Monotony Analysis and Sparse-Grid Integration for nonlinear Chance Constrained Process Optimization

Abebe Geletu, Armin Hoffmann, Klöppel Michael, Pu Li

Monotony Analysis and Sparse-Grid Integration for nonlinear Chance Constrained Process Optimization

Journal:	*Engineering Optimization*
Manuscript ID:	GENO-2010-0113.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	24-Sep-2010
Complete List of Authors:	Geletu, Abebe; Ilmenau University of Technology, Simulation and Optimal Processes Hoffmann, Armin; Ilmenau University of Technology, Dept. of Operations Research and Stochastics Michael, Klöppel; Ilmenau University of Technology, Simulation and Optimal Processes Li, Pu; Ilmenau University of Technology, Simulation and Optimal Processes
Keywords:	chance constraints, nonlinear optimization, monotonicity, sparse-grid integration techniques , chemical process optimization
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
MonoSparseCCOPT-R2.tex	

Engineering Optimization
Vol. 00, No. 00, Month 200x, 1–27

ARTICLE

Monotony Analysis and Sparse-Grid Integration for nonlinear Chance Constrained Process Optimization

Abebe Geletu[(a)]*, Armin Hoffmann[(b)], Michael Klöppel[(a)], Pu Li[(a)]
[(a)] *Department of Simulation and Optimal Processes, Institute for Automation and Systems Engineering,*
[(b)] Department of Operations Research and Stochastic, Institute of Mathematics,
Ilmenau University of Technology, P.O. Box 10 05 65, 98684 Ilmenau, Germany
(*December 2009*)

The numerical solution of a nonlinear chance constrained optimization problem poses a major challenge. The idea of back-mapping (Wendt et al. 2002) is a viable approach for transforming chance constraints on output variables (with unknown distribution) into chance constraints on uncertain input variables (with known distribution) based on a monotony relation. Once transformation of chance constraints has been accomplished, the resulting optimization problem can be solved by using a gradient-based algorithm. However, the computation of values and gradients of chance constraints and the objective function involves the evaluation of multidimensional integrals which is computationally very expensive. This study proposes an easy-to-use method to analyze monotonic relations between constrained outputs and uncertain inputs. In addition, sparse-grid integration techniques are used to reduce the computational time decisively. Two examples from process optimization under uncertainty demonstrate the performance of the proposed approach.

Keywords: chance constraints; nonlinear optimization; monotonicity; sparse-grid techniques; chemical process optimization.

1. Introduction

1.1. *Motivation*

Many practical models from process engineering have uncertain model parameters. Frequently, it is required to guarantee an optimal and robust performance under such uncertainties. The optimization of some performance function under uncertainty is most preferably accomplished through the use of stochastic optimization techniques

*Corresponding author. Email: abebe.geletu@tu-ilmenau.de

ISSN: 0305-215X print/ISSN 1029-0273 online

DOI: 10.1080/0305215YYxxxxxxxx
http://www.informaworld.com

(Prekopa 1995). The conventional way of solving stochastic optimization problems has been by using expected (nominal) values for the uncertain parameters. Nevertheless, the solutions obtained through such simplifications may not be robust. Therefore, stochastic optimization methods are essential to consider the full effect of the uncertain parameters.

In most practical applications, constraint violations are unavoidable as a result of unexpected and extreme events and measurement errors, etc. Therefore, a requirement for strict (deterministic) satisfaction of feasibility may incur high costs of operating the process. In such situations, it is reasonable to demand constraint satisfaction with some degree of probability. This can be achieved by formulating a *chance constrained optimization* problem (CCOPT). In chance constrained optimization, inequality constraints are expected to be satisfied by a certain pre-given probability level. Hence, this approach is less conservative and allows some level of constraint violations. In practice, chance constraints can be used to specify a guaranteed level of fulfilment of product specifications, a guaranteed level of availability of products or outputs, safety conditions and risk-aversion, etc. There are abundant real-life problems with chance constraints widely studied and applied in chemical process engineering (Arellano-Garcia *et al.* 2007, 2008, 2009; Flemming *et al.* 2007; Henrion *et al.* 2001, 2003; Li 2007; Li *et al.* 2002; Straub and Grossmann 1993; Wendet and Wozney 2002); in risk metrics (Rockafellar and Uryasev 2000); in finance and economics (Charnes and Cooper 1959); in water resource management (Prekopa 1995); in reliability based design optimization (Royset and Polak 2004; Royset *et al.* 2006), to mention only a very few.

This paper considers a chance constrained stochastic optimization model of the form

$$(CCOPT) \qquad \min_{u} \left[E(f_1(x,u,\xi)) + \gamma \cdot Var(f_2(x,u,\xi))\right] \tag{1}$$

such that

$$g(x,u,\xi) = 0; \tag{2}$$

$$Pr\{x_i^{min} \leq x_i \leq x_i^{max}\} \geq \alpha_i, i \in I; \tag{3}$$

$$u \in \mathcal{U}; \tag{4}$$

$$x \in \mathcal{X}, \xi \in \mathcal{W}. \tag{5}$$

Here, x, u, ξ are the vectors of output variables, decision (control) variables and uncertain (input) variables, respectively. The set $\mathcal{X} \subset \mathbb{R}^n$ is an open and convex set, while $\mathcal{W} \subset \mathbb{R}^p$ is also an open convex set containing uncertain variables and the set of control variables $\mathcal{U} \subset \mathbb{R}^m$ is compact. It is assumed that the functions $f_1, f_2 : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^n$ are at least one time continuously differentiable functions and $I = \{1, \ldots, n_1\}$, $|I| < n$, where $|I|$ stands for the number of indices in I. The operators $Pr(\cdot)$, $E(\cdot)$ and $Var(\cdot)$ represent probability, expected value and variance, respectively, and $\gamma \geq 0$ is a weighting factor.

Due to the randomness of ξ and the relation (2), for any given value of u, the output variables x are also random. Hence, the constraint sets in (3), $\{x \in \mathbb{R}^n \mid x_i^{min} \leq x_i \leq x_i^{max}\}$, $i \in I$, are random and are given in terms of their probability measures, i.e., $Pr\{x_i^{min} \leq x_i \leq x_i^{max}\} \geq \alpha_i$, $i \in I$. These constraints specify that individual constraints in equation (3) to be satisfied with separately given

probability (reliability) levels α_i, where $\alpha_i \in (0,1), i \in I$. Such chance constraints are called "*separate*" (or *single*) *chance constraints*. In chemical process models, the chance constraints (3), for instance, can specify the holding of minimum and maximum level of the specifications for product x_i the holding of allowed lower and upper volume levels of a buffer-tank, etc.

Optimization problems in process engineering posses special structures, where mass- or energy balance equations of the type $g(x,u,\xi) = 0$ are commonly available as equality constraints. Frequently, these model equations are nonlinear and the (state) variables x can only be represented implicitly in terms of u and ξ, as indicated by $x(u,\xi)$. In general, it is difficult to determine the distribution of the random output variables x directly. As a result, the computation of the values and gradients of the chance constraints (3) poses substantial difficulties in solving the CCOPT problem.

1.2. *A review of methods for computation of chance constraints*

To overcome the difficulties associated with the computation of chance constraints, there are three major approaches (considering only non-heuristic approaches) in the open literature: *analytic approximation*, *approximate discretization* and *back-projection*.

Analytic approximation techniques attempt to replace chance constraints with bounding confidence regions which are usually easy to drive for normal distributions. Nermirovski and Shapiro 2006 suggest convex approximation of chance constraints in order to improve tractability. These considerations are mainly confined to chance constraints whose defining functions are linear with respect to the uncertain variables. As a result, chance constraints are replaced by quadratic constraints and the optimization problem reduces to a *robust optimization* problem (see, for instance, Ben-Tal *et al.* 2009, Kropat *et al.* 2010, Özmen *et al.* 2010, Weber *et al.* 2010). These approaches have the danger of either over or under estimation of chance constraints, leading to conservative or unreliable approximation of chance constraints. Recently, Garnier *et al.* 2009 introduced linearization of chance constraints for the computation of derivatives of chance constraints. But, the suggested technique is shown to work only when the variance of the uncertain variables is very small.

The majority of *discrete approximation techniques* generate random samples for the uncertain input variables through Monte Carlo methods and use averaged sums of function values to approximate chance constraints. Scenario generation methods (Califore and Campi 2005, Henrion *et al.* 2001) are those of such approaches. Despite the fact that this approach is applicable irrespective of the type of distribution function of the uncertain variables (Gaussian or non-Gaussian), the requirement of feasibility of constraints for almost all possible realizations of the uncertain variables leads to a conservative approach and a deterioration in the values of the objective function. The consideration of very large number of scenarios requires also the solution of a very large deterministic optimization problem which is computationally intractable. In particular, scenario generation approaches are less favorable for nonlinear problems. *Variance reduction* techniques like importance sampling, Latin hypercube sampling, Hammersley sequence sampling (Diwekar and Kalagnanam 1997) may provide some improvement over MC methods. Recently, QMC based sample average approximation (Kookos 2003, Pagnoncelli *et al.* 2009, Wang and Ahmed 2008) have been used for the

computation of chance constrained optimization problems. Quasi-Monte Carlo (QMC) methods rely on the generation of uniformly distributed deterministic samples with *low-discrepancy* properties. Thus, there are potential difficulties associated with *sample average approximation* of chance constraints. On the one hand it requires very large sample size to yield efficient estimates for the probability constraints. On the other hand, the discrete approximations of the chance constraints may not be differentiable. Such direct discretization of chance constraints, without considering the special mathematical structures of the problem and the distribution function of the uncertain variables, can lead to redundancy and unnecessary wastage of computational resources (Chen and Merothra 2007, Novak and Ritter 1997, Schürer 2003).

The *back-mapping* (constraint transformation) was proposed by Wendt *et al.* 2002 for the transformation of the chance constraints (3) from the space of output variables into the space of uncertain input variables whose joint distribution is known. The transformation of chance constraints is performed based on the assumption of the existence of strict monotonic relations between a chance constrained variable $x_i, i \in I$, and some uncertain variable ξ_j. Such monotonic relations (probability-measure preserving transformations) are mostly not proved, but frequently assumed or taken for granted through an inspection made on the model under consideration (see, for instance, Arellano-Garcia 2007, 2009; Flemming *et al.* 2007, Li *et al.* 2002, 2008; Wendt *et al.* 2002).

This paper follows the back-mapping approach. To use this approach, two major issues need to be addressed for solving nonlinear chance constrained optimization problems. First, given a nonlinear model, a strict monotonic relation between the constrained output and an uncertain input variable should exist. Second, an efficient integration approach is needed to compute values and gradients of chance constraints and objective functions. First, this paper proposes an easy-to-use method to verify the monotonic relations. This is based on the implicit function theorem. Second, the sparse-grid integration technique is proposed to carry out computation of probability integrals, with which the computation load can be significantly reduced in comparison with full-grid integration or quasi-Monte Carlo techniques.

The rest of the paper is organized as follows. Section 2 presents our solution strategy for CCOPT along with the idea of transformation of chance constraints using monotonicity relations. Section 3 discusses how to determine the required monotonicity relations based on the implicit-function theorem. Section 4 introduces sparse-grid techniques for the computation of chance constraints and their gradients. Section 5 presents two case-studies from process engineering, verifying the applicability of the monotony analysis and the sparse-grid integration technique. The paper concludes with Section 6 with a summary and an outlook for future research endeavors.

2. Solution Strategy for nonlinear CCOPT

The solution strategy for the chance constrained optimization problem (1) - (5) is based on a sequential method with respect to the state equation, i.e., the state (ouput) variables x will be eliminated with the help of the state (model) equations (2). It is assumed that the state equations can be uniquely solved for all state variables. More precisely, we use the following assumption.

ASSUMPTION 2.1 The state equation $g(x,u,\xi)=0$ has a unique solution $x(u,\xi)$, for all (u,ξ) in some open superset V of $\mathcal{U}\times\mathbb{R}^n$ such that the implicit function $(u,\xi)\mapsto x(u,\xi)$ and its partial derivatives $(u,\xi)\mapsto D_{u_k}x(u,\xi), k=1,2,...,m$ are continuous on V. The theoretical existence of such a function can be warranted by the global implicit function theorem (see section 3).

Necessary for such a smooth solvability is the validity of the well-known condition

$$\det D_x g(x,u,\xi)\neq 0$$

for all triplets (x,u,ξ) satisfying (2), (4) and (5) and sufficient smoothness of g with respect to (x,u,ξ), where $D_x g$ denotes the derivative of g with respect to the vector x, i.e., $D_x g:=\left(\frac{\partial}{\partial x_1}g,...,\frac{\partial}{\partial x_n}g\right)$.

First the functions of the objective are described in terms of the control u using the implicit function $(u,\xi)\mapsto x(u,\xi)$:

$$F_1(u)+\gamma F_2(u):=E\left(f_1\left(x(u,\xi),u,\xi\right)\right)+\gamma Var\left(f_2\left(x(u,\xi),u,\xi\right)\right). \tag{6}$$

Then with back-mapping the chance constraints are transformed by studying the non-linear constraints. This requires monotony assumptions which are satisfied in a lot of practical problems (see case studies in Section 5).

DEFINITION 2.2 *The output x_i is **monotonically related** with the uncertain input ξ_j on $\left[x_i^{\min},x_i^{\max}\right]$ if uniformly for arbitrary fixed control $u\in U$ and input $\tilde{\xi}:=(\xi_k)_{k\neq j}$, the following two conditions are satisfied:*

(1) for each output $x_i\in\left[x_i^{\min},x_i^{\max}\right]$ there exists an input $\xi_j\left(x_i,u,\tilde{\xi}\right)$, defined by (2), with $\xi\in\mathcal{X}$;
(2) $x_i^{\min}\leq x_i'<x_i''\leq x_i^{\max}$ implies on the whole interval $\left[x_i^{\min},x_i^{\max}\right]$
a) either $\xi_j\left(x_i',u,\tilde{\xi}\right)<\xi_j\left(x_i'',u,\tilde{\xi}\right)$
b) or $\xi_j\left(x_i',u,\tilde{\xi}\right)>\xi_j\left(x_i'',u,\tilde{\xi}\right)$.

The monotonic relation in case (a) is denoted by $\xi_j\uparrow x_i$ and it is said that ξ_j influences x_i positively. In case (b) it is denoted by $\xi_j\downarrow x_i$ and is described as ξ_j influences x_i negatively. More formally the index j should be written as $j(i)$. In subsequent discussions this is suppressed and simply j is used instead of $j(i)$.

Thus, the interval of outputs $\left[x_i^{\min},x_i^{\max}\right]$ can be back-mapped one-to-one either to the interval of inputs $\left[\xi_j\left(x_i^{\min},u,\tilde{\xi}\right),\xi_j\left(x_i^{\max},u,\tilde{\xi}\right)\right]$ or to the interval of inputs $\left[\xi_j\left(x_i^{\max},u,\tilde{\xi}\right),\xi_j\left(x_i^{\min},u,\tilde{\xi}\right)\right]$. Hence, using the strict monotony relation of x_i with ξ_j and the known joint density function ϕ of ξ, the probability that x_i belongs to the interval $\left[x_i^{\min},x_i^{\max}\right]$ can be written as:

in case a)

$$\Pr\left\{x_i^{\min} \le x_i \le x_i^{\max}\right\} = \Pr\left\{\xi_j\left(x_i^{\min}, u, \tilde{\xi}\right) \le \xi_j \le \xi_j\left(x_i^{\max}, u, \tilde{\xi}\right)\right\}$$
$$= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \int_{\xi_j\left(x_i^{\min},u,\tilde{\xi}\right)}^{\xi_j\left(x_i^{\max},u,\tilde{\xi}\right)} \phi(\xi)\, d\xi_j d\left(\tilde{\xi}\right) =: h_i(u). \tag{7}$$

in case b)

$$\Pr\left\{x_i^{\min} \le x_i \le x_i^{\max}\right\} = \Pr\left\{\xi_j\left(x_i^{\max}, u, \tilde{\xi}\right) \le \xi_j \le \xi_j\left(x_i^{\min}, u, \tilde{\xi}\right)\right\}$$
$$= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \int_{\xi_j\left(x_i^{\max},u,\tilde{\xi}\right)}^{\xi_j\left(x_i^{\min},u,\tilde{\xi}\right)} \phi(\xi)\, d\xi_j d\left(\tilde{\xi}\right) =: h_i(u). \tag{8}$$

Assuming that such transformations exist for each chance constrained output variable x_i, $i \in I$, the resulting chance constraints are functions of u and are denoted by $h_i(u)$, $i \in I$.

The chance constraints have now the form

$$h_i(u) \ge \alpha_i,\ i \in I.$$

Note that it is not necessary to compute the analytic expressions for $\xi_j\left(x_i, u, \tilde{\xi}\right)$ from the system (2). It is sufficient to know theoretically that such a smooth strictly monotone function exists. Section 3 presents a simple monotony analysis based on the implicit function theorem.

Now the chance constrained optimization is transformed to the following nonlinear optimization problem (NLP)

$$(NLP) \qquad \min_u \{F_1(u) + \gamma F_2(u)\} \tag{9}$$

such that

$$h_i(u) - \alpha_i \ge 0,\ i \in I, \tag{10}$$
$$u \in \mathcal{U}. \tag{11}$$

The existence of a solution to this problem follows if its feasible set $\mathcal{M} := \{u \in \mathcal{U} \mid h_i(u) - \alpha_i \ge 0, i \in I\}$ is compact, which depends on the continuity of the functions h_i, $i \in I$. The latter can be shown to follow from the special form of h_i, $i \in I$, in (7) and (8). The NLP can be solved using IpOpt from Biegler and Wächter 2006 which implements an interior-point method and is capable of handling infeasible starting points. Since the solver is gradient oriented, for all functions, the function values and the first partial derivatives with respect to the controls u should be computed. However, all these function are given through the implicit expressions $x(u, \xi)$ or $\tilde{x}\left(x_i, u, \tilde{\xi}\right)$ and $\xi_j\left(x_i, u, \tilde{\xi}\right)$. According to (6), the gradients of F_1 and F_2 are computed using the chain rule with the partial derivatives $D_u x(u, \xi)$ of $x(u, \xi)$ with respect to u. Interchanging the integration and differentiation with respect to u is allowed because of the smoothness of f_1 and f_2, whenever the integrals

with inner differentiation exist, which is; e.g., satisfied for the normal distribution. It follows that

$$D_u F_1(u) = \int_{\mathbb{R}^p} \left[D_x f_1\left(x(u,\xi), u, \xi\right) D_u x(u,\xi) + D_u f_1\left(x(u,\xi), u, \xi\right) \right] \phi(\xi)\, d\xi.$$

For the functions of the objective the integration is done over the full dimension $\mathbb{R}^p$. For the chance constraints, because of variable upper and lower integration limits, the Leibnitz rule yields

$$D_u h_i(u) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \left[\phi(\xi)|_{\xi_j = \xi_j\left(x_i^{\max}, u, \tilde{\xi}\right)} D_u \xi_j\left(x_i^{\max}, u, \tilde{\xi}\right) - \phi(\xi)|_{\xi_j = \xi_j\left(x_i^{\min}, u, \tilde{\xi}\right)} D_u \xi_j\left(x_i^{\max}, u, \tilde{\xi}\right) \right] d\tilde{\xi}.$$

In this work, the computation of integrals is done by cubature formulas over sparse-grids (see Section 4). The grid points and the weights are a priori computed and remain all over the optimization the same. Furthermore, the determination of $D_u \xi_j\left(x_i^{\max}, u, \tilde{\xi}\right)$ and $D_u \xi_j\left(x_i^{\max}, u, \tilde{\xi}\right)$ can be accomplished by the implicit differentiation formula. This requires the solution of a set of nonlinear equations twice. Using a full-grid tensor-product of one dimensional quadrature rule increases the computation time exorbitantly (exponentially) with respect to the dimension of ξ. With sparse-grids integration techniques the approach can be effectively applied for the solution of chance constrained optimization problems. Section 4 gives a brief discussion on how sparse-grid integration techniques can be used to compute chance constraints.

3. A Method to Determine Monotonic Relations

To find monotonically related outputs and inputs is not an easy task. If a guess is known based on a practical knowledge of the process, then one can try to verify monotony relations using methods of simulation. However, in most cases it is easier to prove an algebraic sign than monotony relations by simulation (see Theorem 3.2 below). This section demonstrates that it is enough to test for algebraic signs to facilitate transformation of chance constraints.

The variables x_i and ξ_j are related through the nonlinear equation $g(x, u, \xi) = 0$. This indicates that each variable x_i depends on the variable u and the uncertain variables $\xi_1, \ldots, \xi_p$. Hence, a monotonicity relation of x_i with ξ_j can be termed as a *directional monotonicity*. Note also that the chance constraint on x_i, $Pr\{x_i^{min} \leq x_i \leq x_i^{max}\} \geq \alpha_i$, imposes a level of reliability for x_i to lie in the interval $[x_i^{min}, x_i^{max}]$, still with some probability for x_i to lie outside of this interval. For this purpose, a global implicit-function-theorem is needed to be used as a basis for monotonicity analysis between the x's and ξ's in the equation $g(x, u, \xi) = 0$ for a given u. Global implicit function theorems have been studied, for instance, in Rheinboldt 1969 and Sandberg 1981. In particular, the global implicit function theorem given by Sandberg 1981 will be used here.

Suppose there are open and convex sets $\mathcal{X} \subset \mathbb{R}^n, \mathcal{W} \subset \mathbb{R}^p$ such that x^{min}, $x^{max} \in \mathcal{X}$ and the uncertain variables ξ are from $\mathcal{W}$. The transformation of chance constraints facilitates

the calculation of $h(u)$ and $D_u h(u)$ for each fixed u. Thus, the relation between x's and ξ's should hold true for each fixed $u \in \mathcal{U}$.

THEOREM 3.1 *(Global Implicit Function Theorem, Sandberg 1981) Suppose $u \in \mathcal{U}$ be given and the function $g(\cdot, u, \cdot) : \mathcal{X} \times \mathcal{W} \to \mathbb{R}$ is continuously differentiable. The following four conditions hold true*

- *(a) for each $\xi_0 \in \mathcal{W}$ there is exactly one $x_0 \in \mathcal{X}$ such that $g(x_0, u, \xi_0) = 0$;*
- *(b) for each $(x_0, \xi_0) \in \mathcal{X} \times \mathcal{W}$, there are neighborhoods N_{x_0} and N_{ξ_0} of x_0 and ξ_0, respectively; and a continuous function $\varphi_u : N_{\xi_0} \to N_{x_0}$ such that, for each $\xi \in N_{\xi_0}$, $x = \varphi_u(\xi)$ and $(\varphi_u(\xi), u, \xi)$ is a unique solution of $g(x, u, \xi) = 0$;*
- *(c) $D_\xi g$ has a full rank for each $(x, \xi) \in \mathcal{X} \times \mathcal{W}$ with $g(x, u, \xi) = 0$;*
- *(d) $\det(D_x g) \neq 0$, for each $(x, \xi) \in \mathcal{X} \times \mathcal{W}$ with $g(x, u, \xi) = 0$;*

if and only if there is a unique continuously differentiable function $\varphi_u : \mathcal{W} \to \mathcal{X}$ such that $g(\varphi_u(\xi), u, \xi) = 0$ for all $\xi \in \mathcal{W}$.

In this theorem, the expressions $D_x g$ and $D_\xi g$ represent the Jacobian matrices of g with respect to x and ξ, respectively, and $\det(\cdot)$ represents the determinant operator on matrices. The global implicit function theorem says that: (c) and (d) are necessary for the existence of the above global implicit function φ_u. In general, it is not simple to check (a) and (b) to guarantee the existence of φ_u. However, if $g(x, u, \xi) = 0$ can be written as

$$\begin{aligned} x_1 &= F_1(u, \xi) \\ &\vdots \\ x_{k+1} &= F_k(x_1, x_2, \ldots, x_k, u, \xi), \end{aligned} \tag{12}$$

then (a) and (b) are trivially satisfied. Therefore, the considerations here are confined to checking conditions (c) and (d). The conditions (a) - (d) hold true in many model problems of chemical process engineering. In fact, in the related literature, these *regularity* conditions are taken for granted either implicitly or explicitly (see for instance Arellano-Garcia and Wozny 2009, Hong *et al.* 2006, Li *et al.* 2008, Wendt *et al.* 2002, etc.). Whenever (c) and (d) are satisfied, in this paper, then it is said that the function g is *regular*. Moreover, if g is continuously partially differentiable with respect to u, then the function $\varphi_u(\xi)$ is also continuously differentiable with respect to u.

The importance of the above theorem is to ensure a global functional relation between the $x's$ and ξ's, for a given u. After insuring the conditions (a) - (d) of the global implicit function theorem, then follows an investigation of monotonicity relations between a single coordinate x_i of x and ξ_j of ξ. Thus, the (local) implicit function theorem implies, for $\xi \in \mathcal{W}$ and fixed $u \in \mathcal{U}$, that

$$\frac{dg(\varphi_u(\xi), u, \xi)}{d\xi_j} = \frac{\partial g(\varphi_u(\xi), u, \xi)}{\partial x_i}\frac{\partial x_i}{\partial \xi_j} + \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{\partial g(\varphi_u(\xi), u, \xi)}{\partial x_k}\frac{\partial x_k}{\partial \xi_j} + \frac{\partial g(\varphi_u(\xi), u, \xi)}{\partial \xi_j} = 0, \tag{13}$$

where the column vectors $\frac{dg(\varphi_u(\xi), u, \xi)}{d\xi_j}$ denote the total partial derivative and $\frac{\partial g(\varphi_u(\xi), u, \xi)}{\partial \xi_j}$ the partial derivative with respect to ξ_j. Since the regularity of $D_x g(\varphi_u(\xi), u, \xi)$ is given

for all $\xi \in \mathcal{W}$ and $u \in \mathcal{U}$ with $g(x,u,\xi) = 0$, the linear system of (adjoint) equations

$$\eta^\top \frac{\partial g(\varphi_u(\xi),u,\xi)}{\partial x_k} = \delta_{ki}, k = 1,\ldots,n, \tag{14}$$

has a unique C^1-solution $\eta_u(\xi)$ on $\mathcal{W}$ and $u \in \mathcal{U}$; where $\delta_{ki} = 1$, for $k = i$, and $\delta_{ki} = 0$, otherwise. Multiplying both sides of (13) by $\eta_u(\xi)^\top$ we get that

$$\frac{\partial x_i}{\partial \xi_j} = -\eta_u(\xi)^\top \frac{\partial g(\varphi_u(\xi),u,\xi)}{\partial \xi_j} \text{ for } \xi \in \mathcal{W}, u \in \mathcal{U}$$

which yields necessary and sufficient conditions for strict monotonicity of x_i with respect to ξ_j. Hence, we have the following theorem.

THEOREM 3.2 *Suppose the assumptions for the global implicit function theorem hold true. Then x_i is globally monotonically related to ξ_j if either*

- *(a) $\xi_j \uparrow x_i$, i.e.,*

$$\eta_u(\xi)^\top \frac{\partial g(\varphi_u(\xi),u,\xi)}{\partial \xi_j} < 0,\ \xi \in \mathcal{W},\ u \in \mathcal{U};$$

or

- *(b) $\xi_j \downarrow x_i$, i.e.,*

$$\eta_u(\xi)^\top \frac{\partial g(\varphi_u(\xi),u,\xi)}{\partial \xi_j} > 0,\ \xi \in \mathcal{W},\ u \in \mathcal{U}.$$

In general, it is only needed to solve the linear equations (14) to determine η. Then the type of monotony depends on the algebraic sign of the scalar product

$$\eta^\top \left(\frac{\partial g}{\partial \xi_j}\right).$$

The intention here is not to determine all monotonic pairs of variables x_i and ξ_j. This a combinatorial task, specially for a very large system $g(x,u,\xi) = 0$. The major interest here is to monotonically relate a chance constrained variable x_i with some uncertain variable ξ_j to facilitate constraint transformations.

It needs to be remarked this presented monotonicity analysis is mainly applicable to small or medium-scale problems. Even so, sometimes monotonicity can be also obtained by considering only a part of a large system of nonlinear model equations (see, for instance, Section 5.2). Furthermore, this approach of testing monotonicity might face difficulties if it happens that $\eta^\top \left(\frac{\partial g}{\partial \xi_j}\right) = 0$. This means that, a global monotonicity relation between x_i and ξ_j, as given in Theorem 3.2 above, is not available. In such a case, either monotonicity must be investigated on subsets or it may be the cases that x_i is independent of ξ_j.

4. Computation of Chance Constraints using Sparse-Grid Techniques

The computation of the values and gradients of chance constraints and objective function requires the evaluation of multidimensional integrals. Traditionally multidimensional probability integrals are computed using Quasi-Monte-Carlo (QMC) methods. The QMC methods rely on the generation of *low-discrepancy* integration-points that are uniformly distributed in a bounded domain of integration. Thus, these methods require very large number of integration nodes which result in a very large number of function evaluation and lead to expensive CPU time (see Chen and Merothra 2007, Heiss and Winschel 2006, Gerstner and Griebel 1998, Schürer 2003). The recently introduced QMC based *sample average approximation* (SAA) techniques for the computation of chance constraints (cf. Kookos 2003, Pagnoncelli *et al.* 2009, Wang and Ahmed 2008, Royset and Polak 2004) not only use large sample size but also result in approximation of chance constraints which can be non-differentiable.

On the other hand, deterministic integration techniques for chance constraints such as orthogonal collocation (Arellano-Garcia and Wozny 2009) or recursive integration (Prekopa 1995) used by Arellano-Garcia 2007, Flemming *et al.* 2007, Li *et al.* 2008, lead to the full-grid integration. In full-grid techniques the number of integration nodes grows exponentially with the dimension of the uncertain variables. These techniques incur redundant computations (Mysovkikh 1968) and are known to be ineffective for integrals of higher dimensions, causing *the curse of dimensions*(Novak and Ritter 1997).

In 1963 Smolyak proposed a sparse-grid integration technique to overcome the curse of dimension. The sparse-grid integration technique (Bungartz and Griebel 2004, Gerstner and Griebel 1998) uses a skillful combination of one-dimensional quadrature rules to generate a sparse-grid cubature rule for higher dimensions. Usually, the Gauss-Kronord, Kronord-Patterson, Clenshaw-Kurtis or Gauss-Hermite quadrature rules, which are the most efficient quadrature rules in practice, are used as the underlying rules to construct integration nodes and weights. Since the probability integrals considered here are on unbounded domains, it is necessary to use a convenient sparse-grid technique for a higher dimensional integrals with a Gaussian weight function. Therefore, the computation of values and gradients of chance constraints was accomplished by an extensions of the sparse-grid integration techniques of Genz and Keister 1996. This cubature rule is constructed by a successive (Kronord like) extension of the Gauss-Hermite quadrature rule to higher dimensions. This sparse-grid technique yields good results even when the values of the integrands are given approximately, which is a conducive property for the integral here, since integrands are given implicitly through the solution of a set of nonlinear equations.

In general, the sparse-grid integration technique requires the integration weight function to be a product of one-dimensional weight functions. In our case, this can be achieved by the decorrelation of the normally distributed random variables ξ.

Section 4.1. discusses the transformation (decorrelation) of the normally distributed random variables ξ into standard normal distributed random variables. As a result, the integration weight function becomes a product of one-dimensional weight (probability density) functions. Accordingly, through a variable transformation, the integrals for the objective function and constraints are transformed into integrals having product weight functions. Section 4.2. presents the sparse-grid integration technique and briefly discusses its most salient features that make it valuable in the computation of probability integrals.

4.1. *Decorrelation and variable transformation*

For uncertain input variables ξ with normal distribution the probability density function is

$$\phi(\xi) = \frac{1}{(2\,\pi)^{p/2}\ |\Sigma|^{1/2}}\, e^{-\frac{1}{2}(\xi-\mu)^\top \Sigma^{-1}(\xi-\mu)},$$

with expectation μ and covariance matrix Σ with determinant $|\Sigma| > 0$. Using the Cholesky decomposition $\Sigma = L\,L^\top$, where L is a lower triangular matrix, the transformation $z = L^{-1}\,(\xi - \mu)$ defines the random variables z with a standard normal distribution. Define

$$z_j^{max} = e_j^\top L^{-1}\left((\xi_1,\ldots,\xi_{j-1},\xi_j(u,x_i^{max},\widetilde{\xi}),\xi_{j+1},\ldots,\xi_p)^\top - \mu\right), \tag{15}$$

$$z_j^{min} = e_j^\top L^{-1}\left((\xi_1,\ldots,\xi_{j-1},\xi_j(u,x_i^{min},\widetilde{\xi}),\xi_{j+1},\ldots,\xi_p)^\top - \mu\right), \tag{16}$$

where e_j is the $j-th$ unit vector in $\mathbb{R}^p$. Furthermore, in the expressions (15) and (16), the components of the vector $\widetilde{\xi}^\top = (\xi_1,\ldots,\xi_{j-1},\xi_{j+1},\ldots,\xi_p)$ can be written in terms of z by using $\xi_k = e_k^\top L z + \mu_k, k = 1,\ldots,j-1,j+1,\ldots,p$.

Recalling the integral formulation (7) and (8) of the chance constraints

$$h_i(u) = \int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}\int_{\xi_j(x_i^{min},u,\widetilde{\xi})}^{\xi_j(x_i^{max},u,\widetilde{\xi})}\phi(\xi_1,\ldots,\xi_p)d\xi_j d\widetilde{\xi}, i\in I$$

and using the above variable transformation, it follows that

$$h_i(u) = \int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}^{p-1}}e^{-\frac{1}{2}\widetilde{z}^\top\widetilde{z}}\left(\int_{z_j^{min}}^{z_j^{max}}\frac{1}{\sqrt{2\pi}}e^{-\frac{z_j^2}{2}}dz_j\right)d\widetilde{z}, i\in I, \tag{17}$$

where $\widetilde{z}^\top = (z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_p)$. The one-dimensional integral $\int_{z_j^{min}}^{z_j^{max}}\frac{1}{\sqrt{2\pi}}e^{-\frac{z_j^2}{2}}dz_j$ can be computed by using a library function for standard normal distributions. Moreover, the gradient of the chance constraints is given by

$$D_u h_i(u) = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}^p}\,e^{-\frac{1}{2}\widetilde{z}^\top\widetilde{z}}\left(e^{-\frac{1}{2}\left(z_j^{max}\right)^2}D_u z_j^{max} - e^{-\frac{1}{2}\left(z_j^{min}\right)^2}D_u z_j^{min}\right)d\widetilde{z}. \tag{18}$$

Therefore, to determine the values $h_i(u)$, $D_u h_i(u)$, $i\in I$ for a fixed u, it is needed to compute multidimensional integrals on $\mathbb{R}^{p-1}$ with the Gaussian weight function $W_{\mathbb{R}^{p-1}}:\mathbb{R}^{p-1}\to\mathbb{R}$ such that $W_{\mathbb{R}^{p-1}}(\widetilde{z}) = e^{-\frac{1}{2}\widetilde{z}^\top\widetilde{z}}$.

In addition, the values $F_1(u)$, $F_2(u)$ and gradients $D_u F_1(u)$, $D_u F_2(u)$ also require the computation of the multidimensional integrals. Using the variable transformation introduced above, these require the computation of a set of multidimensional integrals of the form

$$E[f(u,\cdot)] = \int_{\mathbb{R}^p} f(u,\xi)\phi(\xi)d\xi = \frac{1}{\sqrt{2\pi}^p}\int_{\mathbb{R}^p} f(u,z)e^{-\frac{1}{2}z^\top z}\,dz \tag{19}$$

for each fixed u, where in this case the integration weight function $W_{\mathbb{R}^p} : \mathbb{R}^p \to \mathbb{R}$ is given by $W_{\mathbb{R}^p}(z) = e^{-\frac{1}{2}z^\top z} = \Pi_{j=1}^p \exp^{\frac{1}{2}z_j^2}$.

4.2. *Sparse-grid integration techniques for probability integrals*

Now, referring to (17), (18) and (19), the integrals have the general form

$$E[f] = \int_\Omega f(z)W(z)dz,$$

where $\Omega \subset \mathbb{R}^p$ and $W : \mathbb{R}^p \to \mathbb{R}$ is a weight function. For these integrals the domain of integration $\Omega = \Pi_{j=1}^p \Omega_j$ and the weight function $W(z) = \Pi_{j=1}^p W_j(z_j)$ satisfy

$$\Omega_1 = \Omega_2 = \ldots = \Omega_p \text{ and } W_1 \equiv \ldots \equiv W_p$$

which are basic assumptions in sparse-grid techniques. Suppose there is a one-dimensional underlying quadrature rule on each of the sets $\Omega_j \subset \mathbb{R}, j = 1, \ldots, p$ with different *levels of accuracy or precision*[†]. Since all these sets are the same, the index j can be dropped and a sequence of sets of quadrature nodes (grid points) $X^{(i)}, X^{(i+1)}, \ldots$ is considered, such that $X^{(i)} \subset X^{(i+1)}, i = 1, 2, \ldots$ A multi-dimensional cubature rule constructed in this way is known to have a *nested or embedded sequence* of one-dimensional quadratures. That is, there is a sequence $\mathcal{Q}^{(i)}, \mathcal{Q}^{(i+1)}, \ldots$ of one-dimensional quadrature rules, on each of the sets $\Omega_j \subset \mathbb{R}$ with non-decreasing level of precision $N^{(i)}$ with $N^{(i)} \leq N^{(i+1)}, i = 1, 2, \ldots$. Usually, the underlying one-dimensional quadrature nodes are constructed based on the Gauss-Kronord, Kronord-Patterson and Gauss-Hermite rules, which are the most efficient quadrature rules in practice.

Now define a multi-index $\mathbf{i} = (i_1, i_2, \ldots, i_p) \in \mathbb{N}^p$ and its corresponding norm $\|\mathbf{i}\| = i_1 + i_2 + \ldots + i_p$.

DEFINITION 4.1 *(Tensor-Product Sparse-Grid Cubature Technique (Bungartz and Griebel 2004, Gerstner and Griebel 1998, Smolyak 1963)) The Smolyak or tensor-product sparse-grid cubature rule, with accuracy l, for approximation of an p-dimensional integral is given by*

$$\mathcal{A}(l,p)[f] = \sum_{l+1 \leq \|\boldsymbol{i}\| \leq p+l} (-1)^{l+p-\|\boldsymbol{i}\|} \binom{p-1}{\|\boldsymbol{i}\| - l - 1} \left(Q^{(i_1)} \otimes \ldots \otimes Q^{(i_p)}\right)[f]; \qquad (20)$$

where $\boldsymbol{i} = (i_1, i_2, \ldots, i_p)$ and

$$\left(Q^{(i_1)} \otimes \ldots \otimes Q^{(i_p)}\right)[f] = \sum_{k_{i_1}=1}^{N^{(i_1)}} \sum_{k_{i_2}=1}^{N^{(i_2)}} \ldots \sum_{k_{i_p}=1}^{N^{(i_p)}} \left(w_{k_{i_1}} \cdot w_{k_{i_2}} \cdot \ldots \cdot w_{k_{i_p}}\right) f(z_{k_{i_1}}, z_{k_{i_2}}, \ldots, z_{k_{i_p}}).$$

By definition $\mathcal{Q}_j^{(0)}[\cdot] = 0$. The sets of integration nodes $X^{(i_s)} = \left\{z_{k_{i_s}} \mid k_{i_s} = 1, \ldots N^{(i_s)}\right\}$ and their corresponding weights $w_{k_{i_s}}$, $k_{i_s} = 1, \ldots N^{(i_s)}$, are obtained from the set, say

[†]The level of accuracy or precision of a quadrature scheme is equal to the number of nodes used for the approximation of integrals.

Ω_j, and the weight function $W_j(\cdot)$ based on the underlying quadrature rule. For example, if $\Omega \subset \mathbb{R}^3$, $p = 3$, then the cubature scheme (20) with level of precision $l = 5$ will have the following tensor-product as one of its terms

$$\left(Q^{(1)} \otimes Q^{(3)} \otimes Q^{(2)}\right)[f] = \sum_{k_{i_1}=1}^{N^{(1)}} \sum_{k_{i_3}=1}^{N^{(3)}} \sum_{k_{i_2}=1}^{N^{(2)}} \left(w_{k_{i_1}} \cdot w_{k_{i_3}} \cdot w_{k_{i_2}}\right) f(z_{k_{i_1}}, z_{k_{i_3}}, z_{k_{i_2}})$$

with the multi-index $i = (1, 3, 2)$ such that $1 + 5 \leq \|i\| = 1 + 3 + 2 \leq 3 + 5$[‡]. Hence, the terms of (20) are linear combinations of tensor-products of one-dimensional grids of lower to higher levels of precisions[§].

Instead of (20), the *full-grid* technique, for approximate the integral $E[f]$, takes the form

$$\mathcal{Q}[f] = (Q_1 \otimes Q_2 \otimes \ldots \otimes Q_p)[f] = \sum_{k_1=1}^{N_1} \sum_{k_2=1}^{N_2} \ldots \sum_{k_p=1}^{N_p} \left(w_{k_1}^{(1)} \cdot \ldots \cdot w_{k_p}^{(p)}\right) f\left(z_{k_1}^{(1)}, z_{k_2}^{(2)}, \ldots, z_{k_p}^{(p)}\right). \tag{21}$$

The approximation of $E[f]$ by using the sparse-grid cubature rule $\mathcal{A}(l, p)[f]$ has the following properties:

- if the function f is polynomial with respect to z, then the approximation can be made to be exact by using appropriate number of grid points; i.e., $E[f] = \mathcal{A}(l, p)[f]$ (see Barthelmann et. al 2000, Heiss and Winschel 2006, Novak and Ritter 1999, Petras 2003). This is known as *polynomial exactness*.
- A theoretical error-analysis for the sparse-grid approximation of integrals has been already done by Wasilkowski and Woźniakowski 1995 (see also Bungartz and Griebel 2004, Gerstner and Griebel 1998) for bounded integration domains Ω; i.e., $\Omega = [a, b]^p$; so that

  $$|E[f] - \mathcal{A}(l, p)[f]| = O\left(2^{-l\,r} \cdot l^{(p-1)\cdot(r+1)}\right) \tag{22}$$

  for a function f with smoothness of degree r; i.e., $f \in \mathcal{C}_p^r$. Theoretical error analysis can be done by transforming the integral on $\Omega = \mathbb{R}^p$ to an integral on a bounded domain $\Omega = [a, b]^p$. (see Chen and Mehrotra 2007 for such a transformation).
- The number of grid-points in the formula (20) is equal to the cardinality of the set

  $$\mathcal{X}(l, p) := \bigcup_{l+1 \leq \|i\| \leq p+l} \left(X^{(i_1)} \times X^{(i_2)} \times \ldots \times X^{(i_2)}\right) \tag{23}$$

  which, for a fixed level of precision l and embedded one-dimensional quadrature rules, is given by (see Novak and Ritter 1999)

  $$|(\mathcal{X}(l, p))| \approx \frac{(2p)^l}{l!}.$$

[‡]For a more accessible discussion on sparse-grid techniques see Heiss and Winschel 2006.
[§]The level of precision of a quadrature scheme is equal to the number of nodes used for the approximation of integrals.

This means that the number of integration nodes in (20) grows only polynomially with respect to the dimension p of the uncertain variables.

In contrast, the number of integration nodes in the full-grid integration technique (21) depends exponentially on the dimension of the uncertain variables z. That is, the use of all the $N_1 \times N_2 \times \ldots \times N_p$ grid points leads to a full-grid integration rule. If $N_1 = N_2 = \ldots = N_p =: N$, then there are N^p integration nodes in the full-grid. This implies function evaluation grows exponentially with the dimension p.

Summarizing, sparse-gird techniques provide the following advantages:

- a few integration nodes are enough to yield a good approximation of integrals, thus, saving CPU time;
- the construction of the integration nodes and weights is done using the weight function $W(\cdot)$ and the integration domain Ω, independent of the function to be integrated. Thus, grid-points and weights can be computed only once and used repeatedly;
- the multivariate probability density function can be used as the integration weight function;
- functions which are polynomial with respect to the uncertain variable can be integrated exactly.

4.3. *A comparison between full and sparse grid*

To show the efficiency of the sparse-grid techniques the integral

$$\int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi}^n} \left(\sum_{i=1}^{n} x_i \right) e^{-\frac{x^\top x}{2}} dx \tag{24}$$

is evaluated using full- and sparse-grid methods with polynomial exactness equal to 9 and for the dimension n of the integral ranging between 1 and 20. The number of necessary grid points depending on the integral dimension n is displayed in Figure 1. Observe that, the number of grid points for the sparse- and full-grid points are almost equal for $n \leq 3$.

The integral (24) can be evaluated analytically and it is equal to zero. Both sparse- and full-grid method result in values which do not differ more than machine epsilon from the analytical solution. Their computation time is shown in Figure 2. It can be seen that the computation time for the full-grid method grows almost exponentially as a function of the dimension n of the integral, while this time for the sparse-grid technique remains almost constant. In addition, it can be seen that the computation time is almost the same for both cases, if the dimension if the integrals is low. In general, the sparse-grid method performs better for $n > 3$.

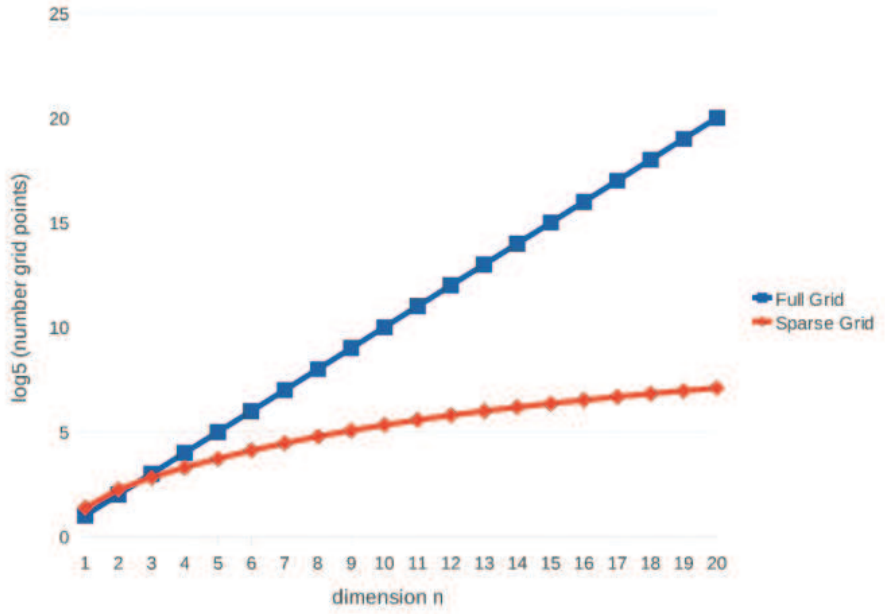


Figure 1. Number of grid points versus dimension

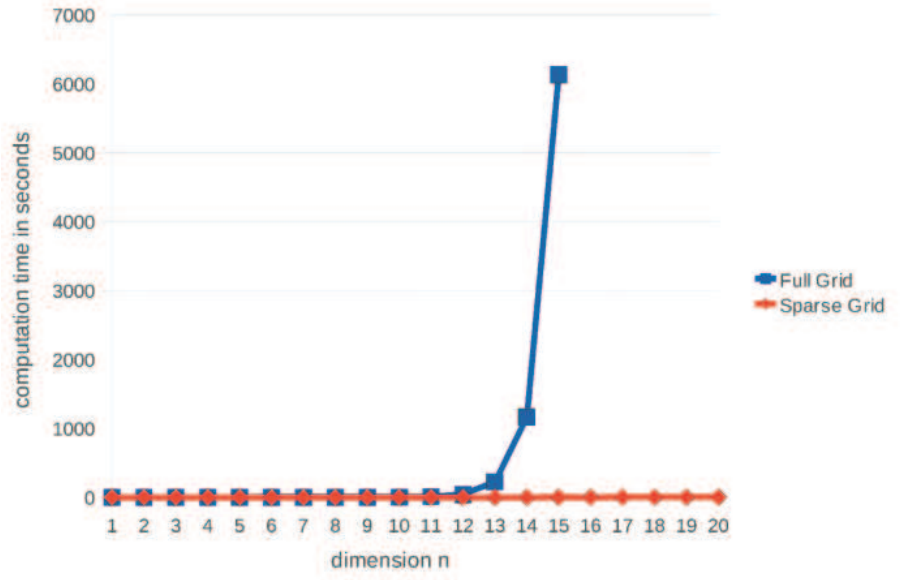


Figure 2. Computation time for integral (24)

5. Case Studies

5.1. *Variance Minimization under Chance Constraints*

In the CSTR (continuously stirred tank reactor) model Figure 3, the objective is to choose values for the design parameters that guarantee minimum fluctuation (variation)

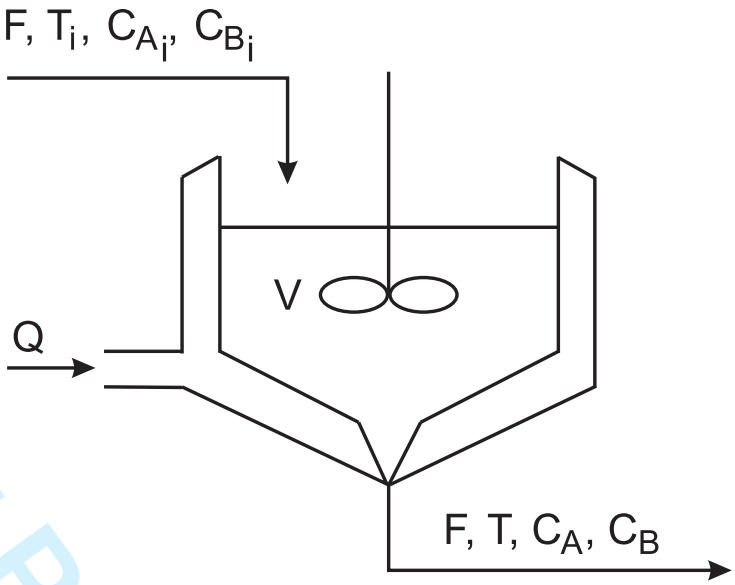


Figure 3. An Isothermal CSTR

of the production rate R_B of species B from the set-point $\overline{R}_B = 60\ mol/min$. Uncertain variables are assumed to be kinetic parameters of the chemical reactions and the inflow concentration.

As shown in Figure 3, the process involves a first order sequential reaction $A \longrightarrow B \longrightarrow C$ (see Kallagnaham and Diwekar 1997). The variables are listed as follows:

- F - feed flow rate;
- C_{Ai}, C_{Bi} - inflow concentrations of the species A and B, respectively;
- T_i - inflow temperature;
- Q - heat from the CSTR;
- V - bulk volume of the mixture in the CSTR;
- T - mixture temperature in the CSTR;
- C_A, C_B - concentrations of A and B in the mixture, respectively.

Furthermore, the kinetics of the reaction process is based on the following parameters for A and B, respectively:

- $-r_A$, $-r_B$ rates of consumption;
- E_A, E_B activation energies;
- H_{RA}, H_{RB} molar-heat constants (independent of temperature);
- k_A^0, k_B^0 pre-exponential Arrhenius constants;
- ρ, C_P density and specific heat constants.

In the reaction process, the rate of production of species B from species A is expected to have a set-point value of at least $\overline{R}_B = r_B V = 60\ \ moles/minutes$. The governing equations for the production of species B in the CSTR are steady state equations. The uncertain variables are supposed to have a known (joint) probability distribution. Hence, the problem can be formulated as a chance constrained nonlinear optimization problem to minimize the variance of R_B.

$$
\begin{aligned}
&\min \quad Var(R_B)\\
&\text{such that}\\
&Q - F\rho C_p\,(T - T_i) - V\,(r_A H_{RA} + r_B H_{RB}) = 0,\\
&C_A\left(1 + k_A^0 \exp(-E_A/RT)\,\tau\right) - C_{A_i} = 0,\\
&C_B\left(1 + k_B^0 \exp(-E_B/RT)\,\tau\right) - C_{B_i} - k_A^0 \exp(-E_A/RT)\,\tau\, C_A = 0,\\
&-r_A - k_A^0 \exp(-E_A/RT) C_A = 0,\\
&-r_B - k_B^0 \exp(-E_B/RT) C_B + k_A^0 \exp(-E_A/RT) C_A = 0,\\
&R_B - r_B V = 0,\\
&Pr\{R_B \geq 60\} \geq 0.9.
\end{aligned}
\tag{25}
$$

where

- $x = (C_A, C_B, r_A, r_B, R_B, T)$ is the vector of state variables;
- $\xi = \left(C_{A_i}, C_{B_i}, T_i, k_A^0, k_B^0\right)$ is the vector of uncertain input variables with a joint normal distribution;
- $u = (Q, V, F)$ is the vector of control variables in some non-negative bounded intervals;
- $\tau = V/F$ represents the average residence time of each species in the reactor.

The chance constraint $Pr\{R_B \geq 60\} \geq 0.9$ requires the process to produce species B at a rate of at least 60 mol/min with a reliability level of $\alpha = 90\%$ and at a minimum variance. Putting all equality constraints into a system of equations $g(x, u, \xi) = 0$, the Jacobian of g with respect to x is

$$
D_x g = \begin{pmatrix}
0 & 0 & -VH_{RA} & -VH_{RB} & 0 & -F\rho C_p\\
1 + k_A^0 \exp(-E_A/RT)\,\tau & 0 & 0 & 0 & 0 & \alpha\\
-k_A^0 \exp(-E_A/RT)\,\tau & 1 + k_B^0 \exp(-E_B/RT)\,\tau & 0 & 0 & 0 & \delta\\
-k_A^0 \exp(-E_A/RT) & 0 & -1 & 0 & 0 & \beta\\
k_A^0 \exp(-E_A/RT) & -k_B^0 \exp(-E_B/RT) & 0 & -1 & 0 & \gamma\\
0 & 0 & 0 & -V & 1 & 0
\end{pmatrix};
$$

with

$$
\begin{aligned}
&\alpha = \frac{E_A C_A}{RT^2}\left[k_A^0 \exp(-E_A/RT)\right]\,\tau, \quad \beta = -\frac{E_A C_A}{RT^2}\left[k_A^0 \exp(-E_A/RT)\right],\\
&\delta = \frac{E_B C_B}{RT^2}\left[k_B^0 \exp(-E_B/RT)\right]\tau - \frac{E_A C_A}{RT^2}\left[k_A^0 \exp(-E_A/RT)\right]\tau,\\
&\gamma = -\frac{E_B C_B}{RT^2}\left[k_B^0 \exp(-E_B/RT)\right] + \frac{E_A C_A}{RT^2}\left[k_A^0 \exp(-E_A/RT)\right],
\end{aligned}
$$

and this matrix is invertible, for each k_A^0, $k_B^0 \geq 0$ and $F\rho C_P \neq 0$. The Jacobian with

Table 1. Monotonic pairs of state and uncertain variables

	CA_i	CB_i	T_i	k_A^0	k_B^0
C_A	$\uparrow$	0	$\downarrow$	$\downarrow$	0
C_B	$\uparrow$	$\uparrow$	0	$\uparrow$	$\downarrow$
r_A	$\downarrow$	0	$\downarrow$	$\downarrow$	0
r_B	$\uparrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
R_B	$\uparrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
T	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$

respect to the uncertain variables ξ will be

$$D_\xi g = \begin{pmatrix} 0 & 0 & F\rho C_p & 0 & 0 \\ -1 & 0 & 0 & C_A \exp(-E_A/RT)\,\tau & 0 \\ 0 & -1 & 0 & -C_A \exp(-E_A/RT)\,\tau & C_B \exp(-E_B/RT)\,\tau \\ 0 & 0 & 0 & -C_A \exp(-E_A/RT) & 0 \\ 0 & 0 & 0 & C_A \exp(-E_A/RT) & -C_B \exp(-E_B/RT) \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Assuming that $F\rho C_P \neq 0, r_B \neq 0, C_A \neq 0$ and $C_B \neq 0$, the matrix $D_\xi g$ has full rank. Using symbolic computations, as explained in Section 3, the monotonicity relations among state and uncertain variables are summarized in Table 1. The table provides a complete list of monotonic pairs of state and uncertain variables. The upward arrows $\uparrow$ indicates that a state variable increases when the corresponding uncertain variables increases; conversely, a down arrow $\downarrow$ indicates that a state variable decreases as the corresponding uncertain variable increases. Nevertheless, for the transformation of the chance constraint $Pr\{R_B \geq 60\} \geq 0.9$, it is enough to use the relation $R_B \uparrow C_{A_i}$. This implies that

$$Pr\{R_B \geq R_B^{min}\} = Pr\{C_A \geq C_A^{min}\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{C_A^{min}}^{+\infty} \phi(\xi) d\xi.$$

Our numerical experiment uses the correlation matrix R, the mean and variance of the uncertain variable as given in Table 2. The numerical computations requires the evaluation of a set of integrals of dimension 5. These integrals are computed with a reasonable number of pre-computed grid-points based the cubature technique of Genz and Keister 1996. Table 3 gives the data used for the numerical experiments and CPU time for both sparse- and full-grid integration techniques. For this example (also for the case-study in Section 5.2) full- and sparse-grid integration techniques are constructed such that the numerical error is $|E[f] - \mathcal{A}(l,p)[f]| < 1e-3$ for a given function f. Note also that the computation of the integrals involves the solution of a large system of nonlinear equations to determine the implicitly defined variables. Therefore, most of the computational effort goes to the solution of the nonlinear equations. In particular, when using the sparse-grid technique, at each step of the optimization algorithm, it is required to solve a set of 6×993 nonlinear equations (for variance) and a set of 6×441 nonlinear equations (for chance constraints). Table 4 shows the optimization results for both the sparse- and full-grid techniques.

$$R = \begin{bmatrix} 1.0 & 0.1 & 0.1 & 0.0 & 0.0 \\ 0.1 & 1.0 & -0.8 & 0.0 & 0.0 \\ 0.1 & -0.8 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

Table 2. Mean and variance of the uncertain variables in the CSTR model

Uncertain variable	T_i	C_{A_i}	C_{B_i}	k_A^0	k_B^0
μ	300.0	3118.0	342.0	$8.4 * 10^5$	$7.6 * 10^4$
σ	1.0	155.9	17.1	100.0	100.0.

Table 3. Data for the chance constrained optimization of the CSTR

	Sparse-grid	Full-grid
Reliability level, i.e., α	0.9	0.9
Number of integration grid points	993 (variance), 441 (constraint)	7776 (variance), 1269 (constraint)
Overall computation time	4s	34s

Table 4. CSTR optimization results using sparse- and full-grid methods

	Q^*	V^*	F^*	$Var(R_B)$
Sparse-grid method	$-1.71e+06$	0.212277	0.0471224	12.1428
Full-grid method	$-1.71e+06$	0.212398	0.0471057	12.1297

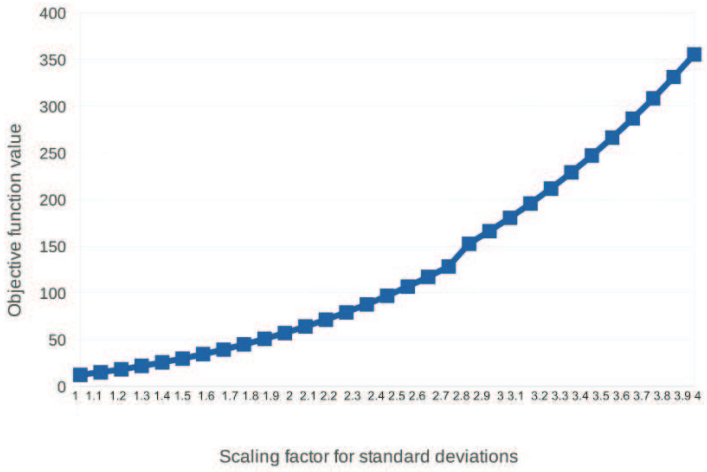


Figure 4. Variance of uncertainties vs. optimal objection values

Figure 4 shows how the variance of the rate of production (R_B) of species B varies when each of the uncertainties are multiplied by an identical factor (horizontal axis). It is easy to see that when the variance of the input uncertainties is increased, then the variance of the output will increase. This implies that higher uncertainties in the inputs result in higher uncertainties in the output.

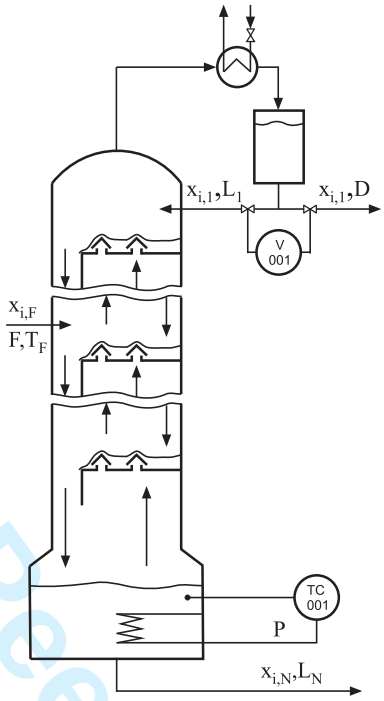


Figure 5. A distillation column

5.2. *A Nonlinear Chance Constrained Model for a Distillation Column*

The second example considers a chance constrained optimization of a distillation column with 20 separation stages (trays); i.e., $j = 1, \ldots, 20$ as shown in Figure 5. The numbering of the trays begins with the condenser as tray 1 and ends with the base as tray $N = 20$. A mixture of methanol and water is to be separated. The mixture is fed to the column at the tray number $j = 14$. The distillation process is highly influenced by a random variation of the feed and its concentration. Random variations can occur due to mixture supply variations. For instance, variations in mixture concentrations can happen when mixtures are supplied by different vendors. As a result there is some random variation in the feed concentration $x_{1,F}$, feed temperature T_F and the feed flow F. In this consideration all the 20 separation stages are supposed to have the same tray *efficiency* η. [‖] In practice, the value of η is determined experimentally; hence, η is uncertain in nature. Consequently, the variables η, F, T_F, $x_{1,F}$ can be considered as uncertain input variables denoted by $\xi^\top = (\eta, F, T_F, x_{1,F})$. In addition, these uncertain input variables are assumed to have a normal distribution with a given correlation matrix.

Products of the distillation process are collected at D and L_1. Uncertainties in the feed flow will propagate through the distillation process and will have an impact on the products (outputs). In the closed-loop framework, the output variables $x_{1,1}$ and D and the amount of energy P, required for the distillation process, are uncertain. This

[‖]It is also possible to consider all 20 separation stages each with different tray efficiency ratios.

problem was also considered by Flemming *et al.* 2007 with a closed-loop optimization under uncertainty.

Each distillation stage results in a set of nonlinear equilibrium equations which connect the input and output variables at the final stage. The model equations $g(x, u, \xi) = 0$ for the above distillation column are well known from the literature (see Flemming *et al.* 2007). Hence, the discussion here concentrates only on the chance constrained optimization of the process.

State variables in the Model equation

- D - the flow amount of the product collected from the accumulator
- $x_{i,j}$ - the concentration of component i in the liquid at stage j
- $y_{i,j}$ - the concentration of component i in the gas at stage j
- L_j, V_j - amount of liquid and vapor leaving stage j, resp.
- L_{j-1}, V_{j+1} - amount of liquid and vapor entering stage j, resp.
- H_j^L, H_j^V - liquid and vapor enthalpy, resp., at stage j
- $h_{i,j}^L$, $h_{i,j}^V$ - liquid and vapor enthalpy, resp., of component i at stage j
- P - flow of the reboiler

The temperature T_N in the reboiler and the reflux ratio r are control variables $u = (T_N, r)$. It means that the setpoints of two control loops will be optimized (see Flemming *et al.* 2007). All other variables are aggregated into the vector x of state variables. The objective is to keep the temperature T_N in the reboiler at a minimum, so that the energy consumption will be minimized, and at the same time product specifications are satisfied up to some pre-given reliability levels. Furthermore, bounds on the boiler temperature T_N are put to guarantee productivity and safety conditions, respectively. Hence, the optimization problem can be formulated as

$$(CCOPT) \quad \min_{T_N,\, r} T_N \tag{26}$$

such that

$$g(x, u, \xi) = 0, \tag{27}$$

$$T_N^{min} \le T_N \le T_N^{max}, \tag{28}$$

$$r^{min} \le r \le r^{max}, \tag{29}$$

$$Pr\{D \ge D^{min}\} \ge \alpha_D, \tag{30}$$

$$Pr\{x_{1,1} \ge x_{1,1}^{min}\} \ge \alpha_{x_{1,1}}, \tag{31}$$

$$Pr\{P \le P^{max}\} \ge \alpha_P. \tag{32}$$

The constraints (30), (31) and (32) specify separate chance constraints. Due to the uncertainty of D, $x_{1,1}$, P, the corresponding constraints are expected to be satisfied up to the pre-given reliability levels α_D, $\alpha_{x_{1,1}}$ and α_P. However, it is difficult to known the distribution of D, $x_{1,1}$ and P a priori. Therefore, it is required to transform the chance constraints on the output variables into chance constraints on the input uncertain variables ξ. This can be accomplished, as explained in Section 3, by using monotonicity relations. Using the energy balance equations

$$(L_{j-1} \cdot x_{1,j-1} - L_j \cdot x_{1,j}) + (V_{j+1} \cdot y_{1,j+1} - V_j \cdot y_{1,j}) + F_j \cdot x_{1,F} = 0, j = 2, \ldots, N-1,$$

with $F_j = 0$ until stage 14, and the component balance equation

$$V_2 \cdot y_{1,2} - (L_1 + D)x_{1,1} = 0,$$

it follows that

$$-D \cdot x_{1,1} + V_N \cdot y_N + F \cdot x_{1,F} = L_N \cdot x_{1,N}.$$

This expression relates the feed flow and the products directly. Consequently, the following strict monotonicity relations hold true

$$x_{1,F} \uparrow x_{1,1} \text{ and } x_{1,F} \uparrow D.$$

Similarly, it can be shown that $x_{1,F} \uparrow P$. This means that, as the feed concentration increases, the amount of energy required for the separation also increases. Therefore, the chance constraints can be transformed as follows

$$Pr\{D \geq D^{min}\} = Pr\{x_{1,F} \geq x_{1,F}^{min}(D^{min}, u, \widetilde{\xi})\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{x_{1,F}^{min}(D^{min},u,\widetilde{\xi})}^{+\infty} \phi(\xi)d\xi,$$

$$Pr\{x_{1,1} \geq x_{1,1}^{min}\} = Pr\{x_{1,F} \geq x_{1,F}^{min}(x_{1,1}^{min}, u, \widetilde{\xi})\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{x_{1,F}^{min}(x_{1,1}^{min},u,\widetilde{\xi})}^{+\infty} \phi(\xi)d\xi,$$

$$Pr\{P \leq P^{max}\} = Pr\{x_{1,F} \leq x_{1,F}^{max}(P^{max}, u, \widetilde{\xi})\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{x_{1,F}^{max}(P^{max},u,\widetilde{\xi})} \phi(\xi)d\xi,$$

where $\widetilde{\xi} = (\eta, F, T_F)$ and $\phi(\xi)$ is the joint normal distribution of the uncertain input variables. The values $x_{1,F}^{min}(D^{min}, u, \widetilde{\xi})$, $x_{1,F}^{min}(x_{1,1}^{min}, u, \widetilde{\xi})$ and $x_{1,F}^{max}(P^{max}, u, \widetilde{\xi})$ can be determined from the system of equation $g(x, u, \xi) = 0$ at each iteration step of the optimization problem.

The following bounds are used for the uncertain output variables in the chance constraints (30) - (32):

$$D_{min} = 6l/h; \quad x_{1,1}^{min} = 0.99mol/mol; \quad P^{max} = 6800W.$$

and the bounds on the control variables are

$$T_N^{min} = 90^0C, \quad T_N^{max} = 100^0C, \quad r^{min} = 1, \quad r^{max} = 6.$$

The uncertain input variables are assumed to have the correlation matrix R, mean μ and variance σ given in Table 5.

Most of the computational effort in the solution of the CCOPT problem goes to the solution of the nonlinear state equations. At each step of the optimization algorithm it is required to solve a block of 193 nonlinear equations 165 times, exploiting the special block structure that arises when using the sparse-grid points. For the solution of the nonlinear equations a two phase approach is used. The first phase calculates a solution to the model equations. From this solution a starting point is generated for the modified

$$R = \begin{bmatrix} 1.0 & 0.3 & 0.1 & 0.1 \\ 0.3 & 1.0 & 0.1 & 0.3 \\ 0.1 & 0.1 & 1.0 & 0.1 \\ 0.1 & 0.3 & 0.1 & 1.0 \end{bmatrix}$$

Table 5. Mean and variance of uncertain variables

Uncertain Variable	η	F	T_F	$x_{1,F}$
μ	0.70	$20.00 l/h$	$30.00\ ^0C$	0.20
σ	0.001	1.000	1.000	0.010

.

Table 6. Data for the chance constrained optimization of the distillation column

	Sparse-Grid	Full-Grid
Reliability level, i.e., $\alpha_D = \alpha_{x_{1,1}} = \alpha_P$	0.95	0.95
Number of Integration Grid Points	165	$6^3 = 216$
Overall computation time	$295s$	219s

Table 7. Optimization results for the distillation column using sparse- and full-grid methods

Solution $\rightarrow$	T_N^*	r^*	Optimal value of objective function
Sparse-grid method	96.24810305	1.412900723	96.24810305
Full-grid method	96.18671087	1.412420353	96.18671087

system

$$g((x_1, \ldots, x_{i-1}, x_i^{min}, x_{i+1}, \ldots, x_n), u, (\xi_1, \ldots, \xi_{j-1}, \xi_j^{min}, \xi_{j+1}, \ldots, \xi_p) = 0.$$

or

$$g((x_1, \ldots, x_{i-1}, x_i^{max}, x_{i+1}, \ldots, x_n), u, (\xi_1, \ldots, \xi_{j-1}, \xi_j^{max}, \xi_{j+1}, \ldots, \xi_p) = 0$$

correspondingly. The data and results of the optimization are given in Table 6 and 7, respectively. Note that all chance constraints are required to hold with a 95% reliability. The numerical experiments are done with an AMD Athlon 64 X2 dual core processor with 2.2 GHz per Core and 2 Gigabyte RAM. Using this platform the total computational time required is $295s$. For the same problem the computation time reported by Flemming *et al.* 2007 is 40.4 minutes which is obtained on a PC with Pentium IV, 3GHz processor and 1GB RAM. In this case, the more than 8 fold improvement of computation time is also attributed to the use of efficient coding and latest hardware platform. In fact, the 3-dimensional unbounded integrals in the chance constraints can be computed equally fast, even with less approximation error, using a full-grid technique. In fact, the comparative advantage of a sparse-grid techniques unfolds itself when the dimension of uncertain variables is high (see Figure 2).

In addition, one of the most important issues in chance constrained optimization is the question of feasibility of the problem versus the cost of the demand for higher reliability.

Table 8. Higher reliability incurs higher objective function values

α	T_N in ^0C
0.500	92.3323
0.550	92.5906
0.600	92.8604
0.650	93.1461
0.700	93.4603
0.750	93.752
0.800	94.1346
0.850	94.5973
0.900	95.2168
0.910	95.3821
0.920	95.5303
0.930	95.7174
0.940	95.9304
0.950	96.2114
0.951	96.2331
0.952	96.2443

This can be studied by investigating the parametric problem

$$\varphi(\alpha) = \min_u \left(F_1(u) + \gamma F_2(u)\right) \tag{33}$$

$$\text{such that} \quad u \in \mathcal{M}(\alpha) = \{u \in \mathcal{U} \mid h_i(u) \geq \alpha_i, i \in I\}, \tag{34}$$

where $\alpha \in (0,1)^{n_1}$ and $n_1 = |I|$. Note that if $\alpha^{(1)} \leq \alpha^{(2)}$ *lexicographically*, then $\mathcal{M}\left(\alpha^{(1)}\right) \supset \mathcal{M}\left(\alpha^{(2)}\right)$ and $\varphi\left(\alpha^{(1)}\right) \leq \varphi\left(\alpha^{(2)}\right)$. In other words, a continuous increasing in the level of reliability increases the minimum value of the objective function. Table 8 shows this fact for CCOPT. At the same time, a continuous increase in the reliability level α results in a shrink on the feasible set. This leads to a maximum reliability level $\alpha^* \in (0,1)$ beyond which it is not possible to hold the required product specifications. That means, there is a maximum α^* for which the problem remains feasible and $\mathcal{M}(\alpha) = \emptyset$, for $\alpha > \alpha^*$. The numerical experiment done on CCOPT justify theses issues. The problem becomes infeasible for $\alpha > \alpha^* = 0.953$.

6. Conclusions

The numerical solution of nonlinear optimization problems subject to chance constraints presents a difficult task. Specifically, this is the case when the chance constrained state (output) variables and the uncertain input variables have nonlinear relations. To circumvent this difficulty chance constraints can be transformed from the space of state (output) variables into chance constraints in the space of uncertain variables, so that the computation of constraint values and their gradients is possible. This idea of transformation of constraints is facilitated if there are strict monotonicity relations between chance constrained state variable and an uncertain input variable. Consequently, this paper gives conditions that guarantee the required monotonicity relations between state (output) and uncertain (input) variable pairs, which lead to an

easy-to-use monotonicity analysis method. This approach can be applied for small to medium-scale problems with respect to the number of uncertain variables. Sometimes monotonicity can also be obtained by investigating part of a large system. If global monotonicity can be found using a small part of a large system of equations, it also holds true for the whole system.

After transformation, the numerical solution of the resulting chance constrained optimization problem requires the evaluation of multi-dimensional integrals. This study indicates how to compute values and gradients of chance constraints by using tensor-product techniques on sparse-grids. With this approach it is possible to save computation time by several fold. However, it still remains to test the potential of the approach for chance constrained optimization problem with a large number of uncertain variables.

The case-studies considered are limited to steady-state optimization problems. But, transformation of chance constraints using monotonicity and sparse-grid integration techniques can still be applied for nonlinear dynamic chance constrained optimization problems from process engineering .

In the case-studies considered above, the chance constrained state variables are monotonically related with at least one uncertain input variable. But this may not be the case, in general. If there is no such uncertain input variable which is globally monotonically related with the given state variable, the transformation of the chance constraint is not straightforward. In several cases, such monotonicity relations can be shown to hold on intervals. This issue will be addressed in our future work.

Acknowledgements

The authors would like to express their indebtedness to the Deutsche Forschungsgemeinschaft (DFG) for supporting this research through the grant No. LI806/8-2.

References

(1) Arellano-Garcia, H., Wendt, M., Barz, T., Wozny, G., 2007. Closed-loop dynamic optimization under probabilistic output-constraints. *In*: R. Findeisen, F. Allgöwer and L. T. Biegler, eds. *Assessment and Future Directions of nonlinear model predictive control*; Springer-Verlag, 305 – 315.

(2) Arellano-Garcia, H., Wozny, G., 2009. Chance constrained optimization of process systems under uncertainity: I. Strict monotonicity. *Comput. Chem. Eng.* 33, 1568 – 1583.

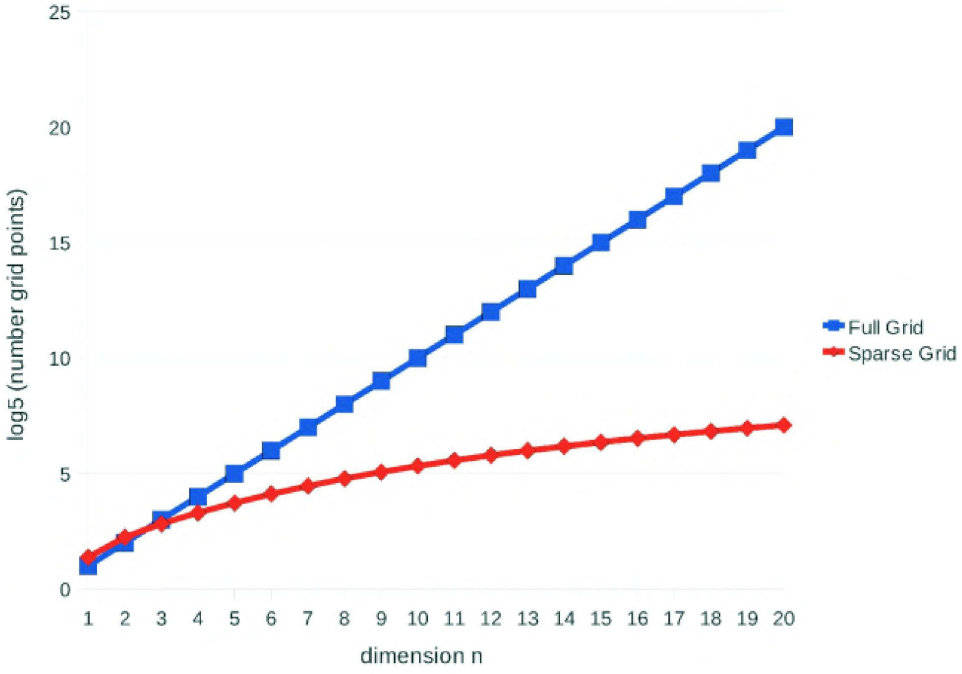
(3) Barthelmann, V., Novak, E., Ritter, K., 2000. High dimensional polynomial interpolation on sparse grids. *Adv. in Comp. Math.*, 12, 273 – 288.

(4) Ben-Tal, A., El Ghaoui, L., Nemirovski, A., 2009. *Robust optimization.* Princeton University Press.

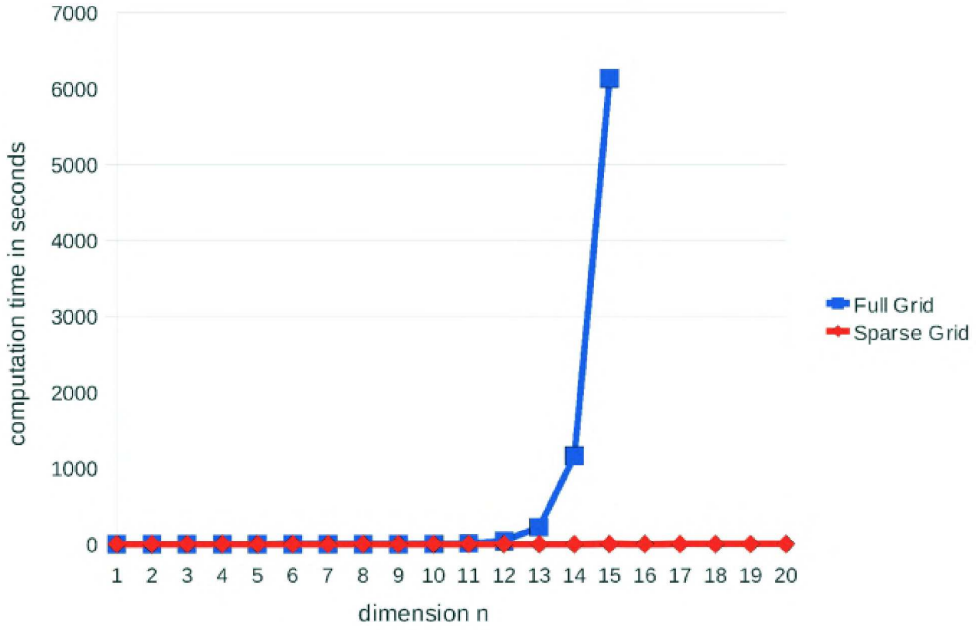
(5) Bungartz, H.-J., Griebel, M., 2004. Sparse grids. *Acta Numerica*, 13, 147 - 269.

(6) Califore, G., Campi, M. C., 2006. Uncertain convex programs: randomized solutions and confidence levels. *Math. Prog., Ser A.*, 102, 25 – 46.
(7) Charnes, A., Cooper, W., 1959. Chance-constrained programming. *Management Science*, 6, pp 73 – 79.
(8) Chen, M., Merothra, S., 2007. Epi-convergent scenario generation method for stochastic problems via sparse grid. Technical report 2007-08 2007; Department of Industrial Engineering and Management Sciences, Robert R. McCormick School of Engineering, Northwesetern University, Evanston, Illinois 60208.
(9) Diwekar, U. M., Kalagnanam, J. R., 1997. Efficient sampling technique for optimization under uncertainty. *AIChE Journal*, *43*, 440 – 447.
(10) Flemming Th., Bartl, M., Li, P., 2007. Set-point optimization for closed-loop control systems under uncertainty. *Ind. Eng. Chem. Res.*, 46, 4930 – 4942.
(11) Genz, A., Keister, B. D., 1996. Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weights. *SIAM J. Numer. Anal.*, 71, 299 – 309.
(12) Gerstner, T., Griebel, M, 1998. Numerical integration on sparse grids. *Numerical Algorithms*, 18, 209 - 232.
(13) Heiss, F., Winschel, V., 2006. Esitimation with numerical integration on sparse grids. *Münchner Wirtschaftswissenschaftliche Beiträge(VWL)*, 2006-15.
(14) Henrion, R., Li, P., Möller, A., Steinbach, M. C., Wendt, M., Wonzy G., 2001. Stochastic optimization for operating chemical processes under uncertainity. *In*: M. Gröthschel, S. Krumke and J. Rambau, eds. *Online Optimization of Large Scale Systems*, ; Springer-Verlag, 455 – 476.
(15) Henrion, R., Möller, A., 2003. Optimization of a continuous distillation process under random inflow rate. *Comput. Math. Appl.*, 45, 247 – 262.
(16) Hinrichs, A., Novak, E., 2007. Cubuature formulas for symmetric measures in high dimensions with few points. *Math. Comput.*, 76, 1357-1372.
(17) Hong, W.R., Wang, S., Li, P., Wozny, G., Biegler, L.T., 2006. A quasi-sequential approach to large-scale dynamic optimization problems. *AIChE Journal*, 52, 255 – 268.
(18) Kallagnaham, J. R., Diwekar, U. M., 1997. An efficient sampling technique for off-line quality control. *Technometrics*, 39, 308 – 319.
(19) Kookos, I. K., 2003. Optimal operation of batch processes under uncertainity: A Monte Carlo simulation deterministic optimization approach. *Ind. Eng. Chem. Res.*, 42, 6815 - 6822.
(20) Kropat, E., Weber, G.-W., Pedamallu, C.S., 2010. Regulatory networks under ellipsoidal uncertainty - optimization theory and dynamical systems, to appear in the book on Data Mining, D. Holmes, ed., Springer.
(21) Mysovkikh, I. P., 1968. On the construction of cubature formulas with the smallest number of nodes. *Soviet Math. Dokl.*, 9, 277 – 280.
(22) Li, P., 2007. *Prozessoptimierung unter* U*nsicherheiten*. Oldenbourg-Verlag.
(23) Li, P., Arellano-Garcia, H., Wozny, G., 2008. Chance constrained programming approach to process optimization under uncertainity. *Comput. Chem. Eng.*, 32, 24 – 45.
(24) Li, P., Wendt, M., Arellano-Garcia, H., Wozny, H., 2002. Optimal operation of distillation process under uncertain inflows accumulated in a feed tank. *AIChE J.*, 48, 1198 – 1211.
(25) Nemirovski, A., Shapiro, A., 2006. Convex approximation of chance constrainained programs. *SIAM J. Optim.*, 969 – 996.

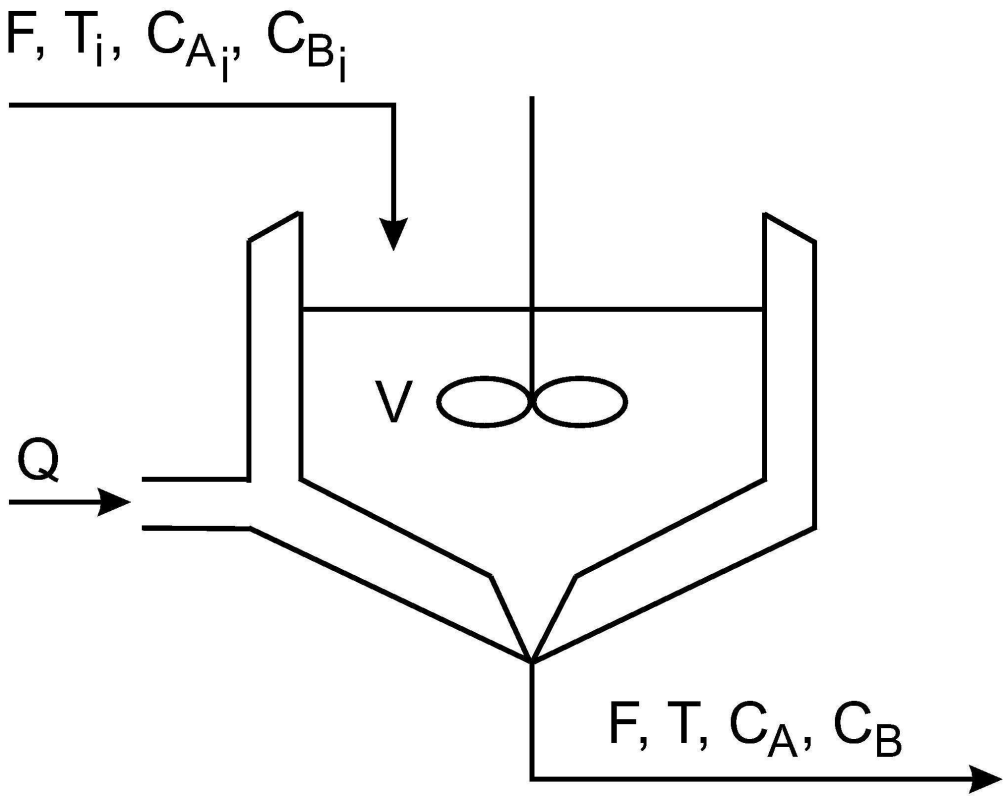
(26) Novak, E., Ritter, K., 1999. Simple cubature formulas with high polynomial exactness. *Constr. Approx.*, 15, 499-522.
(27) Novak, E., Ritter, K., 1997. The curse of dimension and a universal and a universal method for numerical integration. In Multivariate Approximation and Splines, Nürnberger, G., Schmidt, J. W., Walz, G.(eds.); International Series in Numerical Mathematics, Birkhäuser, Basel, V. 125, 177 – 188.
(28) Özmen, A., Weber, G.-W., Batmaz, İ., Kropat, E., 2010. RCMARS: Robustification of CMARS with different scenarios under polyhedral uncertainty set. *Preprint No. 1711, Institute of Applied Mathematics, METU, Turkey.*
(29) Pagnoncelli, B. K., Ahmed, S., Shapiro, A., 2009. Sample avergage approximation method for chance constrained programming: theory and applications. *J. Optim. Theory Appl.*, 142, 399 - 416.
(30) Petras, K., 2003. Smolyak cubature of given polynomial degree with few nodes for increasing dimension. *Numer. Math.*, 93, 729-753.
(31) Prekopa, A., 1995. *Stochastic programming.* Kluwer Academic Publishers.
(32) Rheinboldt, W. C., 1969. Local mapping relations and global implicit function theorems. *Trans. Amer. Math. Soc.*, 138, 183 – 198.
(33) Rockafellar, R. T., Uryasev, S., 200. Optimization of conditional value-at-risk. *Journal of Risk*, 2, 21 – 41.
(34) Royset, J. O., Polak, E., 2004. Implementable algorithm for stochastic optimization using sample average approximation. *JOTA*, 122, 157 – 184.
(35) Royset, J. O., Kiureghian, A. D., Polak, E., 2006. Optimal design with probabilistic objective and constraints. *J. of Eng. Mech.* (ASCE), 32, 107 – 118.
(36) Sandberg, I. W., 1981. Global implicit function theorems. *IEEE Trans. Circuits Syst.*, 28, 145 – 149.
(37) Schürer, R. A., 2003. A comparison between (quasi-)Monte Carlo and cubature based methods for solving high-dimensional integration problems. *Math. Comput. Simulation*, 62, 509 – 513.
(38) Smolyak, S. A., 1963. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Math. Dokl.*, 4, 240 – 243.
(39) Straub, D. A., Grossmann, I. E., 1993. Design optimization of stochastic flexibility. *Comput. Chem. Eng.*, *17*, pp 339 – 354.
(40) Wang, W., Ahmed, S., 2008. Sample average approximation of expected value constrained stochastic programs. *Operations Research Letters*, 36, 515-519.
(41) Wächter A., Biegler, L. T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Prog.*, 106, 25 – 57.
(42) Wasilkowski, G. W., Woźniakowski, H., 1995. Explicit cost bounds of algorithms for multivariate tensor product problems. *J. Complexity*, 11, 1 – 56.
(43) Weber, G.W., Branzei, R., Alparslan Gok, S.Z., 2010. On cooperative elipsoidal games. *Preprint No. 166, Institute of Applied Mathematics, METU, Turkey.*
(44) Wendt, M., Li, P., Wozny, G., 2002. Nonlinear chance-constrained process optimization under uncertainity. *Ind. Eng. Chem. Res.*, 41, 3621 – 3629.



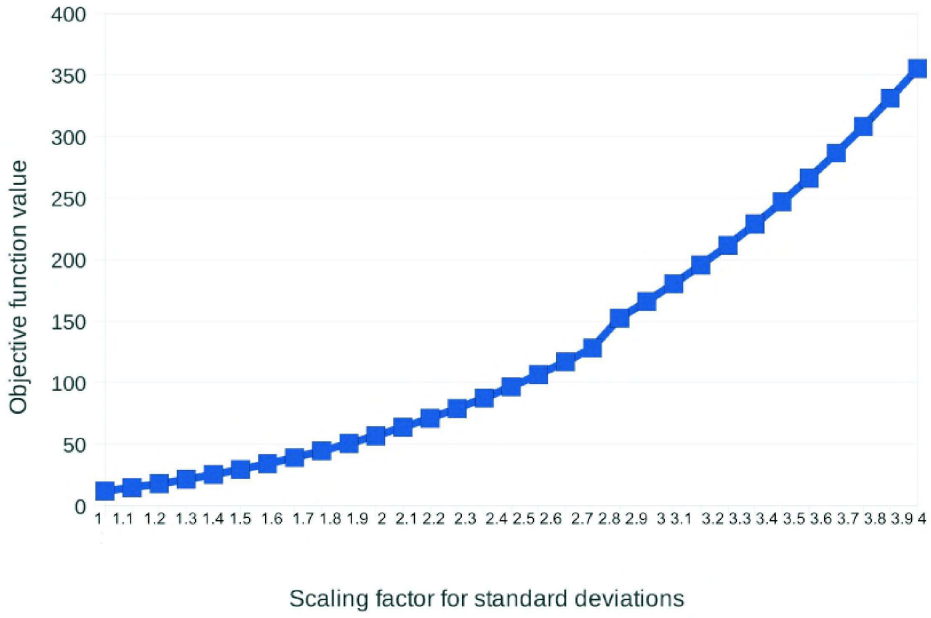
Number of grid points versus dimension
100x70mm (600 x 600 DPI)



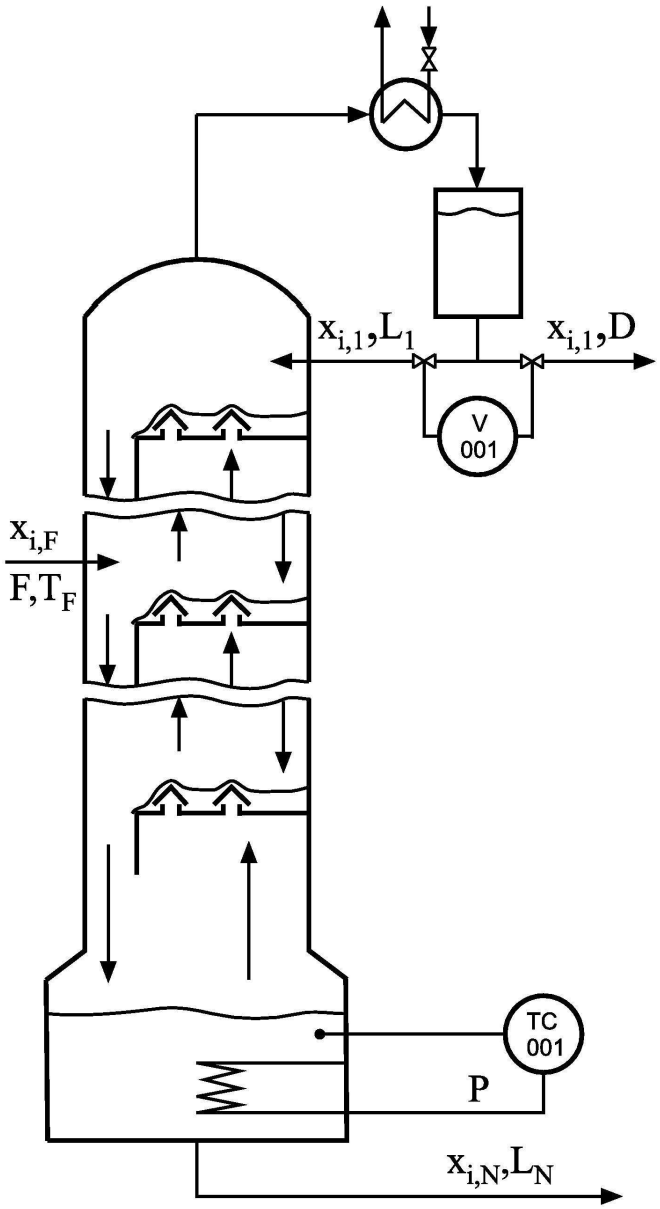
Computation time for integral
168x109mm (600 x 600 DPI)



85x67mm (600 x 600 DPI)



Variance of uncertainties vs. optimal objection values
100x68mm (600 x 600 DPI)



59x108mm (600 x 600 DPI)