

# Quality costs and robustness criteria in chemical process design optimization

Fernando P. Bernardo <sup>a</sup>, Efstratios N. Pistikopoulos <sup>a,\*</sup>, Pedro M. Saraiva <sup>b</sup>

<sup>a</sup> Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College, London SW7 2BY, UK

<sup>b</sup> Department of Chemical Engineering, University of Coimbra, Coimbra, Portugal

Received 5 November 1999; received in revised form 30 June 2000; accepted 13 September 2000

## Abstract

The identification and incorporation of quality costs and robustness criteria is becoming a critical issue while addressing chemical process design problems under uncertainty. This article presents a systematic design framework that includes Taguchi loss functions and other robustness criteria within a single-level stochastic optimization formulation, with expected values in the presence of uncertainty being estimated by an efficient cubature technique. The solution obtained defines an optimal design, together with a robust operating policy that maximizes average process performance. Two process engineering examples (synthesis and design of a separation system and design of a reactor and heat exchanger plant) illustrate the potential of the proposed design framework. Different quality cost models and robustness criteria are considered, and their influence in the nature and location of best designs systematically studied. This analysis reinforces the need for carefully considering/addressing process quality and robustness related criteria while performing chemical process plant design. © 2001 Elsevier Science Ltd. All rights reserved.

*Keywords:* Robust design; Uncertainty; Quality engineering; Stochastic optimization

## 1. Introduction

At the design stage of a process system, decisions have to be made in the presence of high uncertainty level. For instance, equipment configuration and dimensions, and their operating conditions have to be decided on the basis of an available process model, whose parameters may be uncertain, and on external information, which commonly exhibits a random behavior.

Taguchi (1986) approach to quality engineering provides a robust design strategy aimed at determining nominal settings for the design variables (parameter design) and their associated tolerance limits (tolerance design), in order to reduce process sensitivity to uncertainty. The traditional Taguchi methodology, which is based on running statistically designed experiments on a process prototype, is not, however, directly applicable to early process system design. On the other hand,

process design and optimization under uncertainty (Pistikopoulos, 1995) offers a systematic optimization-based vehicle to address process system design issues in the presence of uncertainty. However, in most such optimization studies, robustness issues are not explicitly considered, although attempts to link robustness/quality engineering aspects to stochastic process design optimization have begun to appear in the literature (Straub & Grossmann, 1993; Diwekar & Rubin, 1994; Bernardo & Saraiva, 1998; Samsatli, Papageorgiou & Shah, 1998; Georgiadis & Pistikopoulos, 1999).

In this article, we introduce a systematic design framework for process quality that embeds Taguchi's method and other robustness criteria within a stochastic optimization formulation. Quality related constraints are relaxed and process robustness is guaranteed through the explicit incorporation of robustness criteria in the optimization formulation, such as penalty terms in the objective function and/or limits on the variance of quality variables. With the relaxation strategy mentioned above, feasibility tests are not required, and thus the objective function expected value is obtained through integration over the entire uncer-

\* Corresponding author. Tel.: +44-20-75946620; fax: +44-20-75946606.

E-mail address: e.pistikopoulos@ic.ac.uk (E.N. Pistikopoulos).

tainty space. As a consequence, the original two-stage optimization problem was transformed into a single-level stochastic optimization formulation.

The computation of multiple integrals over the uncertainty space is a critical numerical issue in stochastic process design. Integration techniques applied so far to this kind of problems include Gaussian quadrature and stratified sampling techniques. In the first case, the number of points where the integrand function need to be evaluated increases exponentially with the integral dimension (the number of uncertain parameters), making the problem untreatable for a reasonably large number of uncertain parameters (Pistikopoulos & Ierapetritou, 1995). On the other hand, sampling techniques may be computationally more attractive, since the number of points required does not necessarily increase with the number of uncertain parameters. However, even the most efficient sampling techniques, such as the Hammersley sequence sampling (HSS) introduced by Diwekar and Kalagnanam (1997a,b) require some hundreds of points to achieve a reasonable accuracy.

At the numerical level, the present work employs a cubature technique (Stroud, 1971) to compute the multiple integrals involved in the stochastic problem formulation. When all uncertain parameters are normally distributed, a specialized cubature formula is applied, reducing significantly the number of points needed when compared with other integration strategies, such as product Gauss rules or efficient sampling techniques (Bernardo & Saraiva, 1998).

The remaining parts of this paper are structured as follows. First, the proposed mathematical problem formulation, addressing process quality, is developed, based upon a two-stage stochastic optimization framework. Next, robustness criteria and their implementation are described in more detail. Finally, the proposed formulation is illustrated through two chemical process design examples (synthesis and design of a separation system and design of a reactor and heat exchanger system).

## 2. Stochastic formulation for process quality

The problem of process design under uncertainty can be represented mathematically according to the following general formulation:

$$\begin{aligned} & \max_{d,z,x,y} f(d, z, x, \theta) \\ \text{s.t.} \quad & h(d, z, x, \theta) = 0 \\ & g(d, z, x, \theta) \leq 0 \\ & g_q(y, y^*) \leq 0 \\ & d \in D, \quad z \in Z, \quad x \in X, \quad y \in Y, \quad \theta \in \Theta \end{aligned} \quad (1)$$

where  $d$ ,  $z$  and  $x$  are the vectors of design, control and state variables, respectively;  $y$  stands for the vector of quality-related variables (usually a simple function of state and control variables), with desired values  $y^*$ , and  $\theta$  represents the vector of uncertain parameters over the domain  $\Theta$ . The performance metric to be optimized is defined by the scalar function  $f$ , the model equalities  $h$  refer to process model equations (heat and mass balances, equilibrium relationships, etc.), the inequalities  $g_q$  reflect quality constraints and the inequalities  $g$  other types of constraints.

Several approaches have been suggested to formulate and solve problem (Eq. (1)), differing in how uncertainty is handled, how an operating policy is selected in the face of uncertainty and also in the design objective considered. In the following paragraphs, we will briefly discuss some of these different approaches, while at the same time trying to clarify the assumptions that lie behind our approach.

### 2.1. Uncertainty formulation

With respect to the way uncertainty is handled, three different approaches can be stated, (i) scenario-based approach (Grossmann & Sargent, 1978; Halemane & Grossmann, 1983; Varvarezos, Grossmann & Biegler, 1992; Ahmed & Sahinidis, 1998); (ii) stochastic approach (among others, Pistikopoulos & Ierapetritou, 1995; Bernardo & Saraiva, 1998) and (iii) parametric approach (Acevedo & Pistikopoulos, 1996, 1997; Pertsinidis, Grossmann & McRae, 1998).

In the scenario-based approach, the uncertainty domain  $\Theta$  was approximated by a set of discrete scenarios (periods) with a given probability and, as a result, the original problem (Eq. (1)) was transformed into a multiperiod optimization problem. In the stochastic approach, uncertain parameters were assumed to follow a given joint probability density function (PDF) and an expected average criterion was optimized via a stochastic optimization strategy. In the parametric approach, no assumption was made about the uncertainty model and problem (Eq. (1)) was solved parametrically in the space of the uncertain parameters. The resulting solution was itself a function of the uncertain parameters realizations, providing a full map of the optimal decisions over the uncertainty domain considered. The formulation presented in this article follows a *stochastic approach*, defining the uncertainty domain  $\Theta$  as a probabilistic space of the form  $\Theta = \{\theta: \theta \in j(\theta)\}$ , where  $j$  is a joint PDF for the random vector of uncertain parameters  $\theta$ , which may be related or mutually independent.

### 2.2. Operating policy

The selection of an operating policy in the presence of uncertainty is another issue to be considered when

addressing problem (Eq. (1)). Several approaches are possible, the most conservative one being to assume fixed setpoints for the control variables, regardless of operation information that is available. Under this perspective, both design and control variables are treated as ‘here-and-now’ decisions that remain constant during process operation. The work by Diwekar and Rubin (1994), Bernardo and Saraiva (1998) can be included in this category, but with the additional feature that control variables are taken as random variables following a certain PDF.

The two-stage approach, on the other hand, relies on the assumption of perfect information and control: the design variables are considered ‘here-and-now’ decisions, while the control variables are ‘wait-and-see’ decisions, whose setpoints can be optimally adjusted to uncertainty during operation, according to observed values (among others, Grossmann, Halemane & Swaney, 1983; Pistikopoulos & Ierapetritou, 1995).

While the ‘here-and-now’ approach is conservative, the two-stage approach is rather optimistic. The difference between them quantifies the value of perfect information regarding uncertainty (Ierapetritou, Pistikopoulos & Floudas, 1996). A more realistic approach should fall somewhere between these two extreme approaches, selecting an operating policy that makes use of plant data through available supervisory control systems. The work by Bhatia and Biegler (1997) points in this direction by considering that available information about uncertainty is subject to an assumed feedback control law relating state and control variables. Our formulation in this article is based on the *two-stage approach*, but in the future we intend to define a more effective formulation, taking into account which information is (or should be) available during plant operation.

### 2.3. Design objective

Recent literature on process design under uncertainty has mainly focused on the concept of stochastic flexibility, defined as the probability of feasible process operation under the presence of uncertainty (Pistikopoulos & Mazzuchi, 1990; Straub & Grossmann, 1990). The design objective considered is either to achieve an optimal degree of feasibility/flexibility subject to a maximum investment cost (Straub & Grossmann, 1993), or to maximize an expected economic objective, while simultaneously measuring design feasibility (Pistikopoulos & Ierapetritou, 1995). In our formulation, the design objective integrates process *profitability* and *quality*, incorporating Taguchi loss functions and other robustness criteria in a stochastic formulation. By process quality we refer here to a wide concept that may cover product quality, environmental issues, safety, risk, hazards or any other relevant concerns, through an overall integrated process analysis.

Before discussing the incorporation of robustness/quality issues in our formulation, we will look into the two-stage stochastic formulation, assuming average process performance as the design objective.

$$\text{Design stage: } \max_d E_{R_n(d)}\{f'(d, \theta)\}$$

$$d \in D, \theta \in \Theta$$

$$\text{Operating stage: } f'(d, \theta) = \max_{z,x,y} f(d, z, x, \theta)$$

$$\text{s.t. } h(d, z, x, \theta) = 0$$

$$g(d, z, x, \theta) \leq 0$$

$$g_q(y, y^*) \leq 0$$

$$z \in Z, \quad x \in X, \quad y \in Y \quad (2)$$

The design variables are selected in the first (design) stage and remain fixed in the second (operating) stage, where the goal is to determine an optimal vector of control variables  $z$  for each possible realisation of the uncertain parameters  $\theta$  lying within the corresponding feasible region  $R_n(d)$ :

$$R_n(d) = \{\theta \in \Theta \mid \exists(z, y, z): h(d, z, x, \theta) = 0 \\ \wedge g(d, z, x, \theta) \leq 0 \wedge g_q(y, y^*) \leq 0\} \quad (3)$$

The expectancy operator  $E$  over the region  $R_n$  is defined as the following  $n$ -dimensional integral, where  $n$  is the number of uncertain parameters and  $f$  a general scalar function:

$$E_{R_n(d)}(f) = \int_{R_n(d)} f(\theta) j(\theta) d\theta \quad (4)$$

The above formulation considers both sets of constraints ( $g$  and  $g_q$ ) as hard constraints that must always be verified, with process stochastic flexibility (SF) being defined as the probability that these constraints are satisfied. On the other hand, Taguchi’s approach to quality engineering allows quality variables to take any value in face of uncertainty, penalizing deviations from a given target value, which is equivalent to consider quality constraints  $g_q$  as soft constraints. To clarify the difference between these two perspectives, let us consider the constraint  $y^L \leq y \leq y^U$ . Taguchi’s perspective is *continuous*, considering that a process can operate with  $y$  outside the interval  $[y^L; y^U]$ , although leading to high quality cost, since there are deviations from an optimal target value  $y^*$  for which the quality cost is zero (Fig. 1). The hard constraint perspective is different and *binary*, considering infeasible operation outside the interval  $[y^L; y^U]$  (infinite quality cost) and feasible operation with zero quality cost inside it (Fig. 1). If the quality variable  $y$  is, for instance, a product stream composition, Taguchi’s perspective will in general provide a better representation. In fact, under a realistic scenario, the process will still operate, even if product composition is outside the so-called ‘feasible’ region ( $y^L \leq y \leq y^U$ ), and it is likely for quality to decrease

gradually as one moves away from  $y^*$ . However, for situations where namely market demands or safety constraints translate into well-defined specification intervals, a mathematical representation in the form of hard constraints may be also needed.

Since our main goal here is to include robustness aspects within a generic process design framework, our formulation basically follows the Taguchi's gradual perspective, although hard robustness constraints can also be added to the problem formulation, as we will see later on. Thus, the explicit treatment of hard constraints of the form  $g(d, z, x, \theta) \leq 0$  is not addressed by our formulation, although that is possible by adopting a decomposition approach in order to evaluate the feasible region (Eq. (3)) (Pistikopoulos & Ierapetritou, 1995).

According to the above discussion, our formulation is constructed relaxing hard quality constraints  $g_q$  in Eq. (2) and replacing them by corresponding Taguchi loss functions. If one further assumes that all the constraints leading to infeasibility in the inner optimization problem can be treated as quality constraints, the integration region  $R_n$  becomes the entire uncertainty space  $\Theta$ , and the expectancy operator is then defined as:

$$E_{\Theta}(f) = \int_{\Theta} f(\theta)j(\theta)d\theta \quad (5)$$

From an engineering point of view, this is equivalent to allowing for process operation under every possible scenario of the uncertain parameters, even if quality constraints are not met.

To account for process quality, two different types of robustness criteria can be incorporated in formulation (Eq. (2)), (i) a penalty term in the objective function, i.e. a quality cost term,  $C_q$ , such as a Taguchi loss function; and/or (ii) explicit restrictions over process robustness metrics, such as the variance of a quality related variable. In the first case, the penalized objective function is redefined as:

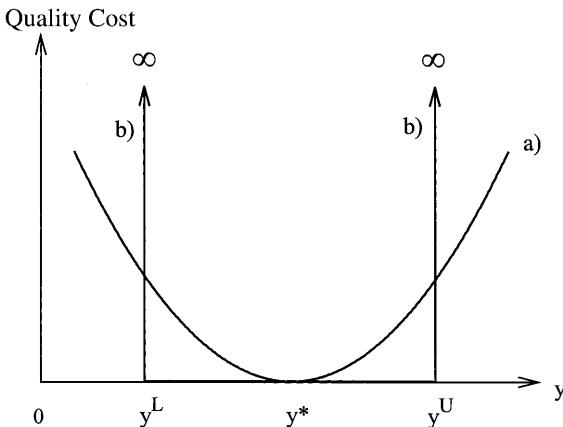


Fig. 1. Quality cost models according to (a) Taguchi's perspective and (b) hard constraint perspective.

$$f_q(d, z, x, y, y^*, \theta) = f(d, z, x, \theta) - C_q(y, y^*) \quad (6)$$

In the second case, a general robustness metric is defined as a function  $r$  of the statistical moments  $m_y$  of the quality variable  $y$ , with the following constraint being added to the problem formulation:

$$r(m_y) \leq \gamma \quad (7)$$

The statistical moments  $m_y$  are easily obtained using the expectancy operator. The first three moments, mean ( $\mu_y$ ), variance ( $\sigma_y^2$ ) and skewness ( $\xi_y$ ), are given by:

$$\begin{aligned} \mu_y &= E_{\Theta}(y) \\ \sigma_y^2 &= E_{\Theta}\{(y - \mu_y)^2\} \\ \xi_y &= E_{\Theta}\left\{\left(\frac{y - \mu_y}{\sigma_y}\right)^3\right\} \end{aligned} \quad (8)$$

As an example, the overall two-stage stochastic formulation for process quality, using variance as robustness metric, becomes as follows:

$$\begin{aligned} \text{Design stage: } & \max_d E_{\Theta}\{f'_q(d, \theta)\} \\ \text{s.t. } & E_{\Theta}\{(y' - \mu_y)^2\} \leq \gamma, \quad \mu_y = E_{\Theta}(y') \\ & d \in D, \quad \theta \in \Theta \\ \text{Operating stage: } & f'_q(d, \theta) = \max_{z, x, y} f_q(d, z, x, y, y^*, \theta) \\ \text{s.t. } & h(d, z, x, \theta) = 0 \\ & g(d, z, x, \theta) \leq 0 \\ & y' = y \\ & z \in Z, \quad x \in X, \quad y \in Y \end{aligned} \quad (9)$$

The formulation for other robustness metrics can be easily obtained using the appropriate Eq. (8) to compute  $r(m_y)$ .

The expectancy operator can be approximated using an integration formula over the uncertainty space, with a grid of  $N_p$  points  $\theta_i$  and the corresponding weights  $w_i$  (which reflect the value of the joint probability density function  $j(\theta_i)$ ):

$$E_{\Theta}(f) = \int_{\Theta} f(\theta)j(\theta) d\theta \cong \sum_{i=1}^{N_p} w_i f(\theta_i) \quad (10)$$

The two-stage formulation (Eq. (9)) may be then simplified to a single-level optimization problem:

$$\begin{aligned} \max_{d, z_i, x_i, y_i} & \sum_{i=1}^{N_p} w_i f_q(d, z_i, x_i, y_i, y^*, \theta_i) \\ \text{s.t. } & \sum_{i=1}^{N_p} w_i (y_i - \mu_y)^2 \leq \gamma, \quad \mu_y = \sum_{i=1}^{N_p} w_i y_i \\ & h(d, z_i, x_i, \theta_i) = 0 \\ & g(d, z_i, x_i, \theta_i) \leq 0 \\ & d \in D, \quad z_i \in Z, \quad x_i \in X, \quad y_i \in Y \\ & i = 1, \dots, N_p \end{aligned} \quad (11)$$

Table 1  
Taguchi loss functions

	Loss function	Example of a quality variable
(L1) Nominal-the-best (symmetric)	$C_q(y, y^*) = k(y - y^*)^2$ , same $k$ for all $y$	Product stream with a target composition $y^*$
(L2) Nominal-the-best (asymmetric)	$C_q(y, y^*) = k(y - y^*)^2$ ; $k = k_1$ , if $y < y^*$ ; $k = k_2$ , if $y \geq y^*$	Product stream with minimum purity requirement $y^*$ ( $k_2 = 0$ )
(L3) Larger-the-better	$C_q(y, y^*) = k(y - y^*)^2$ ; $k = k_1$ , if $y < y^*$ ; $k = 0$ , if $y \geq y^*$	Product stream purity with maximum possible value $y^*$
(L4) Smaller-the-better	$C_q(y, y^*) = k(y - y^*)^2$ ; $k = 0$ , if $y < y^*$ ; $k = k_2$ if $y \geq y^*$	Concentration of a pollutant in a waste stream where the minimum concentration that can be achieved is $y^*$

This formulation has the advantage of keeping the structure of the corresponding deterministic problem, and therefore, if both the deterministic problem and  $r(m_y)$  are convex (which is the case for  $r(m_y) = \sigma_y^2$ ), the resulting stochastic problem of the form (Eq. (11)) is also going to be convex.

The number of points in the integration formula,  $N_p$ , is the critical factor that determines the computational effort required to solve problem (Eq. (11)) and thus adopting an efficient integration technique is crucial. In the Appendix A, we present a cubature formula, especially constructed for the case where all uncertain parameters are normally distributed, that requires a significantly smaller number of points when compared with product Gauss formulae. However, there are problems for which the number of points needed is still significant, leading to computational complexity associated with the direct solution of problem (Eq. (11)). In these cases, a decomposition approach can be advantageous, transforming, for a fixed design, the original large problem in a sequence of smaller problems for each integration point (although in this case constraints  $r(m_y) \leq \gamma$  cannot be considered as such). Recent advances in directly solving problem (Eq. (11)) have been reported by van den Heever and Grossmann (1999).

The optimal solution of problem (Eq. (11)) defines the best design  $d^*$ , together with a robust operating policy  $Z^*$ , under the assumption of perfect information and control. This solution corresponds to the best average process performance that can be achieved considering the robustness criteria imposed in the problem formulation. General trade-offs between process robustness and profitability can be established by varying the weight of the quality cost,  $C_q$ , or the value of the  $\gamma$  parameter.

Although the optimal solution found does not give complete information about the output probability distributions, approximate PDFs for each quality variable can be obtained using the statistical moments estimates (Eq. (8)). The approximation presented by Tørvi and Hertzberg (1997) provides a truncated series based on a normal distribution whose coefficients are a function of the statistical moments for the real probability distribu-

tion. Computed the first three moments, the approximate PDF is given by:

$$\tilde{p}(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-u^2/2} \left[ 1 + \frac{\xi_y}{6}(u^3 - 3u) \right]$$

$$u = \frac{y - \mu_y}{\sigma_y} \quad (12)$$

### 3. Robustness criteria

The formulation presented so far in this article addresses quality related issues by means of two different strategies, penalization of the objective function through Taguchi loss functions *and/or* explicit constraints assumed over certain robustness metrics.

#### 3.1. Taguchi loss functions

Taguchi loss functions are models that express the loss of quality when a variable  $y$  deviates from its desired value  $y^*$  (Phadke, 1989). The most common one is the quadratic loss function, that considers the quality cost,  $C_q$ , to be proportional to the square of the deviation of  $y$  from its desired value  $y^*$ :

$$C_q(y, y^*) = k(y - y^*)^2 \quad (13)$$

The constant of proportionality,  $k$ , is designated as quality loss or penalty coefficient, and must be estimated based on economic considerations.

The quadratic loss function (Eq. (13)) is sufficiently flexible to account for different situations in quality engineering. Table 1 presents four kinds of loss functions based on the quadratic form (Eq. (13)), together with relevant application examples.

The *nominal-the-best* loss function penalizes both positive and negative deviations from the desired value  $y^*$  (Fig. 1), and therefore, can be used, for instance, to evaluate the quality level of a product stream with a certain target composition  $y^*$ . When it is desirable to penalize deviations in different ways, an *asymmetric* loss function can be used, by assuming  $k = k_1$ , if  $y < y^*$ , and  $k = k_2$ , otherwise. If  $k_2 = 0$ , we have a *one-sided*

loss function, well suited to quantify quality losses referring, for instance, to a product stream with a certain minimum purity requirement  $y^*$ .

In Phadke (1989) different forms of loss functions can be found for the *larger-the-better* and *smaller-the-better* types, based on the assumptions that the desired value for  $y$  is, respectively,  $\infty$  and 0. Since these assumptions are not usually applicable in many process engineering problems, we also consider quadratic loss functions for such cases. Under this formulation, in the larger-the-better situation  $y^*$  stands for the maximum possible value for the quality variable that may be reached (Saraiva, 1993), and all the scenarios for which  $y < y^*$  are penalized; in the smaller-the-better loss function,  $y^*$  represents the minimum possible value that one can achieve.

The incorporation of a loss function like (L1) in formulation (Eq. (11)) is straightforward, without introducing any sort of discontinuities. The implementation of an asymmetric loss function, like (L2), (L3) or (L4), needs to be addressed more carefully, because although the loss function and its derivative are continuous, they are defined in a discontinuous way. To handle this problem, we could introduce binary variables for each scenario to decide whether or not  $y_i$  is lower than  $y^*$ . This strategy would, however, increase drastically the complexity of the optimization problem, since the number of binary decision variables would be equal to  $\dim(y) \times N_p$ . An alternative consists of including the following formulation to switch between  $k_1$  and  $k_2$ , with the additional variables  $k_i$  and  $\lambda_i$ ,  $i = 1, \dots, N_p$  (a similar formulation was used by Bhatia and Biegler (1997) to correct an operating policy in the presence of active bounds for control variables):

$$\begin{aligned} \Delta y_i &= y_i - y^* \\ C_{q,i} &= k_i (\Delta y_i)^2 \\ \lambda_i \Delta y_i &\geq 0 \\ \lambda_i \Delta y_i &\geq \Delta y_i \\ 0 &\leq \lambda_i \leq 1 \\ k_i &= (1 - \lambda_i)k_1 + \lambda_i k_2 \\ i &= 1, \dots, N_p \end{aligned} \quad (14)$$

When  $y_i < y^*$ , the switching parameter  $\lambda_i$  is equal to 0 and  $k_i = k_1$ ; when  $y_i > y^*$ ,  $\lambda_i = 1$  and  $k_i = k_2$ . This formulation has the disadvantage of being non-convex and, in the case of  $y_i = y^*$ , there is no guarantee that  $\lambda = 1$ .

In the case of a one-sided or a larger-the-better loss function ( $k_2 = 0$ ), a more reliable and convex formulation is possible:

$$\begin{aligned} C_{q,i} &= k_1 (\Delta y_i)^2 \\ \Delta y_i &\geq y^* - y_i \quad (C1) \\ \Delta y_i &\geq 0 \quad (C2) \\ i &= 1, \dots, N_p \end{aligned} \quad (15)$$

Since  $C_q$  is being minimized, when  $y < y^*$ , (C1) is active; otherwise, (C2) is active and therefore, the quality cost is zero. A similar formulation can be constructed for the case of a smaller-the-better loss function ( $k_1 = 0$ ):

$$\begin{aligned} C_{q,i} &= k_2 (\Delta y_i)^2 \\ \Delta y_i &\geq y_i - y^* \quad (C1) \\ \Delta y_i &\geq 0 \quad (C2) \\ i &= 1, \dots, N_p \end{aligned} \quad (16)$$

### 3.2. Robustness metrics

Taguchi loss functions can describe the quality loss in a wide range of situations, either by varying the penalty coefficient  $k$  or by modifying the functional form of the loss function considered. However, there are problems in which it may be desirable to explicitly limit a given robustness metric, instead of using a Taguchi loss function, or to use both criteria simultaneously. Table 2 presents three examples of robustness metrics that can be applied to a quality variable, such as the purity of a product stream.

Although a Taguchi loss function contributes to center a quality variable PDF around the desired value  $y^*$ , we may want to directly consider the mean value of that variable. In these cases, a *minimum mean* criterion may be used. The maximum variance criterion can be useful in a broad scope of situations; in two-stage planning problems, for instance, operational robustness can be guaranteed by limiting the variance of second-stage cost (Ahmed & Sahinidis, 1998); another situation arises when a one-sided loss function is used (the variability in the region  $y \geq y^*$  is not penalized, and therefore, it may be convenient to explicitly restrict the variance of  $y$ ).

As was discussed, Taguchi's approach to quality engineering is based on a continuous view of quality loss related with the values assumed by quality variables. Imposing bounded hard constraints on quality

Table 2  
Robustness metrics

Robustness metric	Robustness criterion	Constraint $r(m_y) \leq \gamma$
Mean	Minimum mean	$\mu_y \geq \mu_{\min}$
Variance	Maximum variance	$\sigma_y^2 \leq \sigma_{\max}^2$
$q$ Quantile	Minimum $q$ quantile	$\int_{-\infty}^{y_{\min}} p(y) dy \leq q$

variables may lead to overdesigns, because they may behave as continuous random variables with  $p(y) > 0$ ,  $\forall y \in Y$ . Therefore, it may be more appropriate to assume constraints on the variable quantiles (Table 2).

The minimum mean and maximum variance criteria are easily implemented using Eq. (8) to compute the mean and variance and the integration formula (Eq. (10)) to approximate the expectancy operator. The criterion of minimum  $q$  quantile for the output PDF  $p(y)$  is more difficult to implement in a rigorous way, since the corresponding constraint involves the cumulative distribution of  $y$ . However, if  $p(y)$  is not very far from a normal distribution, the criterion can be approximately implemented using the constraint  $\mu_y + z\sigma_y \geq y_{\min}$ . The normal deviation  $z$  is such that  $N(z) = q$ , where  $N$  is the cumulative PDF for the standard normal distribution ( $\mu = 0$  and  $\sigma = 1$ ). The effectiveness of this approximate implementation should be tested a posteriori by evaluating the ‘true’ cumulative probability:

$$p(y \leq y_{\min}) = \int_{-\infty}^{y_{\min}} \tilde{p}(y) dy \quad (17)$$

with the approximate PDF  $\tilde{p}(y)$  being given by Eq. (12).

#### 4. Applications

In this section, we illustrate the usefulness of the proposed stochastic formulation for process quality by considering two process engineering examples — synthesis and design of a separation system and design of a reactor and heat exchanger system. We will perform different analysis in the two examples in order to illustrate different features and capabilities of the proposed design framework.

##### 4.1. Synthesis and design of a separation system

The problem addressed here is a variation of the synthesis and design problem considered previously by Ierapetritou (1995). The superstructure shown in Fig. 2 considers a separation system including three possible configurations (flash drum, distillation column or flash drum and distillation column in parallel), as well as bypass streams. The goal was to select the best separation scheme to be used in the presence of parameter uncertainty. The feedstream ( $F1$ ) of two components (A and B) must be separated into two product streams ( $P1$  and  $P2$ ) with minimum purity requirements of 78% of A in product  $P1$  and 70% of B in product  $P2$ . The mathematical model considered is presented in Table 3 and the values of its deterministic parameters reported in Table 4. The objective function (before penalty) assumed is profit, given by the difference between rev-

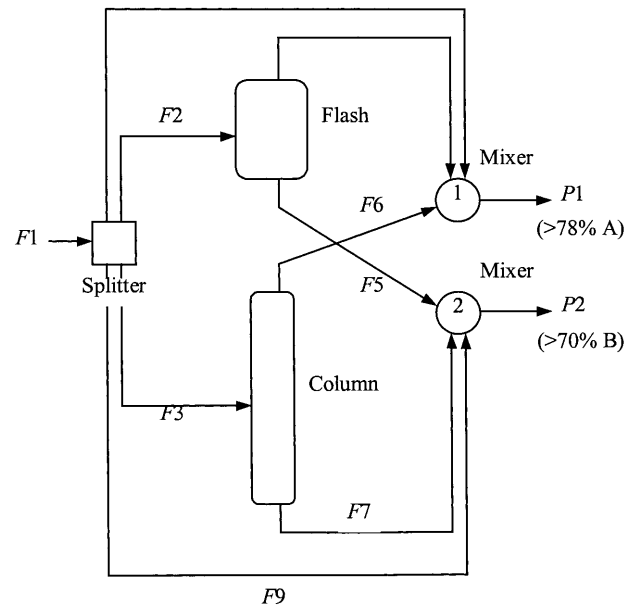


Fig. 2. Separation process superstructure.

enue due to product sales and plant overall cost (investment plus operating costs).

The model considered here includes five uncertain parameters (Table 5) described by normal distributions  $N(\mu_j, \sigma_j)$ . The corresponding joint normal PDF is denoted by  $N(\mu, \Sigma)$ , where  $\mu$  stands for the vector of means and  $\Sigma$  for the covariance matrix. Assuming all uncertain parameters to be mutually independent, matrix  $\Sigma$  becomes equal to the variance matrix,  $V$  (diagonal matrix with the variances as diagonal elements). This uncertainty space can be integrated using the specialized cubature  $SC_{5,1}$ , which for five uncertain parameters leads to only  $N_p = 2 \times 5 + 2^5 = 42$  points,  $u_i$ . The cubature points in the uncertainty space are obtained through the transformation Eq. (A.7).

The synthesis and design problem can now be formulated as a single-level optimization problem of the form Eq. (11), where the quality constraints are not included. Binary decisions, together with the diameters of the flash drum and column, are the design variables,  $d = \{y_F, y_D, D_F, D_C\}$ , while the split fractions stand as control variables,  $z = \{S_1, S_2, S_3, S_4\}$ , that can be optimally adjusted to uncertainty during operation. The quality variables are the compositions of products 1 and 2 in components A and B,  $y = \{x_A, x_B\}$ , with desired values  $y^* = \{0.78; 0.70\}$ . This problem can be convexified by using the substitutions  $d_F = D_F^2$ ,  $d_C = D_C^2$ , and redefining the quality variables as follows:

$$y_1 = P1_A - \frac{0.78}{0.22} P1_B, \quad y_1^* = 0$$

$$y_2 = P2_B - \frac{0.7}{0.3} P2_A, \quad y_2^* = 0$$

Table 3  
Separation system mathematical model<sup>a</sup>

Split mass balances	$F2_A = S_1 F1_A$ $F8_A = S_2 F1_A$ $F9_A = S_3 (F1 - F1_A)$ $F3_A = S_4 F1_A$ $0 \leq S_1 \leq y_F$ $S_1 + S_2 + S_3 + S_4 = 1$	$F2_B = S_1 (F1 - F1_A)$ $F8_B = S_2 (F1 - F1_A)$ $F9_B = S_3 (F1 - F1_A)$ $F3_B = S_4 (F1 - F1_A)$ $0 \leq S_4 \leq y_C$
Flash mass balances	$F4_A = 0.75 F2_A$ $F5_A = 0.25 F2_A$	$F4_B = (1 - k_B) F2_B$ $F5_B = k_B F2_B$
Distillation column mass balances	$F6_A = 0.825 F3_A$ $F7_A = 0.175 F3_A$	$F6_B = 0.125 F3_B$ $F7_B = 0.875 F3_B$
Mixers mass balances	$P1_A = F4_A + F6_A + F8_A$ $P2_A = F5_A + F7_A + F9_A$	$P1_B = F4_B + F6_B + F8_B$ $P2_B = F5_B + F7_B + F9_B$
Design equations	$D_F^2 \geq \frac{1.27M(F4_A + F4_B)}{45v\rho_V}$ $v = 0.064\sqrt{(\rho_L - \rho_V)/\rho_V}$	$D_C^2 \geq \frac{1.27M(F3_A + F3_B)}{35v\rho_V}$
Size restrictions	$2.2y_F \leq D_F \leq 5.5y_F$	$2.2y_C \leq D_C \leq 5.5y_C$
Quality constraints	$P1_A - (0.78/0.22)P1_B \geq 0$	$P2_B - (0.7/0.3)P2_A \geq 0$
Objective function	$\text{Profit} = \text{Pr}_1(P1_A + P1_B) + \text{Pr}_2(P2_A + P2_B)$ $- (20y_F + 2D_F^2 + 50y_C + D_C^2(0.3Np + 2))$ $- (C_F F1 + 0.2(F2_A + F2_B) + 2.5(F3_A + F3_B))$	

<sup>a</sup> Variables —  $y_F, y_C$ , binary variables representing the existence (or not) of the flash tank and column, respectively;  $D_F, D_C$ , diameters of flash tank and column, respectively (m);  $S$ , split fractions;  $F$  and  $P$ , flowrates.

To account for process quality, the problem was solved considering three different robustness criteria, (R1), nominal-the-best Taguchi loss function; (R2), one-sided loss function; and (R3), one-sided loss function plus maximum standard deviation (S.D.) for both quality variables. For all of the above scenarios, the global formulation is convex with loss function and additional constraints as follows:

$$C_{q,i} = k(y_{1,i}^2 + y_{2,i}^2) \quad (\text{R1})$$

$$\begin{cases} C_{q,i} = k_1[(\Delta y_{1,i})^2 + (\Delta y_{2,i})^2] \\ \Delta y_{1,i} \geq -y_{1,i} \\ \Delta y_{2,i} \geq -y_{2,i} \\ \Delta y_{1,i} \geq 0 \\ \Delta y_{2,i} \geq 0 \end{cases} \quad (\text{R2})$$

$$\begin{cases} (\text{R2}) \\ \sigma(y_1) \leq 0.5 \\ \sigma(y_2) \leq 0.5 \end{cases} \quad (\text{R3})$$

For the sake of simplicity, we are assuming equal penalty coefficients for both products, although different values might also be addressed, favoring the quality of one product when compared with the other.

Table 6 shows the results for  $k = k_1 = 64$ , obtained using GAMS/(CONOPT2,OSL) as optimization routines (Brooke, Kendrick & Meerans, 1992). As we can see, process performance in terms of product quality depends on the imposed robustness criteria. A nominal-the-best loss function leads to a solution centered around the desired values together with a low variabil-

ity. The one-sided penalty allows process operation for a wider region, resulting in a less conservative solution, with higher associated profit. The variability of both products' composition is now greater, especially in the case of  $x_B$ , since its PDF is located in the region of no penalization. With criterion (R3), the S.D. of  $x_B$  is reduced from 0.0382 to 0.0096, with a modest decrease in the expected profit. The substantial increase in the CPU time is due to the additional constraints imposed on  $x_A$  and  $x_B$  S.D., that add a new dense region to the optimization problem matrix, which was originally sparse.

General trade-offs between expected profit and quality requirements can be established by solving the problem parametrically in terms of the penalty coefficient. In Fig. 3, expected profit and optimal values for the

Table 4  
Deterministic parameter values

$F1$	Feed flowrate	33
$M$	Molecular weight	92
$\rho_L$	Liquid density of the feed	883
$\rho_V$	Vapor density of the feed	3
$V$	Velocity of vapor phase in drums	1.10
$N_p$	Number of plates in the column	20

Table 5  
Uncertainty model

$F1_A$	Flowrate of A in the feedstream	$N(15, 0.667)$
$k_B$	Separation constant of B in the flash tank	$N(0.84, 0.1)$
$\text{Pr}_1$	Price of product P1	$N(75, 1.25)$
$\text{Pr}_2$	Price of product P2	$N(55, 1.25)$
$C_F$	Cost of feed	$N(8, 0.5)$



Table 6  
Results for different robustness criteria

	(R1) $k = 64$	(R2) $k_1 = 64$	(R3) $k_1 = 64$
$E$ (profit) <sup>a</sup>	1491	1565	1523
$(y_F, y_C)$	(0,1)	(1,1)	(1,1)
$D_F$ (m)	–	3.015	3.194
$D_C$ (m)	4.635	4.388	4.849
$\mu(x_A)$	0.7778	0.7778	0.7778
$\sigma(x_A)$	0.0011	0.0020	0.0018
$\mu(x_B)$	0.7000	0.7937	0.7708
$\sigma(x_B)$	0.0022	0.0382	0.0096
CPU (s)	64.7	85.6	389.1

<sup>a</sup> Including penalty term.

flash and column diameters are represented as a function of the penalty coefficient  $k$ , for the case of criterion (R1). As the penalty coefficient becomes larger, expected profit decreases, with a discontinuity that corresponds to a change in the binary optimal decisions, while investment costs also become larger (except for the higher  $k$  values in the ‘only flash’ region).

The corresponding improvement in product quality can be seen in Fig. 4, where normalized quality costs are plotted as a function of the penalty coefficient — normalized quality cost associated with product A,  $C_{qA}/k$ ; product B,  $C_{qB}/k$ ; and total normalized quality cost,  $C_q/k$ . As can be seen, the decrease in  $C_q/k$  is mainly due to a decrease in the normalized quality cost associated with product A, meaning that the  $x_A$  PDF is critical to overall quality achievement. In the ‘only flash’ region, although a slight increase in  $C_{qB}/k$  can be observed, the total normalized quality cost is always decreasing, indicating an improvement of overall quality.

In Fig. 5, the mean and S.D. estimates obtained for both products are represented as a function of the quality loss coefficient. As it increases, the mean value of both products’ composition moves in the direction of the corresponding target values  $y^* = \{0.78, 0.70\}$ . In the ‘only flash’ region, the  $x_A$  S.D. increases, which at first sight could indicate a decrease in product A related

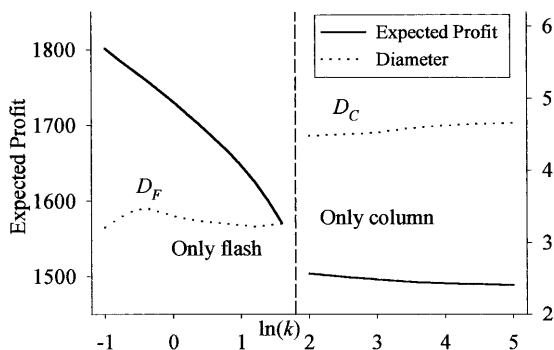


Fig. 3. Expected profit and optimal flash and column diameters vs. penalty coefficient.

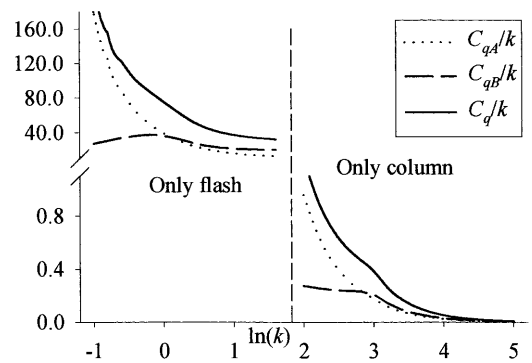


Fig. 4. Normalised quality costs vs. penalty coefficient.

quality. However, as discussed in the previous paragraph, this is not the case; the S.D. increase is compensated by a large mean value increase, resulting globally in smaller A quality costs. In the ‘only column’ region, the mean value of product B composition remains practically constant and equal to the 0.70 target value meaning that in this region quality improvement regarding B is only due to a reduction of  $x_B$  S.D.

#### 4.2. Design of a reactor and heat exchanger system

Fig. 6 presents a flowsheet consisting of a reactor and a heat exchanger, where a first order exothermic reaction  $A \rightarrow B$  takes place (Halemane & Grossmann, 1983; Chacon-Mondragon & Himmelblau, 1996). The goal here is to determine the optimal design (reactor volume,  $V$ , and area of the heat exchanger,  $A$ ) for a minimum  $A$  conversion of 0.9, under the presence of parameter uncertainty. Table 7 describes the system mathematical model, including mass and heat balances, process constraints and quality specifications, while deterministic parameter values are shown in Table 8. The objective function (before penalty) considered is the total plant annual cost including investment and operating costs.

Uncertainties in this case include five parameters (Table 9), described by a normal distribution  $N(\mu, \Sigma)$ , where  $\mu$  is the vector of means and  $\Sigma$  the corresponding covariance matrix. Uncertainty associated with parameter  $j$  is described using the notation  $\mu_j(1 \pm \varepsilon_j)$ ,  $j = 1, \dots, n$ , where  $\varepsilon_j$  is the relative error around the mean value,  $\mu_j$ ; S.D. is such that  $p[\mu_j(1 - \varepsilon_j) \leq \theta_j \leq \mu_j(1 + \varepsilon_j)] = 0.998$  and thus  $\sigma_j = \varepsilon_j \mu_j / 3.09$ . For instance, the feed flowrate has a relative error of 20% around its mean value of  $45.36 \text{ kmol h}^{-1}$ ; and therefore, its S.D. is  $\sigma(F_0) = 0.20 \times 45.36 / 3.09 = 2.936 \text{ kmol h}^{-1}$ .

The feed temperature,  $T_0$ , and the cooling water inlet temperature,  $T_{w1}$ , are assumed to have a positive correlation of 0.7. The covariance matrix is then given by  $\Sigma = V^{1/2} \rho V^{1/2}$ , where  $V$  stands for the variance matrix and  $\rho$  for the correlation matrix:

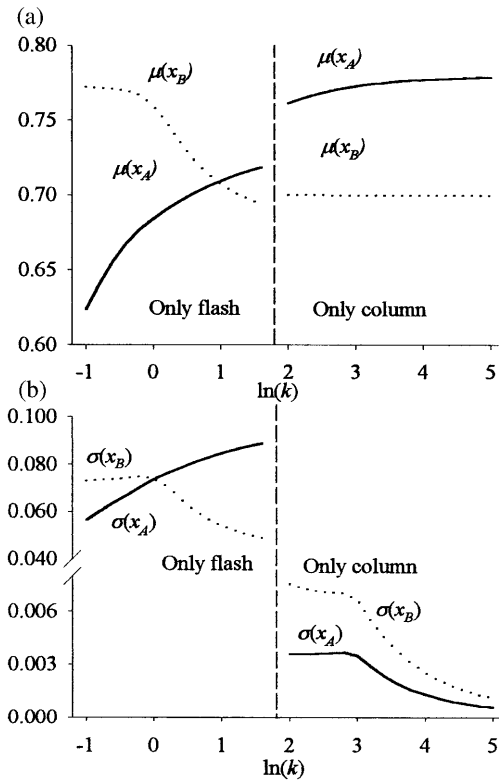


Fig. 5. Mean (a) and S.D. (b) of the quality variables vs. penalty coefficient.

$$\rho = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.7 & 0 & 0 \\ 0 & 0.7 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Once again, integration over the uncertainty space was performed using the specialized cubature formula  $SC_{5,1}$  with 42 points,  $u_i$ , and corresponding  $\theta_i$  points in the uncertainty space obtained according to Eq. (A.7).

The design problem can now be formulated as a single-level optimization problem of the form (Eq. (11)), excluding the quality constraint  $x_A \geq 0.9$ , with the variables classified as follows:  $d = \{V, A\}$ ,  $z = \{F_1, F_w\}$

Table 7  
Reactor and heat exchanger system mathematical model<sup>a</sup>

Reactor material balance	$F_0 x_A - k_R \exp(-E/RT_1) C_{A0} (1 - x_A) V = 0$ , $x_A = (C_{A0} - C_{A1}) / C_{A0}$
Reactor heat balance	$F_0 C_p (T_0 - T_1) - F_1 C_p (T_1 - T_2) + (-\Delta H_R) F_0 x_A = 0$
Heat exchanger design equation	$F_1 C_p (T_1 - T_2) = A U \Delta T_{lm}$ , $\Delta T_{lm} = ((T_1 - T_{w2}) - (T_2 - T_{w1})) / \ln[(T_1 - T_{w2}) / (T_2 - T_{w1})]$
Heat exchanger energy balance	$F_1 C_p (T_1 - T_2) = F_w c_{pw} (T_{w2} - T_{w1})$
Temperature bounds (K)	$311 \leq T_1 \leq 389$ , $311 \leq T_2 \leq 389$ , $294 \leq T_{w2} \leq 323$
Heat exchanger operation constraints	$T_1 - T_2 \geq 0$ , $T_{w2} - T_{w1} \geq 0$ , $T_1 - T_{w2} \geq 11.1$ , $T_2 - T_{w1} \geq 11.1$
Quality constraint	$x_A \geq 0.9$
Objective function (\$ per year)	$Cost = 691.2V^{0.7} + 873.6A^{0.6} + 1.76F_w + 7.056F_1$

<sup>a</sup>  $V$ , reactor volume ( $m^3$ );  $A$ , heat transfer area for the heat exchanger ( $m^2$ );  $F_1$ , reactant flowrate in the heat exchanger ( $kmol h^{-1}$ );  $F_w$ , cooling water flowrate ( $kg s^{-1}$ );  $x_A$ , conversion of A in the reactor;  $T_1$ , reactor temperature (K);  $T_2$ , reactant temperature after cooling (K);  $T_{w2}$ , cooling water outlet temperature (K).

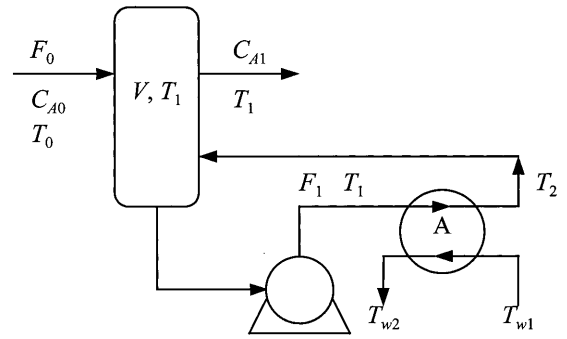


Fig. 6. Reactor and heat exchanger system.

and  $y = x_A$ , with a minimum acceptable value of  $y^* = 0.9$ . Three different robustness criteria were then adopted, (R4) one-sided loss function; (R5) one-sided loss function plus maximum S.D. for  $x_A$  and (R6) one-sided loss function plus minimum 0.05 quantile for  $x_A$ . The loss functions and additional constraints corresponding to each case are:

$$\begin{cases} C_{q,i} = k_1 (\Delta x_{A,i})^2 \\ \Delta x_{A,i} \geq 0.9 - x_{A,i} \\ \Delta x_{A,i} \geq 0 \end{cases} \quad (R4)$$

$$\begin{cases} (R5) \\ \sigma(x_A) \leq \sigma_{max} \end{cases} \quad (R5)$$

$$\begin{cases} (R6) \\ \mu(x_A) - 1.65\sigma(x_A) \geq x_{A,min} \end{cases} \quad (R6)$$

where it should be noted that criterion (R6) was implemented in an approximate way, since the deviation value of  $z = 1.65$  is based on a  $x_A$  normal distribution.

Table 10 shows the results obtained for particular values of  $k_1$ ,  $\sigma_{max}$  and  $x_{A,min}$ , using GAMS/CONOPT2 as optimization routines. One can examine the price 'paid' for when additional quality requirements that correspond to criteria (R5 and R6) are considered —

Table 8  
Deterministic parameter values

$C_{A0}$	Concentration of A in the feed stream	32.04 kmol m <sup>-3</sup>
$E/R$	Ratio of activation energy to the perfect gas constant	555.6 K
$\Delta H_R$	Molar heat of reaction	-23 260 kJ kmol <sup>-1</sup>
$C_p$	Reactant heat capacity	167.4 kJ kmol <sup>-1</sup> K <sup>-1</sup>
$C_{pw}$	Cooling water heat capacity	4.184 kJ kg <sup>-1</sup> K <sup>-1</sup>

Table 9  
Uncertainty model

$F_0$	Feed flowrate	45.36(1 ± 0.20) kmol h <sup>-1</sup>
$T_0$	Feed temperature	333(1 ± 0.04) K <sup>a</sup>
$T_{w1}$	Cooling water inlet temperature	293(1 ± 0.04) K <sup>a</sup>
$k_R$	Arrhenius rate constant	12(1 ± 0.20) h <sup>-1</sup>
$U$	Overall heat transfer coefficient	1635(1 ± 0.20) kJ (m <sup>-2</sup> h <sup>-1</sup> K <sup>-1</sup> )

<sup>a</sup> Positive correlation coefficient,  $r(T_0, T_{w1}) = 0.7$ .

optimal design decisions with higher investment costs are associated with more demanding quality requirements.

In order to show how the PDF of  $x_A$  varies as the robustness criteria become harder, we have solved the problem parametrically in terms of  $k_1$ , for the case of criterion (R4);  $\sigma_{\max}$  (with  $k_1 = 6.4 \times 10^6$ ), for the case of criterion (R5) and  $x_{A,\min}$  ( $k_1 = 6.4 \times 10^6$ ), for the case of criterion (R6) (Fig. 7). As expected, an increase in  $k_1$  shifts the PDF of  $x_A$  in the direction of greater conversion and reduces its variance. In the case of criterion (R5), it is interesting to observe that in order to comply with severe  $\sigma(x_A)$  constraints the optimal solution leads to a distribution for  $x_A$  that lies in the region of high conversion. With criterion (R6), as the parameter  $x_{A,\min}$  increases, not only the PDF of  $x_A$  is shifted in the direction of greater conversion, but also its variance is reduced, confirming that there is a larger process operation robustness associated with high conversion regions.

As stated before, the accuracy of criterion (R6) should be tested after simulation, using equation 17

to estimate the probability of  $x_A \leq x_{A,\min}$ . The results obtained are 0.056, 0.055 and 0.058, for  $x_{A,\min}$  equals to 0.90, 0.93 and 0.96, respectively. Comparing these values with 0.05 (the restricted quantile), we find that criterion (R6) is subject to an error smaller than a cumulative probability of 0.008.

## 5. Conclusions

We have developed a stochastic optimization framework for conducting process design under uncertainty that takes explicitly into account process robustness and product quality issues. Although the formulation is based in a two-stage approach, under the assumption of perfect information and control during process opera-

tion, the design problem has been formulated as a single-level stochastic optimization problem, where a Taguchi's perspective of continuous quality loss is adopted. The incorporation of different robustness criteria and their applicability has also been discussed in the context of process engineering applications.

An efficient cubature technique, suitable to integrate normally distributed uncertainties, has been applied for the estimation of multiple integrals involved in the formulation, reducing significantly the computational effort required, when compared with other integration methodologies, such as product Gauss rules or stratified sampling techniques.

The potential of the proposed design framework is illustrated with two process engineering application examples, where different robustness criteria are studied — nominal-the-best Taguchi loss functions, one-sided Taguchi loss functions, maximum variance and quantile constraints for a quality related variable. From such case studies, one can see quite clearly what implications and consequences are derived from considering different models to express quality related issues while addressing chemical process design problems.

Table 10  
Results for different robustness criteria

	(R4) ( $k_1 = 6.4 \times 10^6$ )	(R5) ( $k_1 = 6.4 \times 10^6$ , $\sigma_{\max} = 0.006$ )	(R6) ( $k_1 = 6.4 \times 10^6$ , $x_{A,\min} = 0.90$ )
$E$ (cost) <sup>a</sup>	13 340	14 099	13 585
$V$ (m <sup>3</sup> )	4.497	5.384	5.151
$A$ (m <sup>2</sup> )	7.760	8.076	8.002
$\mu(x_A)$	0.9007	0.9148	0.9121
$\sigma(x_A)$	0.0082	0.0060	0.0074
$\xi(x_A)$	-0.2136	-0.7753	-0.2141
CPU (s)	7.2	11.6	4.8

<sup>a</sup> Including penalty term.

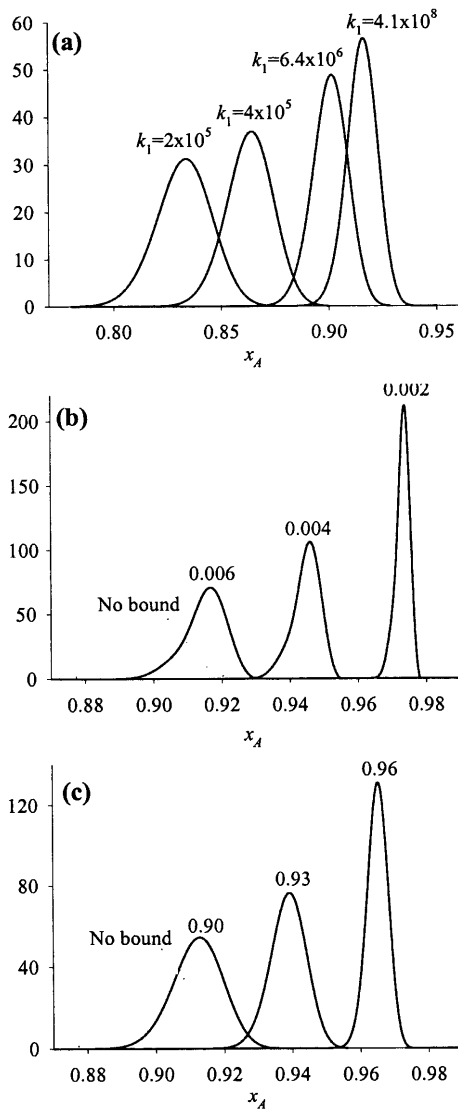


Fig. 7. PDF of  $x_A$  vs. (a) penalty coefficient-criterion (R4); (b) bound on  $\sigma(x_A)$ -criterion (R5); and (c) bound 0.05 quantile of  $x_A$ -criterion (R6).

## Appendix A. Specialized cubature formula

Cubatures are numerical integration techniques that generalize the principles of one-dimensional quadrature to multidimensional integration. Based on the definition of quadrature formula given by Engels (1980), we define cubature formula as follows.

**Definition.** A cubature formula is a numerical rule whereby the value of a multidimensional definite integral is approximated by use of information about the integrand only at a set of discrete points where the integrand is defined.

The mathematical notation that corresponds to such a definition is (Stroud, 1971):

$$\int_{R_n} f(u)w(u)du \approx \sum_{i=1}^{N_p} B_i f(u_i) \quad (\text{A.1})$$

where  $f$  is a scalar function of the vector of independent variables  $u$ ,  $w$  a given weight function and  $R_n$  the integration region in the  $n$ -dimensional Euclidean space,  $E_n$ . The cubature formula has  $N_p$  points  $u_i$  which lie in  $E_n$  and have corresponding coefficients or weights  $B_i$ .

To compute integral (Eq. (5)) using a formula like Eq. (A.1), we need to map the region  $R_n$  into the uncertainty space  $\Theta$ . This is achieved through the general transformation of the form:

$$\theta = \phi(u) \quad (\text{A.2})$$

where  $u \in R_n$  and  $\theta \in \Theta$ . Therefore, each  $u_i$  point in  $R_n$ -space has a correspondent  $\theta_i$  point in  $\Theta$ -space. The transformed weights are given by:

$$B_i^* = B_i |\det \text{Jac}(u_i)|, \quad i = 1, \dots, N_p$$

where  $\text{Jac}(u_i)$  is the Jacobian matrix of function  $\phi(u)$  evaluated at the point  $u_i$  (Davis & Rabinowitz, 1975).

Two different classes of cubature formulas can be distinguished — product formulas constructed using combinations of formulas for regions of dimension less than  $n$  and non-product formulas constructed by other methods. The most common product formula is the product Gauss formula for the  $n$ -dimensional cube, constructed using a Gaussian quadrature in each dimension, which results in a formula with a total number of points that increases exponentially with the dimension  $n$ .

In the present work, we use a non-product cubature formula especially constructed for the case where all uncertain parameters are normally distributed. In this case, integral (Eq. (5)) over the entire probability space takes the following form:

$$E_{\Theta}(f) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(\theta) j_N(\theta) d\theta \quad (\text{A.4})$$

where  $j_N(\theta)$  is the joint normal distribution for the random vector  $\theta$ , with the vector of means  $\mu$  and covariance matrix  $\Sigma$  (Johnson & Wichern, 1988):

$$j_N(\theta) = \frac{1}{(2\pi)^{n/2} (\det \Sigma)^{1/2}} \exp \left[ -\frac{1}{2} (\theta - \mu)^T \Sigma^{-1} (\theta - \mu) \right] \quad (\text{A.5})$$

We will denote this joint distribution by  $N(\mu, \Sigma)$ , and the corresponding uncertainty space by  $\Theta = \{\theta: \theta \sim N(\mu, \Sigma)\}$ .

Our specialized cubature formula is based on the rules reported by Stroud (1971) and labeled as  $E_n^2$ . These rules are constructed to integrate the entire  $n$ -dimensional space with weight function:

$$w_c = \exp(-u^T u) \quad (\text{A.6})$$

To integrate over the uncertainty space  $\Theta = \{\theta: \theta \sim N(\mu, \Sigma)\}$ , we use the following transformation, which allows one to handle correlated uncertainties:

$$\theta(u) = \mu + I_{\sqrt{2}} \sum^{1/2} u \quad (\text{A.7})$$

where  $I_{\sqrt{2}}$  is a diagonal matrix with all the diagonal elements equal to  $\sqrt{2}$ . The substitution of this transformation in Eq. (A.4) leads to:

$$\begin{aligned} E_{\theta}(f) &= \frac{1}{\pi^{n/2}} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f[\theta(u)] w_{\theta}(u) du \\ &\cong \frac{1}{\pi^{n/2}} \sum_{i=1}^{N_p} B_i f[\theta(u_i)] = \sum_{i=1}^{N_p} w_i f[\theta(u_i)] \quad (\text{A.8}) \end{aligned}$$

In the present work, we apply a fifth degree specialized cubature formula, labeled as  $SC_{5,1}$ . The formula degree is defined with respect to the function  $f$ , meaning that the formula is exact when  $f(\theta)$  is a polynomial of degree lower than or equal to 5. The formula is only defined for  $n \geq 3$  and has a total number of points given by  $N_p = 2^n + 2^n$ . The grid of points  $u_i$  and corresponding weights  $w_i$  are of the form:

$$(r, 0, \dots, 0)_{\text{FS}} \quad w_0$$

$$(s, s, \dots, 0)_{\text{FS}} \quad w_1$$

where FS denotes a fully symmetric set of points generated by permutation of the elements and its signs. The values of  $r$ ,  $s$ ,  $w_0$  and  $w_1$  are functions of the integral dimension  $n$  and provided by the following expressions:

$$\begin{aligned} r^2 &= \frac{n+2}{4} & s^2 &= \frac{n+2}{2(n-2)} \\ w_0 &= \frac{4}{(n+2)^2} & w_1 &= \frac{(n-2)^2}{2^n(n+2)^2} \quad (\text{A.9}) \end{aligned}$$

The non-product cubature formula  $SC_{5,1}$  is much more efficient than the fifth degree specialized product Gauss formula ( $N_p = 3^n$ ) or the ninth degree product Gauss formula ( $N_p = 5^n$ , degree defined with respect to  $f(\theta)j(\theta)$ ). For integral dimensions of up to 10 it can compete with the most efficient stratified sampling techniques, such as the HSS technique. Further details on the application of cubature formulae to process design under uncertainty are provided in Bernardo, Pistikopoulos and Saraiva (1999).

## References

- Acevedo, J., & Pistikopoulos, E. N. (1996). A parametric MINLP algorithm for process synthesis problems under uncertainty. *Industrial & Engineering Chemistry Research*, 35, 147.
- Acevedo, J., & Pistikopoulos, E. N. (1997). A multiparametric programming approach for linear process engineering problems under uncertainty. *Industrial and Engineering Chemistry Research*, 36, 717.
- Ahmed, S., & Sahinidis, N. V. (1998). Robust process planning under uncertainty. *Industrial & Engineering Chemistry Research*, 37, 1883.
- Bernardo, F. P., & Saraiva, P. M. (1998). A robust optimization framework for process parameter and tolerance design. *American Institute of Chemical Engineering Journal*, 44, 2007.
- Bernardo, F. P., Pistikopoulos, E. N., & Saraiva, P. M. (1999). Integration and computational issues in stochastic design and planning optimization problems. *Industrial & Engineering Chemistry Research*, 38, 3056.
- Bhatia, T. K., & Biegler, L. T. (1997). Dynamic optimization for batch design and scheduling with process model uncertainty. *Industrial & Engineering Chemistry Research*, 36, 3708.
- Brooke, A., Kendrick, D., & Meerans, A. (1992). *GAMS: a user's guide, release 2.25*. The Scientific Press.
- Chacon-Mondragon, O. L., & Himmelblau, D. M. (1996). Integration of flexibility and control in process design. *Computers & Chemical Engineering*, 20, 447.
- Davis, P. J., & Rabinowitz, P. (1975). *Methods of numerical integration*. New York: Academic Press.
- Diwekar, U. M., & Kalagnanam, J. R. (1997a). Efficient sampling technique for optimization under uncertainty. *American Institute of Chemical Engineering Journal*, 43, 440.
- Diwekar, U. M., & Kalagnanam, J. R. (1997b). Efficient sampling technique for off-line quality control. *Technometrics*, 39, 308.
- Diwekar, U. M., & Rubin, E. S. (1994). Parameter design methodology for chemical processes using a simulator. *Industrial & Engineering Chemistry Research*, 33, 292.
- Engels, H. (1980). *Numerical quadrature and cubature*. London: Academic Press.
- Georgiadis, M. C., & Pistikopoulos, E. N. (1999). An integrated framework for robust and flexible process systems. *Industrial & Engineering Chemistry Research*, 38, 133.
- Grossmann, I. E., & Sargent, R. W. H. (1978). Optimum design of chemical plants with uncertain parameters. *American Institute of Chemical Engineering Journal*, 37, 517.
- Grossmann, I. E., Halemane, K. P., & Swaney, R. E. (1983). Optimization strategies for flexible chemical processes. *Computers & Chemical Engineering*, 7, 439.
- Halemane, K. P., & Grossmann, I. E. (1983). Optimal process design under uncertainty. *American Institute of Chemical Engineering Journal*, 29, 425.
- van den Heever, S. A., & Grossmann, I. E. (1999). Disjunctive multiperiod optimization methods for design and planning of chemical process systems. *Computers & Chemical Engineering*, 23, 1075.
- Ierapetritou, M. G. (1995). Optimization approaches for process engineering problems under uncertainty. Ph.D. thesis. Imperial College, London.
- Ierapetritou, M. G., Pistikopoulos, E. N., & Floudas, C. A. (1996). Operational planning under uncertainty. *Computers & Chemical Engineering*, 20, 1499.
- Johnson, R. A., & Wichern, D. W. (1988). *Applied multivariate statistical analysis*. London: Prentice Hall.
- Phadke, M. S. (1989). *Quality engineering using robust design*. New Jersey: Prentice Hall.
- Pertsinidis, A., Grossmann, I. E., & McRae, G. J. (1998). Parametric optimization of MILP programs and a framework for the parametric optimization of MINLPs. *Computers & Chemical Engineering*, 22(Suppl.), S205.
- Pistikopoulos, E. N. (1995). Uncertainty in process design and operations. *Computers & Chemical Engineering*, 19(Suppl.), S553.
- Pistikopoulos, E. N., & Ierapetritou, M. G. (1995). Novel approach for optimal process design under uncertainty. *Computers & Chemical Engineering*, 19, 1089.
- Pistikopoulos, E. N., & Mazzuchi, T. A. (1990). A novel flexibility approach for processes with stochastic parameters. *Computers & Chemical Engineering*, 14, 991.
- Samsatli, J. N., Papageorgiou, L. G., & Shah, N. (1998). Robustness metrics for dynamic optimization models under parameter uncertainty. *American Institute of Chemical Engineering Journal*, 44, 1993.

- Saraiva, P. M. (1993). Data-driven learning frameworks for continuous process analysis and improvement. Ph.D. Thesis. Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- Straub, D. A., & Grossmann, I. E. (1990). Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties. *Computers & Chemical Engineering*, *14*, 967.
- Straub, D. A., & Grossmann, I. E. (1993). Design optimization of stochastic flexibility. *Computers & Chemical Engineering*, *17*, 339.
- Stroud, A. H. (1971). *Approximate calculation of multiple integrals*. London: Prentice Hall.
- Taguchi, G. (1986). Introduction to quality engineering. Asian Productivity Association.
- Tørvi, H., & Hertzberg, T. (1997). Estimation of uncertainty in dynamic simulation results. *Computers & Chemical Engineering*, *21*(Suppl.), S181.
- Varvarezos, D. K., Grossmann, I. E., & Biegler, L. T. (1992). An outer-approximation method for multiperiod design optimization. *Industrial & Engineering Chemistry Research*, *31*, 1466.