481	0.9423	6.4696	0.9442

From Table 27, it can be found that for both preserved scenario set and original scenario set, the average probability of satisfaction can be smaller than the desired reliability level. This is because the different selection of γ. If γ is larger, the average probability of satisfaction will be smaller.

By applying Algorithm 3, it can be guaranteed to return a feasible solution, and the following results can be obtained, as shown in Table 28.

Table 28. Results for blending problem

		γ		
		0	0.025	0.05
SAA without scenario reduction	N	281	281	281
	Objective	6.914	6.793	6.578
	$p_{satisfy}$	0.9933	0.9821	0.9636
SAA with OSCAR	N	218	218	218
	Objective	6.914	6.727	6.680
	$p_{satisfy}$	0.9868	0.9764	0.9734
	CPU time(s)	2.4	2.3	3.0
SAA	N	218	218	218

with LP Fixed	Objective	6.914	6.727	6.578
	$p_{satisfy}$	0.9933	0.9764	0.9636
	CPU time(s)	3.0	2.0	2.9

From Table 28, it can be seen that the final lb for both methods is 0 for all the γ, which gives a final reduced set size of 218. CPU time represents the time for conducting the scenario reduction, and it is very fast for both methods. It can be found that the results obtained by Algorithm 3 with OSCAR and Algorithm 1 have no big difference, and they are very close to the original solution. Both methods can give a solution that satisfies the desired reliability.

4.5. Nonlinear SAA Problems with Scenario Reduction

In this section, the proposed Algorithm 3 are applied to two nonlinear SAA problems with the MILP based scenario reduction algorithm OSCAR and the LP based scenario reduction Algorithm 1 proposed in Section 3, respectively. The tolerance is set to $\varepsilon = 0.001$ for all the application of Algorithm 1.

4.5.1. Nonlinear Pooling Problem

In this section, the performance of Algorithm 3 on a classical pooling problem is discussed. The formulation given here is closer to the one in Parpas et al.[53]. Figure 47 shows the problem procedure.

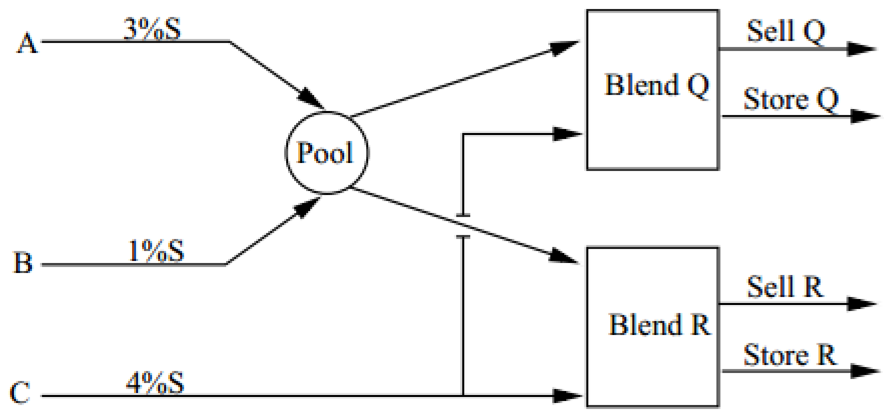


Figure 47. Haverly pooling problem

There are three input chemicals A, B and C. Q and R will be provided when these chemicals blended, and they are assumed uncertain. The cost per unit of raw materials is also uncertain. All the distributions for the uncertainty are assumed to be uniform distribution, and the uncertainty is specified in Table 29.

Table 29. Uncertainty specification

Variable	Support
D_Q : Demand for Q	[50,100]
D_R : Demand for R	[100,200]
C_A : Costs for A	[5,10]
C_B : Costs for B	[10,20]
C_C : Costs for C	[12,25]

The profit per unit for each product is $C_Q = 100$ and $C_R = 150$. Storage cost is $S_Q = 5$ and $S_R = 8$. The formulation of the chance constrained problem is as shown below:

$$
\begin{aligned}
&\max \gamma \\
&s.t.\ \ \Pr\{C_Q \min(\sqrt{x_Q}, \sqrt{D_Q})^2 + C_R \min(\sqrt{x_R}, \sqrt{D_R})^2 - C_A x_A - C_B x_B - C_C x_C \\
&\qquad - S_Q \max(x_Q - D_Q, 0)^2 - S_R \max(x_R - D_R, 0)^2 \geq \gamma\} \geq 1 - \alpha \\
&\quad x_A + x_B = y_Q + y_R \\
&\quad y_Q + x_{CQ} = x_Q \\
&\quad y_R + x_{CR} = x_R \\
&\quad x_{CQ} + x_{CR} = x_C \\
&\quad z \cdot y_Q + 4x_{CQ} \leq 2.5x_Q \\
&\quad z \cdot y_R + 4x_{CR} \leq 1.5x_R \\
&\quad 3x_A + x_B = z\left(x_A + x_B\right) \\
&\quad 1 \leq z \leq 3
\end{aligned}
\tag{48}
$$

where x_i is the amount of chemical or product $i = A, B, C, R, Q$ used, y_j is the flow from the pooling tank to product $j = Q, R$, x_{Cj} is the flow from chemical C to product $j = Q, R$ and z is the Sulphur concentration of the pooling tank. Problem (48) can be reformulated as follows by using SAA method:

$$
\begin{aligned}
&\max \gamma \\
&s.t.\ C_Q v_1 + C_R v_2 - C_{(i,A)} x_A - C_{(i,B)} x_B - C_{(i,C)} x_C - S_Q v_3^2 - S_R v_4^2 - \gamma \geq u_i, i = 1,\ldots,N \\
&\quad D_{(i,Q)} - v_1 \geq u_i, i = 1,\ldots,N \\
&\quad D_{(i,R)} - v_2 \geq u_i, i = 1,\ldots,N \\
&\quad v_3 - x_Q + D_{(i,Q)} \geq u_i, i = 1,\ldots,N \\
&\quad v_4 - x_R + D_{(i,R)} \geq u_i, i = 1,\ldots,N \\
&\quad u_i + \lambda_i \geq 0,, i = 1,\ldots,N \\
&\quad \sum_{i=1}^{N} prob(i) z_i \leq \gamma \\
&\quad \lambda_i \in \{0,1\} \\
&\quad x_Q \geq v_1 \\
&\quad x_R \geq v_2 \\
&\quad v_3 \geq 0 \\
&\quad v_4 \geq 0 \\
&\quad x_A + x_B = y_Q + y_R \\
&\quad y_Q + x_{CQ} = x_Q \\
&\quad y_R + x_{CR} = x_R \\
&\quad x_{CQ} + x_{CR} = x_C \\
&\quad z \cdot y_Q + 4x_{CQ} \leq 2.5x_Q \\
&\quad z \cdot y_R + 4x_{CR} \leq 1.5x_R \\
&\quad 3x_A + x_B = z(x_A + x_B) \\
&\quad 1 \leq z \leq 3
\end{aligned}
\tag{49}
$$

where v_1, v_2, v_3, and v_4 are introduced positive variables, λ_i are the introduced binary variables, u_i are introduced variables.

For this problem, let $\alpha = 0.05$, $\delta = 0.001$, and $\varepsilon' = 1$. There are 16 variables, so $lb = 0$ and $ub = 16$. Then, the lower bound of the reduced set size is 218, and the upper bound is 724. γ is set to 0, 0.025, and 0.05, respectively. If apply the MILP based scenario reduction method OSCAR directly, select 218 scenarios from the original 724 scenarios, and run 50 times for each γ. The results are shown in Figure 48-50.

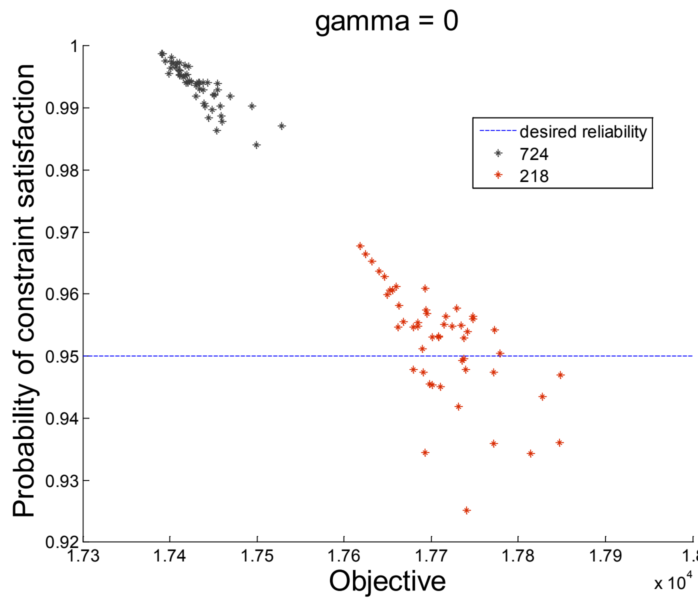


Figure 48. Results for nonlinear pooling problem with OSCAR when γ=0

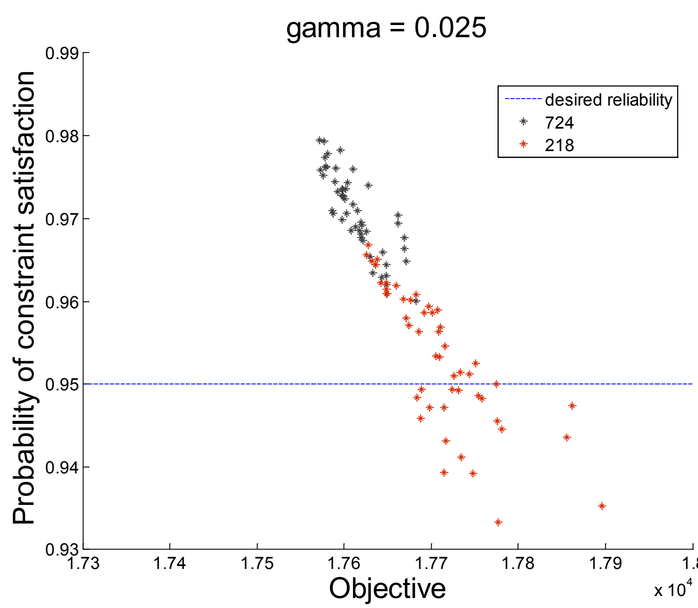


Figure 49. Results for nonlinear pooling problem with OSCAR when γ=0.025

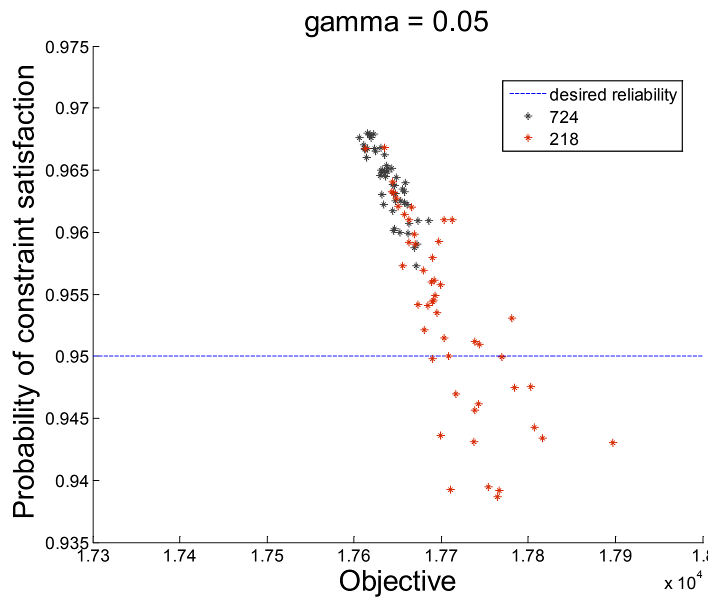


Figure 50. Results for nonlinear pooling problem with OSCAR when γ=0.05

From Figure 48-50, it can be found that if directly select 218 scenarios from the original 724 scenarios, the obtained scenarios may give infeasible solution. The average objective value and average probability of satisfaction are shown in Table 30.

Table 30. Average objective value and probability of satisfaction for nonlinear pooling problem

		N			
		724		218	
		obj_{ave}	p_{ave}	obj_{ave}	p_{ave}
γ	0	17913.3	0.9952	17985.9	0.9579
	0.025	17891.2	0.9760	17967.0	0.9494
	0.05	17945.5	0.9649	18033.7	0.9502

From table 30, it can be found that the average objective value of the reduced scenario set is larger, while the average probability of satisfaction is smaller than the original scenario set. The average probability of satisfaction even can be less than the desired reliability level for the reduced scenario set.

In order to avoid this problem, Algorithm 3 can be applied. By applying Algorithm 3, with the MILP based scenario reduction method OSCAR and Algorithm 1, respectively, the following results can be obtained, as shown in Table 31.

Table 31. Results for nonlinear pooling problem

		γ		
		0	0.025	0.05
SAA without scenario reduction	N	724	724	724
	Objective	17388.5	17596.8	17689.3
	$p_{satisfy}$	0.9986	0.9745	0.9618
SAA with OSCAR	N	281	250	250
	Objective	17670.6	17650.9	17730.9
	$p_{satisfy}$	0.9622	0.9610	0.9569
	CPU time(s)	31.5	24.5	37.1
SAA with LP Fixed	N	218	218	218
	Objective	17741.7	17762.2	17714.0
	$p_{satisfy}$	0.9576	0.9516	0.9585
	CPU time(s)	7.1	11.7	11.4

From Table 31, it can be found that, the results obtained by the MILP based scenario reduction method OSCAR is closer to the solution of the original scenario set. However, the results of the LP based scenario reduction method still satisfy the desired reliability level, 0.95. The objective value when $\gamma = 0,\ 0.025$ is larger than that of MILP based scenario reduction method OSCAR. The CPU time for conducting the scenario reduction is also less for Algorithm 3 with LP based scenario reduction.

For the LP based scenario reduction method, its reduced set size is obtained when $lb = 0$. The solution procedure of Algorithm 3 with the MILP based scenario reduction is shown in the following Table 32.

Table 32. Solution procedure for nonlinear pooling problem with OSCAR

γ					
0		0.025		0.05	
$[lb, ub]$	$[Sizelb, Sizeub]$	$[lb, ub]$	$[Sizelb, Sizeub]$	$[lb, ub]$	$[Sizelb, Sizeub]$
$[0,16]$	$[218,724]$	$[0,16]$	$[218,724]$	$[0,16]$	$[218,724]$
$[0,8]$	$[218,471]$	$[0,8]$	$[218,471]$	$[0,8]$	$[218,471]$
$[0,4]$	$[218,345]$	$[0,4]$	$[218,345]$	$[0,4]$	$[218,345]$
$[0,2]$	$[218,281]$	$[0,2]$	$[218,281]$	$[0,2]$	$[218,281]$

$[1,2]$	$[250,281]$	$[0,1]$	$[218,250]$	$[0,1]$	$[218,250]$

In Table 32, *Sizelb* and *Sizeub* is the lower bound and the upper bound of the returned set size, respectively. Taking $\gamma = 0$ as an example, the final lb equals to 3, then the returned reduced set size is $\Delta = \frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta} + lb + 1\right) = 345$. For $\gamma = 0.025$, the final lb equals to 1, so $\Delta = \frac{1}{\alpha}\frac{e}{e-1}\left(\ln\frac{1}{\delta} + lb + 1\right) = 281$, so as to the case $\gamma = 0.05$.

4.5.2. Continuous Stirred Tank Reactor Design Problem

This case is taken from Ostrovsky et al.[54]. This design problem is defined as the cost minimization under the product specification, as shown in Figure 51.

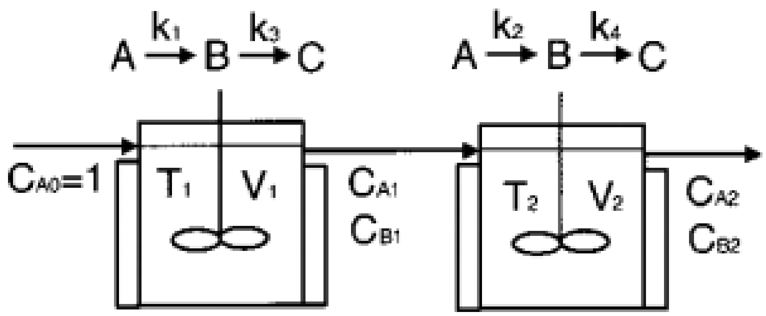


Figure 51. Flowsheet of the reactor network

It is assumed that the uncertainties are from the kinetic parameters (the activation energy and the frequency factor in the Arrhenius equation), while the decision variables are the volumes of both reactors. The chance constrained problem is formulated as[54]:

$$
\begin{aligned}
&\min f = \sqrt{V_1} + \sqrt{V_2} \\
&s.t.\ \ k_1 = k_{10} e^{-E_1/RT_1}, k_2 = k_{10} e^{-E_1/RT_2} \\
&\quad k_3 = k_{20} e^{-E_2/RT_1}, k_4 = k_{20} e^{-E_2/RT_2} \\
&\quad \Pr\left\{ \begin{aligned} &\left(\left(1 - \left(1 + k_1 V_1\right)^{-1} \right) \left(1 + k_3 V_1\right) + \left(1 + k_1 V_1\right) - \left(1 + k_1 V_1\right)^{-1} \left(1 + k_2 V_2\right)^{-1} \right) \cdot \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left(1 + k_4 V_2\right)^{-1} \geq C_{B2}^{SP} \end{aligned} \right\} \geq 1 - \alpha \\
&\quad 0 \leq \left(\left(1 - \left(1 + k_1 V_1\right)^{-1} \right) \left(1 + k_3 V_1\right) + \left(1 + k_1 V_1\right) - \left(1 + k_1 V_1\right)^{-1} \left(1 + k_2 V_2\right)^{-1} \right) \cdot \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left(1 + k_4 V_2\right)^{-1} \leq 1 \\
&\quad 0 \leq V_1, V_2 \leq 16 \\
&\quad 601 \leq T_1 \leq 861 \\
&\quad 542 \leq T_2 \leq 801
\end{aligned}
\tag{50}
$$

where C_{Ai}, C_{Bi}, V_i, and T_i are the concentrations of component *A* and *B*, the volume of both rectors, and the temperature of both reactors, respectively, $i = 1,2$. $RT_1 = 5180.869$, $RT_2 = 4765.169$, and $C_{B2}^{SP} = 0.5$. The random kinetic parameters are assumed to conform to a joint normal distribution, and the data are as shown in Table 33:

Table 33. Joint normal distribution for kinetic parameters

Parameter	Expected value	Standard deviation	Correlation matrix
E_1	6665.948	200	$\begin{bmatrix} 1 & 0.5 & 0.3 & 0.2 \\ 0.5 & 1 & 0.5 & 1 \\ 0.3 & 0.5 & 1 & 0.3 \\ 0.2 & 0.1 & 0.3 & 1 \end{bmatrix}$
E_2	7965.248	240	
k_{10}	0.715	0.0215	
k_{20}	0.182	0.0055	

The SAA problem formulation of problem (50) is:

$$
\begin{aligned}
&\min f = \sqrt{V_1} + \sqrt{V_2} \\
&s.t.\ \ k_{1,i} = k_{10,i} e^{-E_{1,i}/RT_1}, k_{2,i} = k_{10,i} e^{-E_{1,i}/RT_2} \\
&\quad k_{3,i} = k_{20,i} e^{-E_{2,i}/RT_1}, k_{4,i} = k_{20,i} e^{-E_{2,i}/RT_2} \\
&\quad \left(\left(1-\left(1+k_{1,i}V_1\right)^{-1}\right)\left(1+k_{3,i}V_1\right)+\left(1+k_{1,i}V_1\right)-\left(1+k_{1,i}V_1\right)^{-1}\left(1+k_{2,i}V_2\right)^{-1}\right)\cdot \\
&\qquad\qquad\qquad\qquad \left(1+k_{4,i}V_2\right)^{-1} - C_{B2}^{SP} \geq u_i, i=1,\ldots N \\
&\quad 0 \leq \left(\left(1-\left(1+k_{1,i}V_1\right)^{-1}\right)\left(1+k_{3,i}V_1\right)+\left(1+k_{1,i}V_1\right)-\left(1+k_{1,i}V_1\right)^{-1}\left(1+k_{2,i}V_2\right)^{-1}\right)\cdot \\
&\qquad\qquad\qquad\qquad \left(1+k_{4,i}V_2\right)^{-1} \leq 1, i=1,\ldots N \\
&\quad u_i + M\cdot z_i \geq 0 \\
&\quad \sum_{i=1}^{N} prob(i)\cdot z_i \leq \gamma \\
&\quad z_i \in \{0,1\}^N \\
&\quad 0 \leq V_1, V_2 \leq 16 \\
&\quad 601 \leq T_1 \leq 861 \\
&\quad 542 \leq T_2 \leq 801
\end{aligned}
\tag{51}
$$

where $M = 1000$. In order to solve this SAA problem, the parameters are set as: $\alpha = 0.05$, $\delta = 0.001$, $\varepsilon' = 1$. Let $lb = 0$ and $ub = 2$, due to the fact that there are 2 variables. The lower bound of the set size is 218, and the upper bound of the set size is 281. γ is set to 0, 0.025, and 0.05, respectively. If apply the MILP based scenario reduction method OSCAR directly, select 218 scenarios from the original 281 scenarios, and run 50 times for each γ. The results are shown in Figure 52-54.

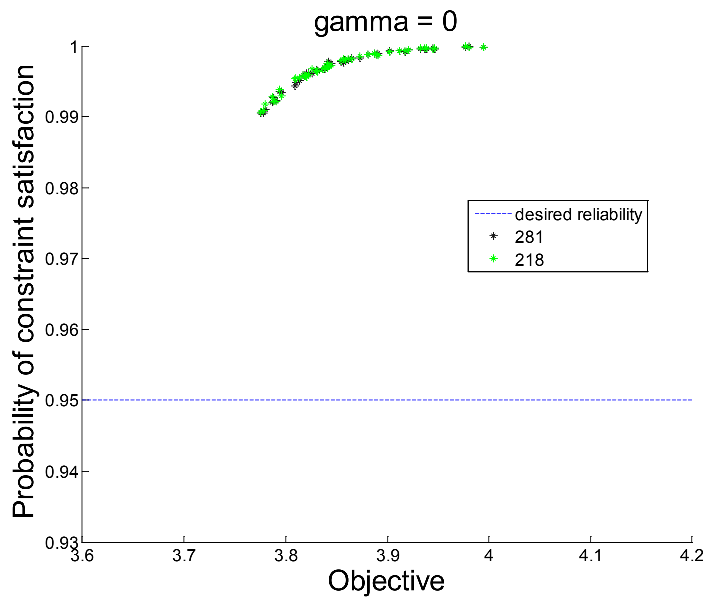


Figure 52. Results for tank reactor design problem with OSCAR when γ=0

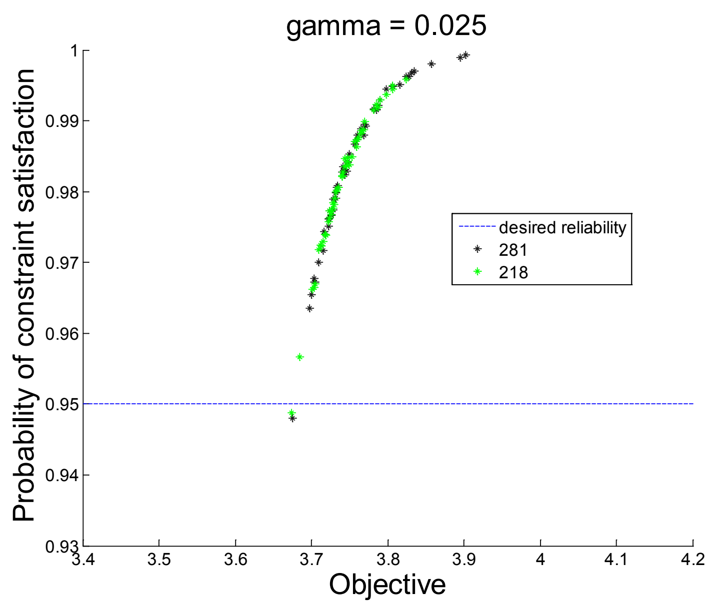


Figure 53. Results for tank reactor design problem with OSCAR when γ=0.025

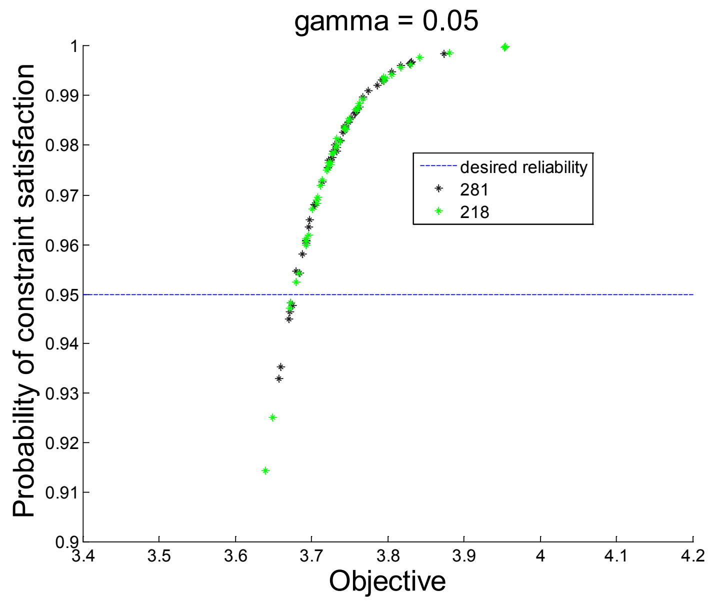


Figure 54. Results for tank reactor design problem with OSCAR when γ=0.05

From Figure 52-54, it can be seen that if directly select 218 scenarios from the original 281 scenarios, the selected scenarios may give infeasible solution. The average objective value and probability of satisfaction for the total 50 runs are shown in Table 34.

Table 34. Average objective value and probability of satisfaction for tank reactor design problem

		N			
		724		218	
		obj_{ave}	p_{ave}	obj_{ave}	p_{ave}
γ	0	3.789	0.9930	3.789	0.9919
	0.025	3.747	0.9848	3.748	0.9823
	0.05	3.713	0.9723	3.713	0.9733

In Table 34, it can be found that the difference between the results obtained by the preserved scenario set and the original scenario set is very small, although there will be some scenarios that lead to infeasible solutions.

By applying Algorithm 3, this problem can be avoided and the following results can be obtained.

Table 35. Results for continuous stirred tank reactor design problem

		γ		
		0	0.025	0.05
SAA	N	281	281	281

without scenario reduction	Objective	3.769	3.719	3.721
	$p_{satisfy}$	0.9893	0.9738	0.9750
SAA with OSCAR	N	218	218	281
	Objective	3.769	3.719	3.721
	$p_{satisfy}$	0.9885	0.9735	0.9761
	CPU time(s)	2.6	2.5	7.5
SAA with LP Fixed	N	218	218	218
	Objective	3.769	3.710	3.736
	$p_{satisfy}$	0.9890	0.9705	0.9808
	CPU time(s)	3.0	2.9	2.3

From Table 35, it can be found that the solution of both scenario reduction methods is very close to the original solution. They both satisfy the desired reliability level. The CPU time for the procedure of scenario reduction is very fast for both methods. For Algorithm 3 with LP based scenario reduction, its final lb is 0 for all the γ. For the one with OSCAR, its final lb equals to 0, 0, and 2, respectively. The solution procedure of Algorithm 3 with OSCAR is shown below in Table 36.

Table 36. Solution procedure for continuous stirred tank reactor design problem with OSCAR

γ					
0		0.025		0.05	
$[lb, ub]$	$[Sizelb, Sizeub]$	$[lb, ub]$	$[Sizelb, Sizeub]$	$[lb, ub]$	$[Sizelb, Sizeub]$
$[0,2]$	$[218,281]$	$[0,2]$	$[218,281]$	$[0,2]$	$[218,281]$
				$[1,2]$	$[250,281]$

4.6. Summary

We have studied sample average approximation for chance constrained optimization problems with scenario reduction, and demonstrated how Algorithm 3 can be used to generate feasible solutions and reduce the computational complexity for three linear problems and two nonlinear problems. Two scenario reduction methods have been considered: the MILP based scenario reduction method OSCAR and the proposed LP based scenario reduction method. It can be found through the study that OSCAR may be intractable for large original scenario set, even though it can provide optimal reduced scenario set. The proposed LP based scenario reduction method is able to provide feasible solution for the original chance constrained optimization problem, and still works well for large initial scenario set. The corresponding solution is very close to the optimal one. By applying scenario reduction method, the resulted sample average approximation problem can be solved efficiently with less computational complexity. Future work in this area will be finding a new method to update the preserved scenarios to obtain the optimal transportation plan (i.e. the optimal preserved scenario subset).

5. Conclusion and Future Work

5.1. Conclusion

Solving optimization problems under uncertainty has received increasing attention in recent years. Chance constraint is one way to formulate optimization problems under uncertainty, which allows the constraint to be satisfied with a certain probability level. However, chance constrained optimization problem is challenging. In order to solve chance constrained problems, the development went into two major directions: analytical approach based approximation method and sampling approach based approximation. In this thesis, two approximation methods are considered: robust counterpart optimization approximation is one of the analytical approximation methods, while sample average approximation is a sampling based approximation method.

There are two contributions of this thesis. The first one is an optimal robust counterpart optimization method. Based on a branch and bound method, the two-stage algorithm can provide the optimal size of the uncertainty set to solve the robust optimization problem, so as to find the optimal robust solution, while still satisfies the desired reliability level. The main drawback of this method is that the computational time can be very large due to the application of the branch and bound method. While computational time is not a practical restriction, the proposed optimal robust optimization method will be the best since it leads to the least conservative robust solution. Section 2 first reviews the set induced robust counterpart optimization problem and three methods for improving the solution quality while it is used to approximate the chance constrained problem. Then, an optimal approximation method is proposed to identify the optimal set size that leads to the optimal robust solution. The proposed two-step algorithm is not restricted by the uncertainty distribution. The obtained robust solution's optimality can be greatly improved while still satisfies the desired reliability level. As a tradeoff, the computational time of the proposed method can be a little large due to the usage of a branch and bound method to find the optimal set size. However, it can balance the solution quality and the computational time.

The other contribution is the LP based scenario reduction method. The MILP based optimal scenario reduction method OSCAR will have a very large computational time when the original number of scenarios is very large, which may be unacceptable. A LP based scenario reduction method is proposed in Section 3. It uses the Kantorovich distance between the original scenario set and the selected sub scenario set to update the scenarios in the subset. By reducing the Kantorovich distance until the desired tolerance reached, the selected subset will become more representative. In the comparison between the proposed LP based scenario reduction method and the GAMS scenario reduction tool Scenred2, both the CPU time and the results of the proposed LP scenario reduction method are better than that of Scenred2. The proposed method is not as good as the optimal scenario reduction method. However, it does not need too much time when the original scenario set has a very large size, while the MILP based optimal scenario reduction method is very hard to get the results due to the high computational complexity. In OSCAR, it considers all the possible scenario subset and selects the optimal one, which will result in intractable problem for large initial scenario set. The LP based scenario reduction method just considers one scenario subset in each iteration to obtain the Kantorovich distance, so the computational complexity can be decreased efficiently. In Section 4, sample average approximation method is applied to three linear optimization problems and two nonlinear optimization problems with scenario reduction. The proposed Algorithm 3 uses the MILP based scenario reduction method OSCAR and the proposed LP based scenario reduction method in Section 3, respectively. Although the proposed LP based scenario reduction method is not the optimal one, it can still give a solution that satisfy the desired reliability, even very close to the optimal solution obtained by the MILP based scenario reduction method, and the computational time for solving the sample average approximation problem and conducting scenario reduction can be significantly reduced when the original number of scenarios is very large.

5.2. Future Work

The optimal robust optimization method proposed in Section 2 is just applied to single chance constrained problem, and the application to joint chance constrained problem will be conducted in the future.

The proposed LP based scenario reduction method is not the optimal one, as shown in Section 3. The final transportation plan has a larger Kantorovich distance and it is not the same as the optimal one obtained by the MILP based scenario reduction method. Based on the principle that minimizing the transportation cost, the proposed scenario reduction method selects the one with a less cost to update the whole selected scenario set. However, it cannot return the optimal transportation plan. In order to get the optimal solution, i.e., the optimal transportation plan, a new method of updating the selected scenarios needs to be addressed in the future.

Bibliography

[1]. Bitran GR, Haas EA, Matsuo H. Production planning of style goods with high setup costs and forecast revisions[J]. Operations Research. 1986,34(2):226-36.

[2]. Lenstra JK, Kan AR, Stougie L. A framework for the probabilistic analysis of hierarchical planning systems[J]. Annals of Operations Research. 1984,1(1):23-42.

[3]. Escudero LF, Kamesam PV, King AJ, Wets RJ. Production planning via scenario modelling[J]. Annals of Operations Research. 1993,43(6):309-35.

[4]. Liu ML, Sahinidis NV. Optimization in process planning under uncertainty[J]. Industrial & Engineering Chemistry Research. 1996,35(11):4154-65.

[5]. Acevedo J, Pistikopoulos EN. Stochastic optimization based algorithms for process synthesis under uncertainty[J]. Computers & Chemical Engineering. 1998,22(4):647-71.

[6]. Gupta A, Maranas CD. A two-stage modeling and solution framework for multisite midterm planning under demand uncertainty[J]. Industrial & Engineering Chemistry Research. 2000,39(10):3799-813.

[7]. Birge JR, Dempstert M. Stochastic programming approaches to stochastic scheduling[J]. Journal of Global Optimization. 1996,9(3-4):417-51.

[8]. Tayur SR, Thomas RR, Natraj N. An algebraic geometry algorithm for scheduling in presence of setups and correlated demands[J]. Mathematical Programming. 1995,69(1-3):369-401.

[9]. Dupačová J, Gaivoronski A, Kos Z, Szantai T. Stochastic programming in water management: A case study and a comparison of solution techniques[J]. European Journal of Operational Research. 1991,52(1):28-44.

[10]. Ouarda T, Labadie J. Chance-constrained optimal control for multireservoir system

optimization and risk analysis[J]. Stochastic environmental research and risk assessment. 2001,15(3):185-204.

[11]. Sreenivasan K, Vedula S. Reservoir operation for hydropower optimization: a chance-constrained approach[J]. Sadhana. 1996,21(4):503-10.

[12]. Dhar A, Datta B. Chance constrained water quality management model for reservoir systems[J]. ISH Journal of Hydraulic Engineering. 2006,12(3):39-48.

[13]. Markowitz H. Portfolio selection*[J]. The journal of finance. 1952,7(1):77-91.

[14]. Ermoliev YM, Ermolieva TY, MacDonald G, Norkin V. Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks[J]. Annals of Operations Research. 2000,99(1-4):207-25.

[15]. Bonami P, Lejeune MA. An exact solution approach for portfolio optimization problems under stochastic and integer constraints[J]. Operations research. 2009,57(3):650-70.

[16]. Pagnoncelli BK, Ahmed S, Shapiro A. Computational study of a chance constrained portfolio selection problem[J]. Optimization Online. 2008,2008.

[17]. Charnes A, Cooper WW, Symonds GH. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil[J]. Management Science. 1958,4(3):235-63.

[18]. Prékopa A. Stochastic programming: Springer; 1995.

[19]. Alizadeh F, Goldfarb D. Second-order cone programming[J]. Mathematical programming. 2003,95(1):3-51.

[20]. Lagoa C, editor On the convexity of probabilistically constrained linear programs. Decision and Control, 1999 Proceedings of the 38th IEEE Conference on; 1999: IEEE.

[21]. Calafiore G, El Ghaoui L. Distributionally robust chance-constrained linear programs with applications[J]. Techical Report, DAUIN, Politecnico di Torino, Torino, Italy. 2005.

[22]. Stambaugh F. Risk and value at risk[J]. European Management Journal. 1996,14(6):612-21.

[23]. Pritsker M. Evaluating value at risk methodologies: accuracy versus computational time[J]. Journal of Financial Services Research. 1997,12(2-3):201-42.

[24]. Philippe J. Value at risk: the new benchmark for controlling market risk[J]. Chicago: Irwin Professional. 1996.

[25]. Artzner P, Delbaen F, Eber JM, Heath D. Coherent measures of risk[J]. Mathematical finance. 1999,9(3):203-28.

[26]. Ogryczak W, editor Stochastic dominance relation and linear risk measures. Financial Modelling–Proc 23rd Meeting EURO WG Financial Modelling, Cracow; 1998.

[27]. Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk[J]. Journal of risk. 2000,2:21-42.

[28]. Rockafellar RT, Uryasev S. Conditional value-at-risk for general loss distributions[J]. Journal of Banking & Finance. 2002,26(7):1443-71.

[29]. Nemirovski A, Shapiro A. Convex approximations of chance constrained programs[J]. SIAM Journal on Optimization. 2006,17(4):969-96.

[30]. Verderame PM, Floudas CA. Operational planning of large-scale industrial batch plants under demand due date and amount uncertainty: II. Conditional value-at-risk framework[J]. Industrial & engineering chemistry research. 2009,49(1):260-75.

[31]. Chebyshev P. On mean values[J]. Complete collected works. 1867,2.

[32]. Bernstein S. Modification of Chebyshev's inequality and error of Laplace's formula[J]. Uch Zap Nauchno-Issled Kafedr Ukr, Otd Mat,(1). 1924:38-48.

[33]. Bernstein SN. On certain modifications of Chebyshev's inequality[J]. Doklady Akademii Nauk SSSR. 1937,17(6):275-7.

[34]. Bernstein S. Teoriia Veroiatnostei.[The Theory of Probabilities.][J]. Moskva-Leningrad: Gosudarstvennoe Izdatel'stvo. 1927.

[35]. Hoeffding W. Probability inequalities for sums of bounded random variables[J]. Journal of the American statistical association. 1963,58(301):13-30.

[36]. Soyster AL. Technical note—convex programming with set-inclusive constraints and applications to inexact linear programming[J]. Operations research. 1973,21(5):1154-7.

[37]. Li Z, Ding R, Floudas CA. A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization[J]. Industrial & engineering chemistry research. 2011,50(18):10567-603.

[38]. Luedtke J, Ahmed S. A sample approximation approach for optimization with probabilistic constraints[J]. SIAM Journal on Optimization. 2008,19(2):674-99.

[39]. Atlason J, Epelman MA, Henderson SG. Optimizing call center staffing using simulation and analytic center cutting-plane methods[J]. Management Science. 2008,54(2):295-309.

[40]. Pagnoncelli B, Ahmed S, Shapiro A. Sample average approximation method for chance constrained programming: theory and applications[J]. Journal of optimization theory and applications. 2009,142(2):399-416.

[41]. Li Z, Floudas CA. Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization[J]. Computers & Chemical Engineering. 2014,70:50-66.

[42]. Calafiore GC, Campi MC. The scenario approach to robust control design[J]. Automatic Control, IEEE Transactions on. 2006,51(5):742-53.

[43]. Nemirovski A, Shapiro A. Scenario approximations of chance constraints. Probabilistic and randomized methods for design under uncertainty: Springer; 2006. p. 3-47.

[44]. Li Z, Tang Q, Floudas CA. A comparative theoretical and computational study on robust counterpart optimization: II. Probabilistic guarantees on constraint satisfaction[J]. Industrial & engineering chemistry research. 2012,51(19):6769-88.

[45]. Li Z, Floudas CA. A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: III. Improving the Quality of Robust Solutions[J]. Industrial & Engineering Chemistry Research. 2014,53(33):13112-24.

[46]. Wang W, Ahmed S. Sample average approximation of expected value constrained stochastic programs[J]. Operations Research Letters. 2008,36(5):515-9.

[47]. Dupačová J, Gröwe-Kuska N, Römisch W. Scenario reduction in stochastic programming[J]. Mathematical Programming. 2003,95(3):493-511.

[48]. Heitsch H, Römisch W. Scenario Reduction Algorithms in Stochastic Programming[J]. Computational Optimization and Applications. 2003,24(2-3):187-206.

[49]. Kantorovich LV, editor On the transfer of masses. Dokl Akad Nauk SSSR; 1942.

[50]. MacQueen J, editor Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability; 1967: Oakland, CA, USA.

[51]. Alamo T, Tempo R, Luque A, editors. On the sample complexity of randomized approaches to the analysis and design under uncertainty. American Control Conference (ACC), 2010; 2010: IEEE.

[52]. Carè A, Garatti S, Campi MC. FAST-Fast Algorithm for the Scenario Technique[J]. Operations Research. 2014,62(3):662-71.

[53]. Parpas P, Rustem B, Pistikopoulos EN. Global optimization of robust chance constrained problems[J]. Journal of Global Optimization. 2009,43(2-3):231-47.

[54]. Ostrovsky G, Ziyatdinov N, Lapteva T, Zaitsev I. Optimization of chemical processes

with dependent uncertain parameters[J]. Chemical Engineering Science. 2012,83:119-27.