AIChE WILEY InterScience® DISCOVER SOMETHING GREAT

# Flexibility Analysis: Taking into Account Fullness and Accuracy of Plant Data

**G. M. Ostrovsky, and Yu. M. Volin**
Karpov Institute of Physical Chemistry, Vorontsovo Pole 10, Moscow 103064, Russia

*DOI 10.1002/aic.10931*
*Published online July 6, 2006 in Wiley InterScience (www.interscience.wiley.com).*

*The modern foundations of flexibility analysis were developed in the 1980s by Grossmann and co-workers. They formulated the approaches to solve the main problems of flexibility analysis, namely, the feasibility test, flexibility index, and two-stage optimization problem (TSOP). All these formulations are based on the supposition that during the operation stage there is enough plant data to accurately determine all uncertain parameter values. However, in practice, this assumption is likely to be untrue. To deal with this problem, the feasibility test and TSOP were extended to account for the possibility of accurately estimating only some of the uncertain parameters. However, these extensions have a drawback. They do not take into account all the data that are available at the operation stage. A formulation is discussed that will remove this drawback of the TSOP. The formulation requires working out a modification of the split and bound approach, which had been developed for solving the TSOP. Then, two computation experiments are performed to demonstrate the advantage of taking the additional plant data from the operation stage into account.* © 2006 American Institute of Chemical Engineers *AIChE J,* 52: 3173–3188, 2006
*Keywords: flexibility, uncertainty, optimization, design*

## Introduction

Here we will consider the problem of determining optimal equipment sizes and regimes of chemical processes (CP).[1,2] This is done at the design stage of the CP, after design constraints are specified. The constraints are often ecological or related to safety and performance.

Satisfying these design constraints is complicated by the presence of uncertainties in mathematical models. They are induced by the inexactness of physical and chemical laws, plant data, and changes in internal and external factors during the CP operation. Hence, one can give the following general formulation of a CP optimization problem at the design stage taking into account the uncertainties in mathematical models: it is necessary to find the optimal equipment sizes and regimes of the CP to guarantee its flexibility (that is, the ability to satisfy all the constraints at the operation stage). This problem is solved with the help of CP flexibility theory.

Foundations of modern flexibility theory were laid in the 1980s[3-5] by Grossmann and his coworkers. They formulated the main problems of flexibility analysis—the feasibility test, flexibility index, and two-stage optimization problem—and gave approaches to their solving. A number of clever approaches to the evaluation of the feasibility test, flexibility index metrics, and the solution of the TSOP have been suggested in the literature. When certain conditions of convexity are met, Halemane and Grossmann[3] suggested an enumeration approach, which identified one of the vertices of an uncertainty region (multidimensional rectangle) as the appropriate evaluation point for the feasibility test. In this approach there was need to enumerate all the vertices of the uncertainty region. To avoid complete enumeration of the vertices, Swaney and Grossmann[4] developed the *branch and bound* method for calculation of the *flexibility index*. Grossmann and Floudas[5] developed the *active constraints set* method, which permits evaluating the feasibility test at a point that is not necessarily a

Correspondence concerning this article should be addressed to G. Ostrovsky at ostralex@yandex.ru.

vertex. Besides, they reduced the problem of calculation of the feasibility function to an MINLP problem. Halemane and Grossmann[3] developed a method for solving the two-stage optimization problem based on the *outer approximation algorithm*, in which an explicit enumeration method is used for calculation of the feasibility function.

The theory has developed greatly in the last 20 years. Pistikopoulos and Grossmann[6] developed a method for the retrofit design problem under uncertainty; the strategy was based on the active constraints set method. In the ensuing years, there was significant development in flexibility analysis. Ierapetritou and Pistikopoulos[7] formulated the optimization problem under uncertainty as a two-stage stochastic programming problem in which the tradeoff between the process economics and flexibility is explored. Straub and Grossmann[8] introduced the concept of *stochastic flexibility* and formulated a problem in which the stochastic flexibility is maximized under the presence of some cost constraints. Ostrovsky et al.[9,10] proposed the branch and bound method for estimation of the feasibility test and the *split and bound* method for solving the TSOP.

Calculation of the statistical expectation (which involves the evaluation of multidimensional integrals) is one of the most computationally intensive parts of the stochastic programming problem. To make the calculations computationally tractable, Acevedo and Pistikopoulos[11] suggested three alternative integration schemes. For normally distributed uncertain variables, Bernardo et al.[12] obtained a special quadrature formula, which significantly decreases the number of required evaluation points (knot points) for the multidimensional integrals. Bansal et al.[13] developed methods for solving design optimization problems under uncertainty, based on piecewise linear approximations of the feasibility test and flexibility index. Ahmad et al.[14] improved the algorithm from ref. [7]. Raspanti et al.[15] suggested new formulations for the feasibility test and flexibility index based on the use of the constraint aggregation function. Ierapetritou[16] describes an approach for identifying a CP's feasible region and suggests a new metric for evaluating process flexibility based on the convex hull, inscribed within the feasible region. A global optimization approach was proposed by Floudas et al.[17] for determination of the globally optimal solution for the feasibility test and the flexibility index problems. The new formulation of the feasibility test was developed in refs. [1,2] for the case when there is incomplete information about the uncertain parameters at the operation stage.

The following factors affect the formulations of the CP flexibility condition and the optimization problem under uncertainty:

1. *The level of uncertainty at the design stage.* The following three variants are possible: (i) We know the parameters of mathematical models accurately; (ii) At the operation stage, the uncertain parameters can take any values from the uncertainty region; (iii) We know the probability distribution functions for the uncertain parameters.

2. *The level of uncertainty at the operation stage.* The level of uncertainty depends strongly on the completeness and accuracy of process data available at the operation stage. It is clear that the completeness and accuracy of process data depend on measurement errors and availability of sensors. Therefore, the set of uncertain parameters can be divided into three groups. The first group is a subset of the uncertain parameters that can be determined with enough precision at the operation stage. The second group is a subset of the uncertain parameters such that their intervals of uncertainty at the operation stage remain the same as at the design stage. Finally, the third group is a subset of the uncertain parameters that cannot be measured with sufficient accuracy during the operation stage.

3. *The type of process constraints.* The constraints can be "hard" or "soft." Hard constraints must not be violated during the operation stage. On the other hand, if some violations are allowed, the constraints are called soft.

4. *The tools for ensuring the flexibility of the CP.* When constructing a flexible CP one can use: (a) design and control variables, (b) only the design variables, or (c) only the control variables. If a change of the control variables is exploited at the operation stage, then the optimization problem is called the two-stage optimization problem (TSOP).

We suppose that at each time instant during the operation stage, CP optimization is carried out using process models, and at least one uncertain parameter is made more precise using the new plant data available at this time instant. We refer to the CP optimization problem at the current time instant as an *internal optimization problem.*

We suppose that all the constraints are hard. For this case we will consider two-stage formulations of the optimization problem under uncertainty for different levels of fullness and accuracy of the plant data available at the operation stage. In order to formulate these optimization problems, we will use the average cost function as an objective function and a flexibility condition as a constraint. We will consider here four cases. The two-stage optimization problem corresponding to the *i*-th case will be designated as TSOP*i.*

### *Case 1*

At the operation stage, plant data permit us to determine precisely all the uncertain parameter values. In this case the two-stage optimization problem under uncertainty is given as[3]:

*TSOP1:*

$$f_1 = \min_{d} E_{\boldsymbol{\theta}}\{f_1^*(\boldsymbol{d}, \boldsymbol{\theta})\} \tag{1}$$

$$\chi_1(\boldsymbol{d}) \le 0, \tag{2}$$

where $E_{\boldsymbol{\theta}\in T}\{f_1^*(\boldsymbol{d}, \boldsymbol{\theta})\}$ is the mathematical expectation of the $f_1^*(\boldsymbol{d}, \boldsymbol{\theta})$ over the region $T$ and $f_1^*(\boldsymbol{d}, \boldsymbol{\theta})$ is the solution of *the internal optimization problem for TSOP1:*

$$f_1^*(\boldsymbol{d}, \boldsymbol{\theta}) = \min_{z} f(\boldsymbol{d}, z, \boldsymbol{\theta}) \tag{3}$$

$$g_j(\boldsymbol{d}, z, \boldsymbol{\theta}) \le 0, \quad j = 1, \ldots, m$$

where $\boldsymbol{d}$ is the vector of the design variables, $\boldsymbol{z}$ is the vector of the control variables, and $\boldsymbol{\theta}$ is the vector of the uncertain parameters. The feasibility function $\chi_1(\boldsymbol{d})$ is of the form

$$\chi_1(d) \equiv \max_{\boldsymbol{\theta}\in T} \min_{z} \max_{j\in J} g_j(\boldsymbol{d}, z, \boldsymbol{\theta}) \le 0 \tag{4}$$

where $J = (1, \ldots m)$. The reduced objective function $f(\boldsymbol{d}, z, \boldsymbol{\theta})$ and the reduced process constraints

$$g_j(\boldsymbol{d}, z, \boldsymbol{\theta}) \le 0, \quad j = 1, \ldots, m \tag{5}$$

are obtained from the original optimization problem

$$\min_{d,z,x} \bar{f}(\boldsymbol{d}, \boldsymbol{x}, z, \boldsymbol{\theta})$$

$$\varphi_j(\boldsymbol{d}, \boldsymbol{x}, z, \boldsymbol{\theta}) = 0, \quad j = 1, \ldots, r \tag{6}$$

$$\bar{g}_j(\boldsymbol{d}, \boldsymbol{x}, z, \boldsymbol{\theta}) \le 0, \quad j = 1, \ldots, m \tag{7}$$

where $\boldsymbol{x}$ is the vector of state variables. Equation 6 describes the state of the CP (that is, material and energy balances), whereas the inequalities in Eq. 7 are the design specifications. Here $r = \dim \boldsymbol{x}$ and $\boldsymbol{x}$ can be obtained from Eq. 6 either analytically or numerically using fixed values of the variables $\boldsymbol{d}, \boldsymbol{z}, \boldsymbol{\theta}$. Thus, the variables $\boldsymbol{x}$ are implicit functions of $\boldsymbol{d}, \boldsymbol{z}, \boldsymbol{\theta}$:

$$\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{d}, z, \boldsymbol{\theta}) \tag{8}$$

Substituting this expression into $\bar{f}(\boldsymbol{d}, \boldsymbol{x}, z, \boldsymbol{\theta})$ and Eq. 7, we obtain the reduced objective function

$$f(\boldsymbol{d}, z, \boldsymbol{\theta}) \equiv \bar{f}(\boldsymbol{d}, \boldsymbol{x}(\boldsymbol{d}, z, \boldsymbol{\theta}), z, \boldsymbol{\theta}) \tag{9}$$

and the reduced process constraints

$$g_j(\boldsymbol{d}, z, \boldsymbol{\theta}) \equiv \bar{g}_j(\boldsymbol{d}, \boldsymbol{x}(\boldsymbol{d}, z, \boldsymbol{\theta}), z, \boldsymbol{\theta}) \le 0, \quad j = 1, \ldots, m \tag{10}$$

Constraint (2) in Eq. 1 is the CP flexibility condition (feasibility test). The internal optimization problem should be solved at each time instant during the operation stage after the new values of all the uncertain parameters are determined. The optimal values of the control variables obtained by solving the internal optimization problem must then be used to control the CP. The drawback of this formulation is that it is based on the supposition that at each time instant at the operation stage we can determine accurate values of all the uncertain parameters. Therefore, we will now consider the cases when at the operation stage we have incomplete or inaccurate information about the uncertain parameters.

### *Case 2*

At the operation stage, plant data permit us to determine accurate values only of a subset $\boldsymbol{\theta}^1$ of the set of the uncertain parameters: $\boldsymbol{\theta}^1 \in T^1$, $T^1 = \{\boldsymbol{\theta}^1 : \boldsymbol{\theta}^{1,L} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{1,U}\}$. Let $\boldsymbol{\theta}^2$ be the subset of the uncertain parameters that cannot be determined at the operation stage: $\boldsymbol{\theta}^2 \in T^2 \quad T^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^{2,L} \le \boldsymbol{\theta} \le \boldsymbol{\theta}^{2,U}\}$. It is clear that $\boldsymbol{\theta} = (\boldsymbol{\theta}^1, \boldsymbol{\theta}^2)$. We will suppose that $\boldsymbol{\theta}^1$ and $\boldsymbol{\theta}^2$ are independent. As such, the probability density function (pdf) can be expressed as $\mu(\boldsymbol{\theta}^1, \boldsymbol{\theta}^2) = \mu(\boldsymbol{\theta}^1)\mu(\boldsymbol{\theta}^2)$ where $\mu(\boldsymbol{\theta}^1)$ and $\mu(\boldsymbol{\theta}^2)$ are the individual pdf's. In this case the optimization problem under uncertainty is given as[1]:

*TSOP2:*

$$f_2 = \min_d E_{\theta^1}\{f_2^*(\boldsymbol{d}, \boldsymbol{\theta}^1)\} \tag{11}$$

$$\chi_2(\boldsymbol{d}) \le 0 \tag{12}$$

where $E_{\boldsymbol{\theta}^1}\{f^*_2(\boldsymbol{d}, \boldsymbol{\theta}^1)\}$ is the mathematical expectation of the $f^*_2(\boldsymbol{d}, \boldsymbol{\theta}^1)$ over the region $T^1$ and $f_2^*(\boldsymbol{d}, \boldsymbol{\theta}^1)$ is the solution of *the internal optimization problem for TSOP2:*

$$f_2^*(\boldsymbol{d}, \boldsymbol{\theta}^1) = \min_z E_{\boldsymbol{\theta}^2}\{f(\boldsymbol{d}, z, \boldsymbol{\theta}^1, \boldsymbol{\theta}^2) \tag{13}$$

$$\max_{\boldsymbol{\theta}^2 \in T^2} g_j(\boldsymbol{d}, z, \boldsymbol{\theta}^1, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \ldots, m$$

The feasibility function $\chi_2(\boldsymbol{d})$ is of the form[1,2]:

$$\chi_2(\boldsymbol{d}) \equiv \max_{\boldsymbol{\theta}^1 \in T^1} \min_z \max_{\boldsymbol{\theta}^2 \in T^2} \max_{j \in J} g_j(\boldsymbol{d}, z, \boldsymbol{\theta}^1, \boldsymbol{\theta}^2) \tag{14}$$

We showed[1] that

$$\chi_2(\boldsymbol{d}) \ge \chi_1(\boldsymbol{d}) \tag{15}$$

$$f_2 \ge f_1 \tag{16}$$

Equation 11 can be transformed to the form[1]:

$$f_2 = \min_{\boldsymbol{d}, z(\boldsymbol{\theta}^1)} \int_{T^1} \int_{T^2} f(\boldsymbol{d}, z(\boldsymbol{\theta}^1), \boldsymbol{\theta}^1, \boldsymbol{\theta}^2)\mu(\boldsymbol{\theta}^1)\mu(\boldsymbol{\theta}^2)d\boldsymbol{\theta}^1 d\boldsymbol{\theta}^2 \tag{17}$$

$$\max_{\boldsymbol{\theta}^2 \in T^2} g_j(\boldsymbol{d}, z(\boldsymbol{\theta}^1), \boldsymbol{\theta}^1, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \ldots, m$$

$$\chi_2(\boldsymbol{d}) \le 0$$

Since the search variables $z$ in Eq. 11 depend on $\boldsymbol{\theta}^1$, $z$ are multivariate functions $z(\boldsymbol{\theta}^1)$.

Using Gaussian quadrature,[18] we obtain *the discrete variant of the TSOP2:*

$$f_2 = \min_{\boldsymbol{d}, z^i} \sum_{i \in I_1} w_i \sum_{l \in L_1} \nu_l f(\boldsymbol{d}, z^i, \boldsymbol{\theta}^{1i}, \boldsymbol{\theta}^{2l}) \tag{18}$$

$$\max_{\boldsymbol{\theta}^2 \in T^2} g_j(\boldsymbol{d}, z^i, \boldsymbol{\theta}^{1i}, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \ldots, m \quad i \in I_1 \tag{19}$$

$$\chi_2(\boldsymbol{d}) \le 0 \tag{20}$$

where $w_i$ and $v_i$ are weight coefficients, satisfying the conditions

$$\sum_{i \in I_1} w_i = 1, \quad \sum_{l \in L_1} \nu_l = 1 \tag{21}$$

If the probability density functions are not known, the weight coefficients and the sets of approximation points $S_1 = \{\boldsymbol{\theta}^{1i} : i$

$\in I_1\}$ and $S_2 = \{\boldsymbol{\theta}^{2l}\colon l \in L_1\}$ need to be specified a priori by a user.

### *Case 3*

Here as in case 2 some parameters $\boldsymbol{\theta}^2$ cannot be determined at the operation stage, and other parameters $\boldsymbol{\theta}^1$ can be measured with some error, determined by the accuracy of the sensors. In this case $\boldsymbol{\theta}^1$ can be represented in the form $\boldsymbol{\theta}^1 = \tilde{\boldsymbol{\theta}}^1 + \boldsymbol{\theta}^3$, where components of the vector $\tilde{\boldsymbol{\theta}}^1$ are the values of the uncertain parameters $\boldsymbol{\theta}^1$ as measured by sensors and the components of the vector $\boldsymbol{\theta}^3$ are errors inherent to the measurement. The region $\tilde{T}^1$ of possible values of the vector $\tilde{\boldsymbol{\theta}}^1$ is given as

$$\tilde{T}^1 = \{\tilde{\boldsymbol{\theta}}^1\colon \boldsymbol{\theta}^{1,L} - \boldsymbol{\theta}^{3,U} \le \tilde{\boldsymbol{\theta}}^1 \le \boldsymbol{\theta}^{1,U} + \boldsymbol{\theta}^{3,U}\}$$

This case can be reduced to the previous one. Indeed, here we can distinguish the following two groups of the uncertain parameters. The first group consists of the parameters $\tilde{\boldsymbol{\theta}}^1$ and the second group consists of the parameters $\boldsymbol{\theta}^3, \boldsymbol{\theta}^2$. It is clear that the first group corresponds to the set $\boldsymbol{\theta}^1$ in case 2 and the second group corresponds to the set $\boldsymbol{\theta}^2$ in case 2. We now introduce the following designations $\bar{\boldsymbol{\theta}}^2 = (\boldsymbol{\theta}^3, \boldsymbol{\theta}^2)\bar{T}^2 = (\boldsymbol{\theta}^3, \boldsymbol{\theta}^2 : \boldsymbol{\theta}^3 \in T^3, \boldsymbol{\theta}^2 \in T^2)$. By substituting the expression for $\boldsymbol{\theta}^1$ in Eqs. 11-13 and using $\bar{\boldsymbol{\theta}}^2$ instead of $\boldsymbol{\theta}^2$ and $\bar{T}^2$ instead of $T^2$, we obtain the formulations of the internal optimization problem, flexibility condition, and the optimization problem under uncertainty:

*the internal optimization problem for TSOP3:*

$$f_3^*(\boldsymbol{d}, \tilde{\boldsymbol{\theta}}^1) = \min_{z} E_{\boldsymbol{\theta}^2}\{f(\boldsymbol{d}, \boldsymbol{z}, \tilde{\boldsymbol{\theta}}^1 + \boldsymbol{\theta}^3, \boldsymbol{\theta}^2)$$

$$\max_{\boldsymbol{\theta}^2 \in T^2} g_j(\boldsymbol{d}, \boldsymbol{z}, \tilde{\boldsymbol{\theta}}^1 + \boldsymbol{\theta}^3, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \ldots, m$$

*the flexibility condition for TSOP3:*

$$\chi_3(\boldsymbol{d}) \equiv \max_{\tilde{\boldsymbol{\theta}}^1 \in T^1} \min_{z} \max_{\bar{\boldsymbol{\theta}}^2 \in \bar{T}^2} \max_{j \in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \tilde{\boldsymbol{\theta}}^1 + \boldsymbol{\theta}^3, \boldsymbol{\theta}^2)$$

*TSOP3:*

$$f_3 = \min_{d} E_{\tilde{\boldsymbol{\theta}}^1}\{f_3^*(\boldsymbol{d}, \tilde{\boldsymbol{\theta}}^1)\}$$

$$\chi_3(\boldsymbol{d}) \le 0$$

where $E_{\tilde{\boldsymbol{\theta}}^1}\{f_3^*(\boldsymbol{d}, \tilde{\boldsymbol{\theta}}^1)\}$ is the mathematical expectation of the $f_3^*(\boldsymbol{d}, \tilde{\boldsymbol{\theta}}^1)$ over the region $\tilde{T}^1$.

TSOP1 is based on the supposition that at the operation stage we can determine accurate values of all the uncertain parameters. This condition is very restrictive and it is often not satisfied in practice.[19] In connection with this, the new formulations of the two-stage optimization problem (TSOP2) and the feasibility test were developed in refs.[1,2] The formulation of TSOP2 is based on the assumption that at the operation stage we can calculate the exact values only of a subset of the uncertain parameters. However, the formulation of TSOP2 has another drawback. It does not permit taking into account the plant data that may be available at the operation stage. We will introduce a new formulation of the TSOP that will not have this drawback.

### *New feasibility test and TSOP formulation*

Later on we consider the case when at the operation stage there is not enough plant data for correction of some uncertain parameters. The set of the uncertain parameters can be split into two subsets: measured parameters $\bar{\boldsymbol{\theta}}$ (the parameters that are measured directly with help of sensors) and unmeasured ones $\bar{\bar{\boldsymbol{\theta}}}$. We suppose that the parameters $\bar{\boldsymbol{\theta}}$ are measured without mistakes. The parameters $\bar{\bar{\boldsymbol{\theta}}}$ are not measured directly. Let $\bar{x}$ and $\bar{\bar{x}}$ designate the sets of the state variables that can and cannot be measured at the operation stage, respectively.

$$\boldsymbol{x} = \begin{pmatrix} \bar{\boldsymbol{x}} \\ \bar{\bar{\boldsymbol{x}}} \end{pmatrix}$$

Consider at first the case when there exists the subset $J_1$ of Eq. 6 depending on the measured state variables only. Thus, Eq. 6 can be written in the following form:

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, z, \bar{\boldsymbol{\theta}}, \bar{\bar{\boldsymbol{\theta}}}) = 0, \quad j \in J_1$$
$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}, z, \bar{\boldsymbol{\theta}}, \bar{\bar{\boldsymbol{\theta}}}) = 0, \quad j = 1, \ldots, r, \quad j \notin J_1 \qquad (22)$$

Consider some time instant at the operation stage. Suppose for some $z = z^*$ we have measured the values of the parameters $\bar{\boldsymbol{\theta}}$ and the state variables $\bar{\boldsymbol{x}}$. If

$$|J_1| > \dim \bar{\bar{\boldsymbol{\theta}}}, \qquad (23)$$

where $|J_1|$ is a number of elements in the set $J_1$, then the values of the parameters $\bar{\bar{\theta}}$ can be found by solving the following inverse problem:

$$\min_{\bar{\bar{\boldsymbol{\theta}}}} \sum_{j \in J_1} \alpha_j[\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, z^*, \bar{\boldsymbol{\theta}}, \bar{\bar{\boldsymbol{\theta}}})]^2, \qquad (24)$$

where $\boldsymbol{\alpha}_j$ are weight coefficients. In this case we must use TSOP1 to solve the optimization problem under uncertainty.

If Eq. 23 is not met, then we cannot determine the values of the parameters $\bar{\bar{\theta}}$. In this case it is possible to solve TSOP2 with

$$\boldsymbol{\theta}^1 = \bar{\boldsymbol{\theta}}, \quad \boldsymbol{\theta}^2 = \bar{\bar{\boldsymbol{\theta}}} \qquad (25)$$

using Eq. 22 in the following form:

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, z, \boldsymbol{\theta}^1, \boldsymbol{\theta}^2) = 0, \quad j \in J_1$$
$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}, z, \boldsymbol{\theta}^1, \boldsymbol{\theta}^2) = 0, \quad j = 1, \ldots, r, \quad j \notin J_1 \qquad (26)$$

However, in this case we do not take into account the plant data (the state variables $\bar{\boldsymbol{x}}$) that may be obtained at the operation stage. In connection with this, we will give here a new formulation of the two-stage optimization problem that takes into account this information. As above, we will assume that the sets of the measured and unmeasured parameters are $\boldsymbol{\theta}^1$, $\boldsymbol{\theta}^2$, respectively.

At first we will formulate the internal optimization problem. Let the parameters $\bar{\boldsymbol{\theta}}^1$, $\bar{\boldsymbol{\theta}}^2$ correspond to some time instant at the operation stage. Suppose for some $z = z^*$ we have measured values of the parameters $\boldsymbol{\theta}^1$ and the state variables $\bar{\boldsymbol{x}}$. Designate the measured values as $\bar{\boldsymbol{\theta}}^1$ and $\bar{\boldsymbol{x}}^*$, respectively. The parameters $\boldsymbol{\theta}^2$ should satisfy the conditions (see Eq. 26):

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}^*, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0, \quad j \in J_1. \tag{27}$$

Introduce the region $T^2$:

$$\bar{T}^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^2 \in T^2, \quad \varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}^*, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0, \quad j \in J_1\} \tag{28}$$

For this time instant, the value $\bar{\boldsymbol{\theta}}^1$ is known while $\theta^2$ is not known. However, $\boldsymbol{\theta}^2$ should satisfy Eq. 27. Therefore, for fixed $\boldsymbol{z}$ and $\bar{\boldsymbol{\theta}}^1$, the following flexibility conditions should be met:

$$\forall\ \boldsymbol{\theta}^2 \in \bar{T}^2\ \ \forall j \in J[\, g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0, \quad j \in J]. \tag{29}$$

Transform the logical condition to an analytical form. According to equivalent relations (A.1) we can replace the logical condition $\forall j \in J[\, g_j(d, z, \theta) \le 0, j \in J]$ in Eq. 29 with the following inequality:

$$\max_{j \in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0.$$

Then, Eq. 29 takes the form:

$$\forall\ \boldsymbol{\theta}^2 \in \bar{T}^2 \max_{j \in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0$$

Using the equivalent relation (A1) again, we transform this inequality to the form:

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}^2} \max_{j \in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0$$

We can rewrite this inequality as (see (A3)):

$$\max_{j \in J} \max_{\boldsymbol{\theta}^2 \in \bar{T}^2} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0 \tag{30}$$

Using equivalent relation (A1) again, we can replace Eq. 30 with the following $m$ inequalities:

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}^2} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \ldots, m \tag{31}$$

We will take the mathematical expectation of $f(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2)$ with respect to $\boldsymbol{\theta}^2$ as an objective function. Uniting this objective function and flexibility Eq. 31, we obtain *the internal optimization problem for TSOP4:*

$$f_4^*(d, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = \min_z E_{\theta^2}\{f(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2)\}$$
$$\max_{\boldsymbol{\theta}^2 \in \bar{T}^2} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \ldots, m \tag{32}$$

Let $\bar{z}$ be the solution of the problem. In contrast to internal optimization Eq. 13 for TSOP2, this formulation takes into account the plant data $\bar{\boldsymbol{x}}^*$, that can be obtained at the operation stage. The mathematical expectation

$$E_{\bar{\boldsymbol{\theta}} \in T}\{f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)\} \tag{33}$$

characterizes a future performance of CP (here, $\bar{\boldsymbol{\theta}} = (\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$). Therefore, it is advisable to use the value $E_{\bar{\boldsymbol{\theta}} \in T}\{ f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)\}$ as the objective function for the two-stage optimization problem that must be solved at the design stage. However, in order to use Eq. 33 we should be able to calculate the value $f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ (that is, to solve Eq. 32) at any point $(\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \in T$. Unfortunately, we cannot do this. Indeed, in Eq. 32 we use the value $\bar{\boldsymbol{x}}^*$, which is measured at the operation stage. It is clear that at the design stage we cannot know results of a future measurement and, consequently, we do not know the value $\bar{\boldsymbol{x}}^*$. . However, the future measurement can be replaced by the simulation of the chemical process using the full mathematical Eq. 26. This means that for fixed $\boldsymbol{\theta}^1 = \bar{\boldsymbol{\theta}}^1$, $\boldsymbol{\theta}^2 = \bar{\boldsymbol{\theta}}^2$ , and $z = z^*$ , the value $\bar{\boldsymbol{x}}^*$ can be found by solving the system of nonlinear Eq. 26 with respect to $\bar{\boldsymbol{x}}$, $\bar{\bar{\boldsymbol{x}}}$. Thus, we must solve the following system of the nonlinear equations with respect to $\bar{\boldsymbol{x}}$, $\bar{\bar{\boldsymbol{x}}}$ :

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = 0, \quad j \in J_1$$
$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = 0, \quad j = 1, \ldots, r, \quad j \notin J_1 \tag{34}$$

Let $\bar{\boldsymbol{x}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2), \bar{\bar{\boldsymbol{x}}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ be the solution of Eq. 34. Replacing the value $\bar{\boldsymbol{x}}^*$ with $\bar{\boldsymbol{x}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ in Eq. 28, we obtain the following form of the region $\bar{T}^2$ :

$$\bar{T}^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^2 \in T^2, \quad \varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2), \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0, \quad j \in J_1\} \tag{35}$$

Thus, we can calculate the value $f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ at all points $\bar{\boldsymbol{\theta}}^1$, $\bar{\boldsymbol{\theta}}^2 \in T$ and, consequently, Eq. 33 can be used as the objective function for the two-stage optimization problem.

Now we will formulate the flexibility condition (the feasibility test). For fixed $\bar{\boldsymbol{\theta}}^1$, $\bar{\boldsymbol{\theta}}^2$, the flexibility condition takes the form:

$$\exists\, z\ \ \forall\ \boldsymbol{\theta}^2 \in \bar{T}^2\ \ \forall j \in J[\, g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \le 0]$$

This condition means that for any fixed $\bar{\boldsymbol{\theta}}^1$, $\bar{\boldsymbol{\theta}}^2$ there exists such vector $z$ that all the constraints $g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \le 0$ are met for all $\boldsymbol{\theta}^2 \in \bar{T}^2$. It is clear that the flexibility of the chemical process is guaranteed if this condition is met at all points $\bar{\boldsymbol{\theta}}^1$, $\bar{\boldsymbol{\theta}}^2$ belonging to $T$:

$$\forall\ \bar{\boldsymbol{\theta}}^1 \in T^1\ \ \forall\ \bar{\boldsymbol{\theta}}^2 \in T^2\ \ \exists\, z\ \ \forall\ \boldsymbol{\theta}^2 \in \bar{T}^2\ \ \forall j \in J(g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0). \tag{36}$$

This is the logical form of the *flexibility condition* of the chemical process.

As indicated above, Eq. 29 can be replaced by Eq. 30. Therefore, Eq. 36 can be rewritten in the form:

$$\forall\, \bar{\boldsymbol{\theta}}^1 \in T^1 \;\; \forall\, \bar{\boldsymbol{\theta}}^2 \in T^2 \;\; \exists\, z \max_{j\in J} \max_{\boldsymbol{\theta}^2\in \bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0 \quad (37)$$

Using relation ($A2$) we can replace the inequality $\exists z \max_{j\in J}\max_{\boldsymbol{\theta}^2\in\bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0$ with the following inequality:

$$\min_z \max_{j\in J} \max_{\boldsymbol{\theta}^2\in\bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0$$

Hence, Eq. 37 can be replaced by the following one:

$$\forall\, \bar{\boldsymbol{\theta}}^1 \in T^1 \quad \forall\, \bar{\boldsymbol{\theta}}^2 \in T^2 \min_z \max_{j\in J} \max_{\boldsymbol{\theta}^2\in\bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0$$

Using relation ($A1$) we can reduce this condition to the following form:

$$\max_{\bar{\boldsymbol{\theta}}^1\in T^1} \max_{\bar{\boldsymbol{\theta}}^2\in T^2} \min_z \max_{\boldsymbol{\theta}^2\in\bar{T}^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0$$

Since the maximization operators are commutative (see (A.3)), we can transform this inequality to the following form of *the flexibility condition:*

$$\max_{\bar{\boldsymbol{\theta}}^1\in T^1} \max_{\bar{\boldsymbol{\theta}}^2\in T^2} \min_z \max_{j\in J} \max_{\boldsymbol{\theta}^2\in\bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0 \quad (38)$$

It is the new formulation of the feasibility test. If it is met, then for all values of the uncertain parameters $\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2 \in T$ , one can find such values of the control variables $z$ that Eq. 5 will be met. Introduce the designations

$$\chi_4(\boldsymbol{d}) = \max_{\bar{\boldsymbol{\theta}}^1\in T^1} \max_{\bar{\boldsymbol{\theta}}^2\in T^2} \min_z \max_{j\in J} \max_{\boldsymbol{\theta}^2\in\bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \quad (39)$$

$$\chi_4^i(\boldsymbol{d}) = \max_{\bar{\boldsymbol{\theta}}^1\in T_i^1} \max_{\bar{\boldsymbol{\theta}}^2\in T^2} \min_z \max_{j\in J} \max_{\boldsymbol{\theta}^2\in\bar{T}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \quad (40)$$

Here $\chi_4(\boldsymbol{d})$, $\chi_4^i(\boldsymbol{d})$ are the feasibility functions for the whole uncertainty region $T$ and for the subregion $T_i = \{\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2 : \bar{\boldsymbol{\theta}}^1 \in T_i^1, \bar{\boldsymbol{\theta}}^2 \in T^2\}$ where $T_i^1$ is the $i$-th subregion of $T^1$ ($T_i^1 \in T^1$). Taking into account Eq. 39, we can write down the feasibility test (Eq. 38) in the following form:

$$\chi_4(\boldsymbol{d}) \le 0 \quad (41)$$

Condition

$$\chi_4^i(\boldsymbol{d}) \le 0 \quad (42)$$

denotes that the chemical process will be flexible for all $\bar{\boldsymbol{\theta}} \in T_i$. Using equality (A.3) we can unite the first two maximization operators in Eq. 39. Consequently, we can rewrite $\chi_4(\boldsymbol{d})$ as follows:

$$\chi_4(\boldsymbol{d}) = \max_{\bar{\boldsymbol{\theta}}\in T} \min_z \max_{\boldsymbol{\theta}^2\in\bar{T}^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \quad (43)$$

where $\bar{\boldsymbol{\theta}} = (\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$.

Now we can formulate the two-stage optimization problem using Eq. 33 as the objective function and flexibility Eq. 41 as the constraint. Uniting objective Eq. 33 and flexibility Eq. 41, we obtain

*TSOP4:*

$$f_4 = \min_d E_{\bar{\theta}\in T}\{f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)\} \quad (44)$$

$$\chi_4(\boldsymbol{d}) \le 0$$

Let $\boldsymbol{d}^*$ be the solution of this problem. If in the chemical process $\boldsymbol{d} = \boldsymbol{d}^*$, then at the operation stage the average cost Eq. 9 will be minimal and for each $\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2 \in T$ Eq. 5 can be satisfied by tuning the control variables $z$. Of course, this statement is valid only when one uses the values $\bar{z}$ (obtained by solving internal optimization Eq. 32) for control of the chemical process.

Using Gaussian quadrature ([18]) we transform the objective function of Eq. 44:

$$E_{\bar{\theta}\in T}\{f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2\} = \sum_{i\in I_1} w_i \sum_{l\in L_1} \nu_l f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^{1i}, \bar{\boldsymbol{\theta}}^{2l}) \quad (45)$$

Here $w_i$ and $v_i$ are weight coefficients ($w_i \ge 0$, $\nu_l \ge 0$, $\Sigma_{i\in I_1} w_i = 1$, $\Sigma_{l\in L_1}\nu_l = 1$); $\bar{\boldsymbol{\theta}}^{1i}$, ($i \in I_1$) are the approximation points in the region $T^1$; $\bar{\boldsymbol{\theta}}^{2l}$, ($l \in L_1$) are the approximation points in the region $T^2$; $f_4^*$ ($\boldsymbol{d}, \bar{\boldsymbol{\theta}}^{1i}, \bar{\boldsymbol{\theta}}^{2l}$) is the solution of Eq. 32 for $\bar{\boldsymbol{\theta}}^1 = \bar{\boldsymbol{\theta}}^{1i}$ and $\bar{\boldsymbol{\theta}}^2 = \bar{\boldsymbol{\theta}}^{2l}$. Designate region $\bar{T}^2$ for this problem as $\bar{T}_{il}^2$. It has the form:

$$\bar{T}_{il}^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^2 \in T^2, \quad \varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2l}), z^*, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2) = 0, \quad j \in J_1\}$$

where $\bar{\boldsymbol{x}}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2l})$ together with $\bar{\bar{\boldsymbol{x}}}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2l})$ is the solution of the following system of nonlinear equations:

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, z^*, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2l}) = 0, \quad j \in J_1$$
$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}, z^*, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2l}) = 0, \quad j = 1, \ldots, r, \quad j \notin J_1$$

Since $f_4^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^{1i}, \bar{\boldsymbol{\theta}}^{2l})$ is the solution of Eq. 32 for $\bar{\boldsymbol{\theta}}^1 = \bar{\boldsymbol{\theta}}^{1i}$ and $\bar{\boldsymbol{\theta}}^2 = \bar{\boldsymbol{\theta}}^{2l}$ , we can rewrite the right-hand side of Eq. 45 as follows:

$$\sum_{i\in I_1} w_i \sum_{l\in L_1} \nu_l \min_z [E_{\boldsymbol{\theta}^2\in T^2}\{f(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \bar{\boldsymbol{\theta}}^2)\} / \max_{\boldsymbol{\theta}^2\in\bar{T}_{il}^2} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2) \le 0, \quad j \in J] \quad (46)$$

Using Gaussian quadrature[18] we transform the integral

$$E_{\boldsymbol{\theta}^2\in T^2}\{f(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2)\} = \int_{\boldsymbol{\theta}^2} f(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2)\mu(\boldsymbol{\theta}^2)d\boldsymbol{\theta}^2 \quad (47)$$

to the form

$$E_{\boldsymbol{\theta}^2\in T^2}\{f(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2)\} = \sum_{q\in Q_1} p_q f(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2q}),$$

$$\sum_{q\in Q_1} p_q = 1 \quad p_q \geq 0$$

Substitute this expression into Eq. 46:

$$\sum_{i\in I_1} w_i \sum_{l\in L_1} \nu_l \min_z \left[\sum_{q\in Q_1} p_q f(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \bar{\boldsymbol{\theta}}^{2q}) / \max_{\boldsymbol{\theta}^2\in \bar{T}^2_{il}} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2\right) \leq 0, \quad j \in J \quad (48)$$

It is easy to see that the control variables vectors $z$ corresponding to different pairs of the parameters $\bar{\theta}^{1i}$, $\bar{\theta}^{2l}$ do not depend on each other. Since condition A.8 (see theorem 2A) is met, then we can change the order of the minimization and summation operators in Eq. 48. Changing the order of these operators in Eq. 48 and substituting the obtained expression in Eq. 46, we deduce the final form of $E_{\bar{\boldsymbol{\theta}}\in T}\{f^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2\}$:

$$E_{\bar{\boldsymbol{\theta}}\in T}\{f^*(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2\} = \min_{z^{IL}} \sum_{i\in I_1} w_i \sum_{l\in L_1} \nu_l \sum_{q\in Q_1} p_q f(\boldsymbol{d}, z^{il}, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2q}) \max_{\boldsymbol{\theta}^2\in \bar{T}^2_{il}} g_j(\boldsymbol{d}, z^{il}, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2) \leq 0 \quad j = 1, \ldots, m, \quad i \in I_1, \quad l \in L_1 \quad (49)$$

where $\boldsymbol{z}^{il}$ is the vector of the control variables corresponding to the approximation point $\bar{\boldsymbol{\theta}}^{1i}$, $\bar{\boldsymbol{\theta}}^{2l}$; $z^{IL}$ is the set of the control variables vectors corresponding to all the approximation points $\bar{\boldsymbol{\theta}}^{1i}$, $\bar{\boldsymbol{\theta}}^{2l}$ ($\forall i \in I_1, \forall l \in L_1$): $z^{IL} = \{z^{il}, \forall i \in I_1, \forall l \in L_1\}$. . Substituting Eq. 49 in Eq. 44, we obtain *the discrete variant of TSOP4:*

$$f_4 = \min_{\boldsymbol{d}, z^{IL}} \sum_{i\in I_1} w_i \sum_{l\in L_1} \nu_l \sum_{q\in Q_1} p_q f(\boldsymbol{d}, z^{il}, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^{2q}) \max_{\boldsymbol{\theta}^2\in \bar{T}^2_{il}} g_j(\boldsymbol{d}, z^{il}, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2) \leq 0 \quad j = 1, \ldots, m, \quad i \in I_1, \quad l \in L_1 \chi_4(\boldsymbol{d}) \leq 0 \quad (50)$$

Consider the particular case of this problem when for the approximation of Eq. 47 we use only one approximation point $\boldsymbol{\theta}^{2q}$ coinciding with the point $\bar{\boldsymbol{\theta}}^{2l}$. In this case $Q_1 = \{l\}$ and from Eq. 50 we obtain *the simplified discrete variant of TSOP4:*

$$\tilde{f}_4 = \min_{\boldsymbol{d}, z^{IL}} \sum_{i\in I_1} w_i \sum_{l\in L_1} \nu_l f(\boldsymbol{d}, z^{il}, \bar{\boldsymbol{\theta}}^{1i}, \bar{\boldsymbol{\theta}}^{2l}) \quad (51)$$

$$\max_{\boldsymbol{\theta}^2\in \bar{T}^2_{il}} g_j(\boldsymbol{d}, z^{il}, \bar{\boldsymbol{\theta}}^{1i}, \boldsymbol{\theta}^2) \leq 0, \quad j = 1, \ldots, m, \quad i \in I_1, \quad l \in L_1$$

$$\chi_4(\boldsymbol{d}) \leq 0 \quad (52)$$

where $\tilde{f}_4$ is some approximation of $f_4$. Of course, the approximation can be rough enough. Therefore, it can be made more precise by solving Eq. 37 using the solution of Eq. 51 as an initial approximation.

Compare the discrete variant of TSOP2 (Eq. 18) and the simplified discrete variant of TSOP4 (Eq. 51). It should be noted that the control variables vector $z^{il}$ in Eq. 51 is associated with a single approximation point $\boldsymbol{\theta}^{il} = (\boldsymbol{\theta}^{1i}, \boldsymbol{\theta}^{2l})$. On the other hand, it is seen from Eq. 18 that in contrast to Eq. 51, the vector $z^i$ in TSOP2 is associated with all the approximation points from the subregion

$$T(\boldsymbol{\theta}^{1i}) = \{\boldsymbol{\theta} : \boldsymbol{\theta}^2 \in T^2; \boldsymbol{\theta}^1 = \boldsymbol{\theta}^{1i}\}.$$

Thus, if the number of the approximation points in Eq. 51 is equal to the number of the approximation points in TSOP2, then Eq. 51 has more search variables than TSOP2. At first we consider the case when the control variables in Eq. 51 do not depend on the superscript $l$, that is, the following equalities are valid:

$$z^{il} = z^i, \quad \forall\, l \in L_1 \quad (53)$$

In this case, Eqs. 18 and 51 have the identical objective functions. Since the relation $\bar{T}^2_{il} \subseteq T^2$ holds, we have

$$\max_{\theta^2\in \bar{T}^2_{il}} g_j(\boldsymbol{d}, z, \boldsymbol{\theta}^{1i}, \boldsymbol{\theta}^2) \leq \max_{\theta^2\in T^2} g_j(\boldsymbol{d}, z, \boldsymbol{\theta}^{1i}, \boldsymbol{\theta}^2) \quad \forall\, l \in L_1, \quad \forall\, i \in I_1. \quad (54)$$

Compare $\chi_2(\boldsymbol{d})$ and $\chi_4(\boldsymbol{d})$. For the sake of simplicity, we represent $\chi_2(\boldsymbol{d})$ in the following form:

$$\chi_2(\boldsymbol{d}) \equiv \max_{\bar{\boldsymbol{\theta}}^1\in T^1} \min_z \max_{\boldsymbol{\theta}^2\in T^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \quad (55)$$

Since for any $\bar{\boldsymbol{\theta}}^1 \in T^1$ and $\bar{\boldsymbol{\theta}}^2 \in T^2$ we have $\bar{T}^2 \subseteq T^2$, then for $\forall \bar{\boldsymbol{\theta}}^1 \in T^1$, $\forall \bar{\boldsymbol{\theta}}^2 \in T^2$ the following inequality holds:

$$\max_{\boldsymbol{\theta}^2\in \bar{T}^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \leq \max_{\boldsymbol{\theta}^2\in T^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2)$$

Hence, for $\forall \bar{\boldsymbol{\theta}}^1 \in T^1$, $\forall \bar{\boldsymbol{\theta}}^2 \in T^2$ we have

$$\min_z \max_{\boldsymbol{\theta}^2\in \bar{T}^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \leq \min_z \max_{\boldsymbol{\theta}^2\in T^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2)$$

Since the inequality holds for any $\bar{\boldsymbol{\theta}}^2$, then we have

$$\max_{\bar{\boldsymbol{\theta}}^2\in T^2} \min_z \max_{\boldsymbol{\theta}^2\in \bar{T}^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \leq \min_z \max_{\boldsymbol{\theta}^2\in T^2} \max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \quad \forall\, \bar{\boldsymbol{\theta}}^1 \in T^1$$

Since the inequality holds for any $\bar{\boldsymbol{\theta}}^1$, then we have

$$\max_{\bar{\boldsymbol{\theta}}^1\in T^1}\max_{\bar{\boldsymbol{\theta}}^2\in T^2}\min_{z}\max_{\boldsymbol{\theta}^2\in\bar{T}^2}\max_{j\in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le \max_{\bar{\boldsymbol{\theta}}^1\in T^1}\min_{z}\max_{\boldsymbol{\theta}^2\in T^2}\max_{j\in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \quad (56)$$

Taking into account Eqs. 55 and 39, we obtain

$$\chi_4(\boldsymbol{d}) < \chi_2(\boldsymbol{d}). \quad (57)$$

It follows from Eqs. 54 and 57 that the feasible region of Eq. 18 is contained in the feasible region of Eq. 51. Therefore:

$$\tilde{f}_4 \le f_2. \quad (58)$$

The optimal value of the objective function of Eq. 51 without Eq. 53 can only be less than $\tilde{f}_4$. Therefore, in the general case Eq. 58 is valid. Thus, taking into account the additional plant data available at the operation stage enables creating a more economical CP.

Consider now the general case when it is impossible to represent the mathematical model in Eq. 22. In other words, the mathematical model is of the form

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}, z^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0, \quad j = 1, \ldots, r \quad (59)$$

Suppose for $\boldsymbol{z} = \boldsymbol{z}^*$ we have measured the values of $\bar{\boldsymbol{x}}^*$ and $\bar{\boldsymbol{\theta}}^1$. We will solve Eq. 32 similarly to the previous case. However, here we cannot use the region $\bar{T}^2$ in Eq. 28 since we consider the case when it is impossible to select a subset of equations that depend only on the measured state variables. Therefore, we should construct the region $\bar{T}^2$ differently. Let us solve the set of nonlinear Eq. 59 with respect to the variables $\bar{\boldsymbol{x}}$, $\bar{\bar{\boldsymbol{x}}}$ for fixed $\bar{\boldsymbol{\theta}}^1$, $\boldsymbol{\theta}^2$ and $\boldsymbol{z} = \boldsymbol{z}^*$:

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}, z^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0, \quad j = 1, \ldots, r, \quad (60)$$

Designate the solution of this set of nonlinear equations as $\bar{\boldsymbol{x}}(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2)$, $\bar{\bar{\boldsymbol{x}}}(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2)$. Since at the given time instant, the values of the state variables $\bar{\boldsymbol{x}}$ are equal to $\bar{\boldsymbol{x}}^*$, then the following equality must be met:

$$\bar{\boldsymbol{x}}(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = \bar{\boldsymbol{x}}^*$$

Hence, region $\bar{T}^2$

will have the following form:

$$\bar{T}^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^2 \in T^2, \quad \bar{\boldsymbol{x}}(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) - \bar{\boldsymbol{x}}^* = 0\}. \quad (61)$$

So at each time instant at the operation stage, we will solve Eq. 32 in which region $\bar{T}^2$ has the form of Eq. 61.

Consider now a formulation of the two-stage optimization problem that should be solved at the design stage. Using the same reasoning as in the previous case, one can show that TSOP4 has the form of Eq. 44. In addition, the region $T^2$ (Eq. 61) is used in the internal optimization problem (Eq. 32) and in the feasibility function (Eq. 39). However, here we encounter the same difficulty as in the previous case, namely, we use the value $\bar{\boldsymbol{x}}^*$, which is not known at the design stage, in the expression for $\bar{T}^2$ In connection with this, we will obtain the value $\bar{\boldsymbol{x}}^*$ by solving Eq. 59 for $\boldsymbol{\theta}^1 = \bar{\boldsymbol{\theta}}^1$, $\boldsymbol{\theta}^2 = \bar{\boldsymbol{\theta}}^2$ , and $\boldsymbol{z} = \boldsymbol{z}^*$:
:

$$\varphi_j(\boldsymbol{d}, \bar{x}, \bar{\bar{x}}, z^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = 0, \quad j = 1, \ldots, r, \quad (62)$$

Designate as $\bar{\boldsymbol{x}}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2, \bar{\bar{\boldsymbol{x}}}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ the solution of Eq. 62. Substituting $\bar{\boldsymbol{x}}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ in Eq. 61 instead of $\bar{\boldsymbol{x}}^*$, we obtain the following expression for $\bar{T}^2$:

$$\bar{T}^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^2 \in T^2, \quad \bar{x}(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) - \bar{x}^*(\boldsymbol{d}, z^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = 0\} \quad (63)$$

Thus, similarly to the previous case, we replace the future measurement of the plant data by the simulation of the chemical process. So the two-stage optimization problem is completely defined.

Comparing TSOP1, TSOP2, and TSOP4, we note that fullness and accuracy of plant data available at the operation stage influences the design of the chemical process and taking into account this information permits creating a more economical process.

## Split and Bound Method of Solving Two-Stage Optimization Problem

To solve TSOP4 we will modify the split and bound (SB) method developed for solving TSOP2.[1] For the sake of simplicity, we will use the simplified discrete variant of TSOP4 (Eq. 51) in which we will omit the tilde in $\tilde{f}_4$. We will suppose that mathematical Eq. 6 can be represented as Eq. 22. The SB algorithm is a two-level iterative procedure that employs a partitioning of $T^1$ into subregions $T_i^{(1,(k)}$, where $k$ is an index of the iteration of the SB method and the subscript $i$ is the index of a subregion. Suppose that at the $k$-th iteration, there are $N_k$ subregions $T_i^{1,(k)}$. Let $T^{(k)}$ be the set of the subregions $T_i^{1,(k)}$, $i = 1, \ldots, N_k$. Partitioning of some subregions is carried out on the upper level.

The lower level is used to calculate upper and lower bounds of $f_4$. The split and bound method has the following three main operations:

1. The algorithm for estimating an upper bound of the objective function of TSOP4.
2. The algorithm for estimating a lower bound of the objective functions of TSOP4.
3. The partitioning strategy.

Let us start with obtaining an expression for an upper bound of the feasibility function. The feasibility function can be represented as:

$$\chi_4(\boldsymbol{d}) = \max_{\bar{\boldsymbol{\theta}}\in T} \psi^{(4)}(d, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \quad (64)$$

where $\bar{\boldsymbol{\theta}} = (\bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ and

$$\psi^{(4)}(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = \min_{z}\max_{\boldsymbol{\theta}^2\in\bar{T}^2}\max_{j\in J} g_j(\boldsymbol{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \quad (65)$$

Introduce the auxiliary functions $\chi_4^U(\boldsymbol{d})$ and $\chi_4^{U,i}(\boldsymbol{d})$ for the region $T^1$ and its subregions $T_i^1$ :

$$\chi_4^U = \min_z \max_{\bar{\boldsymbol{\theta}}^1 \in T^1} \max_{\bar{\boldsymbol{\theta}}^2 \in \bar{T}^2} \max_{\boldsymbol{\theta}^2 \in \bar{T}^2} \max_{j \in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \tag{66}$$

$$\chi_4^{U,i} = \min_z \max_{\bar{\boldsymbol{\theta}}^1 \in T_i^1} \max_{\bar{\boldsymbol{\theta}}^2 \in \bar{T}^2} \max_{\boldsymbol{\theta}^2 \in \bar{T}^2} \max_{j \in J} g_j(\boldsymbol{d}, \boldsymbol{z}, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \tag{67}$$

The function $\chi_4^U(\boldsymbol{d})$ is obtained by the rearrangement of the operator $\min_z$ and the operator $\max_{\bar{\boldsymbol{\theta}}^1 \in T^1} \max_{\bar{\boldsymbol{\theta}}^2 \in T^2}$ in Eq. 39. Similarly, $\chi_4^{U,i}(\boldsymbol{d})$ is obtained by the rearrangement of the operator $\min_z$ and the operator $\max_{\bar{\boldsymbol{\theta}}^1 \in T_i^1} \max_{\bar{\boldsymbol{\theta}}^2 \in T^2}$ in Eq. 40. Consider some properties of $\chi_4^U(\boldsymbol{d})$ and $\chi_4^{U,i}(\boldsymbol{d})$ .

*Property 1.* The function $\chi_4^U(\boldsymbol{d})$ is an upper bound of $\chi_4(\boldsymbol{d})$ on $T^1$.

$$\chi_4^U(\boldsymbol{d}) \ge \chi_4(\boldsymbol{d}) \tag{68}$$

Indeed, the function $\chi_4^U(\boldsymbol{d})$ is obtained by the rearrangement of the operator $\min_z$ and the operator $\max_{\bar{\boldsymbol{\theta}}^1 \in T^1} \max_{\bar{\boldsymbol{\theta}}^2 \in T^2}$ in Eq. 39. Hence, according to inequality A.4, we obtain Eq. 68. Taking into account Eq. 64, we have:

$$\chi_4^U(\boldsymbol{d}) \ge \psi^{(4)}(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \quad \forall\, \bar{\boldsymbol{\theta}}^1 \in T^1, \bar{\boldsymbol{\theta}}^2 \in T^2 \tag{69}$$

Similarly, for $\chi_2^{U,i}$, we have

$$\chi_4^{U,i}(\boldsymbol{d}) \ge \psi^{(4)}(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \quad \forall\, \bar{\boldsymbol{\theta}}^1 \in T_i^1, \bar{\boldsymbol{\theta}}^2 \in T^2 \tag{70}$$

*Property 2.* Let $T^1$ be partitioned into $N_k$ subregions $T_i^{1,(k)}$ :

$$T^1 = T_1^{1,(k)} \cup T_2^{1,(k)} \cup \ldots T_{N_k}^{1,(k)} \tag{71}$$

Then there exists the following inequality:

$$\max_i \chi_4^{U,i} \ge \chi_4(\boldsymbol{d}) \tag{72}$$

Indeed, since Eq. 70 holds for all the subregions $T_i^1$, then

$$\max_i \chi_4^{U,i}(\boldsymbol{d}) \ge \psi^{(4)}(\boldsymbol{d}, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) \quad \forall\, \bar{\boldsymbol{\theta}}^1 \in T^1, \bar{\boldsymbol{\theta}}^2 \in T^2$$

Using Eq. 64, we obtain Eq. 72. Thus, the value $\chi_4^{U,i}$ is an upper bound of $\chi_4(\boldsymbol{d})$ in $T^1$.

## Upper Bound of TSOP4 Objective Function

Let the region $T^1$ be partitioned into $N_k$ subregions $T_1^{1,(k)}$ (see Eq. 71). Replace Eq. 52 in Eq. 51 with the constraints $\chi_4^{U,1}(\boldsymbol{d}) \le 0, \ldots \chi_4^{U,N_k}(\boldsymbol{d}) \ge 0$. We obtain the following problem:

$$f_4^{U,(k)} = \min_{d, z^{IL}} \sum_{i \in I_1} w_i \sum_{l \in L_1} v_l f(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \bar{\boldsymbol{\theta}}^{2l}) \tag{73}$$

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}_{il}^2} g_j(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \boldsymbol{\theta}^2) \le 0 \quad j = 1, \ldots m; \quad i \in I_1, l \in L_1$$

$$\chi_4^{U,1}(\boldsymbol{d}) \le 0, \ldots \chi_4^{U,N_k}(\boldsymbol{d}) \le 0 \tag{74}$$

where $\boldsymbol{z}^{IL} = \{\boldsymbol{z}^{il}, \forall i \in I_1, \forall l \in L_1\}$ and $\chi_4^{U,i}(\boldsymbol{d})$ is determined by Eq. 67. Since Eq. 72 is met, then the feasibility region of Eq. 73 is contained in the feasibility region of TSOP4 (see Eq. 51). Consequently, the following inequality holds:

$$f_4^{U,(k)} \ge f_4 \tag{75}$$

Thus, the value $f_4^{U,(k)}$ is an upper bound of the optimal value of the objective function in TSOP4.

Substituting the expression for $\chi_4^{U,i}$ into Eq. 73, we obtain:

$$f_4^{U,(k)} = \min_{d, z^{IL}} \sum_{i \in I_1} w_i \sum_{l \in L_1} v_l f(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \bar{\boldsymbol{\theta}}^{2l}) \tag{76}$$

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}_{il}^2} g_j(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \boldsymbol{\theta}^2) \le 0 \quad j = 1, \ldots m; \quad i \in I_1, l \in L_1$$

$$\min_z \max_{\bar{\boldsymbol{\theta}}^1 \in T_s^1} \max_{\bar{\boldsymbol{\theta}}^2 \in T^2} \max_{j \in J} \max_{\boldsymbol{\theta}^2 \in \bar{T}^2} g_j(\mathbf{d}, z, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0, \quad s = 1, \ldots, N_k \tag{77}$$

where $\bar{T}^2$ in Eq. 77 has the following form:

$$\bar{T}^2 = \{\boldsymbol{\theta}^2 : \boldsymbol{\theta}^2 \in T^2, \varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0, j \in J_1\} \tag{78}$$

Here $\bar{\boldsymbol{x}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ together with $\bar{\bar{\boldsymbol{x}}}^*(\boldsymbol{d}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2)$ is the solution of the following set of nonlinear equations:

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = 0, \quad j \in J_1$$

$$\varphi_j(\boldsymbol{d}, \bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}, \boldsymbol{z}^*, \bar{\boldsymbol{\theta}}^1, \bar{\boldsymbol{\theta}}^2) = 0, \quad j = 1, \ldots, r, \quad j \ne J_1$$

Transform Eq. 76 using Theorem A.1. Adding new search variables $z^s$, $s = 1, \ldots, N_k$ (the vector $z^s$ of the control variables corresponds to the $s$-th subregion), we can transform Eq. 76 into the form:

$$f_4^{U,(k)} = \min_{d, z^{IL}, z^S} \sum_{i \in I_1} w_i \sum_{l \in L_4} v_l f(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \bar{\boldsymbol{\theta}}^{2l}) \tag{79}$$

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}_{il}^2} g_j(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \boldsymbol{\theta}^2) \le 0 \quad j = 1, \ldots m; \quad i \in I_1, l \in L_1$$

$$\max_{\bar{\boldsymbol{\theta}}^1 \in T_s^1} \max_{\bar{\boldsymbol{\theta}}^2 \in T^2} \max_{j \in J} \max_{\boldsymbol{\theta}^2 \in \bar{T}^2} g_j(\boldsymbol{d}, z^s, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) \le 0, \quad s = 1 \ldots, N_k \tag{80}$$

where $z^s$ is the set of the control variables vectors corresponding to all the subregions $T_i^{1,(k)} : \boldsymbol{z}^S = \{z^s, s = 1, \ldots, N_k\}$.

According to relation (A.3) Eq. 80 can be reduced to the following form:

$$\max_{j\in J}\max_{\bar{\boldsymbol\theta}^1\in T_s^1}\max_{\bar{\boldsymbol\theta}^2\in T^2}\max_{\boldsymbol\theta^2\in\bar T^2} g_j(\boldsymbol d, z^s, \bar{\boldsymbol\theta}^1, \boldsymbol\theta^2) \le 0, \quad s = 1\ldots, N_k$$

Using relation (A.1) we can transform these inequalities into the following ones:

$$\max_{\bar{\boldsymbol\theta}^1\in T_s^1}\max_{\bar{\boldsymbol\theta}^2\in T^2}\max_{\boldsymbol\theta^2\in\bar T^2} g_j(\boldsymbol d, z^s, \bar{\boldsymbol\theta}^1, \boldsymbol\theta^2) \le 0, \quad s = 1\ldots, N_k,$$

$$j = 1, \ldots, m \quad (81)$$

Replacing Eq. 80 in Eq. 79 with Eq. 81, we obtain the final form of *the upper bound problem for TSOP4:*

$$f_4^{U,(k)} = \min_{\boldsymbol d, \boldsymbol z^{lL}, \boldsymbol z^S} \sum_{i\in I_1} w_i \sum_{l\in L_1} v_l f(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \bar{\boldsymbol\theta}^{2l}) \quad (82)$$

$$\max_{\boldsymbol\theta^2\in\bar T_{il}^2} g_j(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \boldsymbol\theta^2) \le 0 \quad j = 1, \ldots m; \quad i\in I_1, l\in L_1$$

$$\max_{\bar{\boldsymbol\theta}^1\in T_s^1}\max_{\bar{\boldsymbol\theta}^2\in T^2}\max_{\boldsymbol\theta^2\in\bar T^2} g_j(\mathbf d, z^s, \bar{\boldsymbol\theta}^1, \boldsymbol\theta^2) \le 0,$$

$$s = 1\ldots, N_k, j = 1, \ldots, m \quad (83)$$

Let $d^{(k)}$, $z^{il,(k)}$, $z^{s,(k)}$ be the solution of Eq. 82. It is the semi-infinite programming problem. To solve it, we will use the outer approximation algorithm ([20]) (see more detail in ref[1]).

## Lower Bound of TSOP4 Objective Function

Taking into account Eq. 64 of $\chi_4(d)$ and using relation (A1), we can reduce Eq. 51 to the form:

$$f_4^{L,(k)} = \min_{\boldsymbol d, \boldsymbol z^{lL}} \sum_{i\in I_1} w_i \sum_{l\in L_1} v_l f(\mathbf d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \bar{\boldsymbol\theta}^{2l}) \quad (84)$$

$$\max_{\boldsymbol\theta^2\in\bar T_{il}^2} g_j(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1i}, \boldsymbol\theta^2) \le 0 \quad j = 1, \ldots m; \quad i\in I_1, l\in L_1$$

$$\psi^{(4)}(d, \bar{\boldsymbol\theta}^1, \bar{\boldsymbol\theta}^2) \le 0 \quad \forall\, \bar{\boldsymbol\theta}^1 \in T^1, \bar{\boldsymbol\theta}^2 \in T^2 \quad (85)$$

This is an optimization problem with an infinite number of constraints. Let $d^*$, $z^{il*}$ be the solution of Eq. 84. Introduce some arbitrary set $S^{(k)} = \{\bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p};\ r\in R^{(k)},\ p\in P^{(\mathrm{k})}\}$ of points in the region $T$, where $R^{(k)}$ and $P^{(\mathrm{k})}$ are some sets of indices. These points will be referred to as *critical points*. Consider the following problem:

$$f_4^{L,(k)} = \min_{\boldsymbol d, \boldsymbol z^{lL}} \sum_{i\in I_1} w_i \sum_{l\in L_1} v_l f(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \bar{\boldsymbol\theta}^{2l}) \quad (86)$$

$$\max_{\boldsymbol\theta^2\in\bar T_{il}^2} g_j(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \boldsymbol\theta^2) \le 0 \quad j = 1, \ldots m; \quad i\in I_1, l\in L_1$$

$$\psi^{(4)}(d, \bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p}) \le 0 \quad \forall\, (\bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p}) \in S^{(k)}$$

In contrast to Eq. 84, this problem has a finite number of constraints. Since $S^{(k)} \in T$, , the feasible region of Eq. 84 is contained within the feasible region of Eq. 86. Therefore:

$$f_4^{L,(k)} \le f_4 \quad (87)$$

Thus, the value $f_4^{L,(k)}$ is a lower bound of the optimal value of the TSOP4 objective function. Substitute Eq. 65 of $\psi^{(4)}(\boldsymbol d, \bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p})$ into Eq. 84:

$$f_4^{L,(k)} = \min_{\boldsymbol d, \boldsymbol z^{lL}} \sum_{i\in I_1} w_i \sum_{l\in L_1} v_l f(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \bar{\boldsymbol\theta}^{2l}) \quad (88)$$

$$\max_{\boldsymbol\theta^2\in\bar T_{il}^2} g_j(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \boldsymbol\theta^2) \le 0 \quad j = 1, \ldots m; \quad i\in I_1, l\in L_1$$

$$\min_{z}\max_{\boldsymbol\theta^2\in\bar T_{rp}^2}\max_{j\in J} g_j(\boldsymbol d, \boldsymbol z, \bar{\boldsymbol\theta}^{1,r}, \boldsymbol\theta^2) \le 0 \quad (\bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p}) \in S^{(k)}$$

where

$$\bar T_{rp}^2 = \{\boldsymbol\theta^2 : \boldsymbol\theta^2 \in T^2, \varphi_j(d, x^*(\boldsymbol d, \boldsymbol z^*, \bar{\boldsymbol\theta}^{1r}, \bar{\boldsymbol\theta}^{2p}), \boldsymbol z^*, \bar{\boldsymbol\theta}^{1r}, \boldsymbol\theta^2) = 0, \quad j\in J_1\}$$

Using Theorem A.1 we can transform Eq. 88 to the form:

$$f_4^{L,(k)} = \min_{\boldsymbol d, \boldsymbol z^{lL}, \boldsymbol z^R} \sum_{i\in I_1} w_i \sum_{l\in L_1} v_l f(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \bar{\boldsymbol\theta}^{2l}) \quad (89)$$

$$\max_{\boldsymbol\theta^2\in\bar T_{il}^2} g_j(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \boldsymbol\theta^2) \le 0 \quad j = 1, \ldots m; \quad i\in I_1, l\in L_1$$

$$\max_{j\in J}\max_{\boldsymbol\theta^2\in\bar T_{rp}^2} g_j(\boldsymbol d, \boldsymbol z^r, \bar{\boldsymbol\theta}^{1,r}, \boldsymbol\theta^2) \le 0,$$

$$j = 1, \ldots, m \quad (\bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p}) \in S^{(k)} \quad (90)$$

where $z^r$ is the control variables vector corresponding to the point $\bar{\boldsymbol\theta}^{1,r}$, and $\boldsymbol z^R$ is the set of the control variables vectors $z^r$ that correspond to all the points $\bar{\boldsymbol\theta}^{1,r}$. Using relation (A.1) we can transform the following inequalities:

$$\max_{j\in J}\max_{\boldsymbol\theta^2\in\bar T_{rp}^2} g_j(\boldsymbol d, \boldsymbol z^r, \bar{\boldsymbol\theta}^{1,r}, \boldsymbol\theta^2) \le 0, \quad (\bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p}) \in S^{(k)}$$

to the following form:

$$\max_{\boldsymbol\theta^2\in\bar T_{rp}^2} g_j(\boldsymbol d, \boldsymbol z^r, \bar{\boldsymbol\theta}^{1,r}, \boldsymbol\theta^2) \le 0, \quad j = 1, \ldots, m \quad (\bar{\boldsymbol\theta}^{1,r}, \bar{\boldsymbol\theta}^{2,p}) \in S^{(k)}$$

$$(91)$$

Replacing Eq. 90 in Eq. 89 with Eq. 91, we obtain the final form of *the lower bound problem for TSOP4:*

$$f_4^{L,(k)} = \min_{\boldsymbol d, \boldsymbol z^{lL}, \boldsymbol z^R} \sum_{i\in I_1} w_i \sum_{l\in L_1} v_l f(\boldsymbol d, \boldsymbol z^{il}, \bar{\boldsymbol\theta}^{1,i}, \bar{\boldsymbol\theta}^{2l}) \quad (92)$$

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}^2_{il}} g_j(\boldsymbol{d}, \boldsymbol{z}^{il}, \bar{\boldsymbol{\theta}}^{1,i}, \boldsymbol{\theta}^2) \le 0 \quad j = 1, \dots m; \quad i \in I_1, l \in L_1$$

$$\max_{\boldsymbol{\theta}^2 \in \bar{T}^2_{rp}} g_j(\boldsymbol{d}, \boldsymbol{z}^{r}, \bar{\boldsymbol{\theta}}^{1,r}, \boldsymbol{\theta}^2) \le 0, \quad j = 1, \dots, m \quad (\bar{\boldsymbol{\theta}}^{1,r}, \bar{\boldsymbol{\theta}}^{2,p}) \in S^{(k)}$$

We will use the set of critical points obtained by solving the upper bound problem (Eq. 76) (see ref. [1]) as $S^{(k)}$
.

## The Partitioning Strategy

We apply the following heuristic rule of partitioning: at the $k$-th iteration, only those subregions $T_i^{(k)}$ $(i = 1, \dots, N_k)$ are partitioned, for which Eq. 74 in upper bound Eq. 73 is active. Since Eq. 82 is equivalent to Eq. 73, the heuristic is as follows: at the $k$-th iteration, the subregion $T_l^1$ is partitioned if the following condition is met:

$$\exists\, j \in J \max_{\bar{\boldsymbol{\theta}}^1 \in T_l^1} \max_{\bar{\boldsymbol{\theta}}^2 \in T^2} \max_{\boldsymbol{\theta}^2 \in \bar{T}^2} g_j(\boldsymbol{d}, z^l, \bar{\boldsymbol{\theta}}^1, \boldsymbol{\theta}^2) = 0 \tag{93}$$

where $J = (1, \dots, m)$.

## Algorithm of Solving TSOP4

*Step 1.* Set $k = 1$. Choose an initial partition of $T$ into subregions $T_i^{(1)}$, $(i = 1, \dots, N_1)$, initial value $u^{(0)}$, $z^{l,(0)}$ $(l = 1, \dots, N_1)$, and a small value $\varepsilon_1 > 0$. Put $f_1^{(0)} = -a$, where $a$ is a large enough value $(a > -f_1)$.

*Step 2.* Solve upper bound Eq. 73.

Let $[\boldsymbol{d}^{(k)}, \boldsymbol{z}^{l,(k)}]$ $(l = 1, \dots, N_k)$ be the solution.

*Step 3.* Determine the set $(Q^{(k)} = \{T_l^{(k)} : l \in I_Q^{(k)}\}$ of subregions $T_i^{(k)}$ having active constraints in upper bound Eq. 73:

$$\chi_4^{U,i}(\boldsymbol{d}^{(k)}) = 0, \quad T_l^{(k)} \in Q^{(k)}$$

*Step 4.* If $Q^{(k)}$ is empty, then the solution of TSOP4 is found.[1] STOP.

*Step 5.* Calculate the lower bound by solving Eq. 92.

*Step 6.* If

$$f_4^{U,(k)} - f_4^{L,(k)} \le \varepsilon_1 |f_4^{U,(k)}| \tag{94}$$

then the solution of TSOP4 is found. STOP.

*Step 7.* Partition each subregion $T_l^{(k)} \in Q^{(k)}$ into two subregions such that $T_{l_1}^{(k+1)} \cup T_{l_2}^{(k+1)} = T_l^{(k)}$. Form a new collection $T^{(k+1)}$ of subregions from the old collection $T^{(k)}$ by replacing each subregion $T_l^{(k)}$ with $T_{l_1}^{(k+1)}$ and $T_{l_2}^{(k+1)}$.

*Step 8.* Set $k = k + 1$ and go to Step 2.

In Step 7, the subregion $T_l^{(k)}$ $(T_l^{(k)} \in Q^{(k)})$ is partitioned into two subregions $T_{l_1}^{(k+1)}$ and $T_{l_2}^{(k+1)}$, such that one subregion has the constraint $\theta_s \le a_s$ and another has the conflicting constraint $\theta_s \ge a_s$:

$$T_{l_1}^{(k+1)} = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in T_l^{(k)}, \theta_s \le a_s\}, T_{l_2}^{(k+1)} = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in T_l^{(k)}, \theta_s \ge a_s\}$$

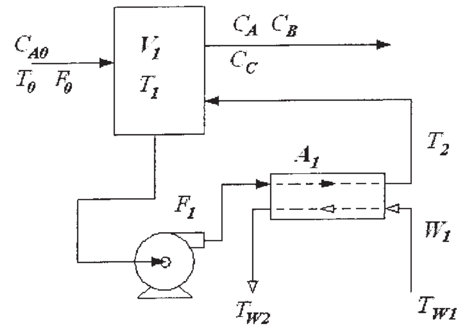


**Figure 1. Reactor—heat exchanger flowsheet.**

Parameter $a_s$ is chosen as $a_s = 0.5(\bar{\theta}_s + \bar{\bar{\theta}}_s)$, where $\bar{\theta}_s$ and $\bar{\bar{\theta}}_s$ are the lower and upper bounds of $\theta_s$ in the subregion. We refer to $\theta_s$ as a *branching variable*. Each $\theta_i$ will take a turn at being the branching variable.

Generally speaking, the SB method obtains only a local minimum of TSOP4 when a local optimization solver is used for solving upper and lower bound Eqs. 82 and 92. To obtain the global minimum, the following conditions should be satisfied[21]:

1. Each $g_j(\boldsymbol{d}, z, \boldsymbol{\theta})$ is quasi-concave in $\boldsymbol{\theta}$.
2. Each $g_j(\boldsymbol{d}, z, \boldsymbol{\theta})$ is quasi-convex in $\boldsymbol{d}$ and $z$.
3. The function $f(\boldsymbol{d}, z, \boldsymbol{\theta})$ is convex in $\boldsymbol{d}$ and $z$.

## Computational Experiments

We consider two examples. The uncertainty region is given as

$$T(\gamma) = \{\boldsymbol{\theta}_i : \boldsymbol{\theta}_i^N(1 - \gamma\delta\boldsymbol{\theta}_i) \le \boldsymbol{\theta}_i \le \boldsymbol{\theta}_i^N(1 + \gamma\delta\boldsymbol{\theta}_i)\} \tag{95}$$

where $\theta_i^N$ is a nominal value of the parameter $\theta_i$, $\delta\theta_i$ is a maximal deviation from the nominal value, and $\gamma$ is a scalar coefficient. For solving NLP problems in the algorithms for calculation of the upper and lower bounds, we employed CFSQP.[22] All the examples were solved on a 2GHz Pentium 4 PC.

### *Example 1: Continuous stirred-tank reactor (CSTR) and external heat exchanger*

Suppose we have to design a CP consisting of a CSTR and an external countercurrent heat exchanger[3] (Figure 1).

The reaction is assumed to be first-order exothermal of the type $A \rightarrow B \rightarrow C$ with $B$ as a product. The flow rate through the heat exchanger loop is adjusted to maintain the reactor temperature $T_1$ below some $T_{1max}$. The process model of the reactor (material and energy balance) is:

$$\begin{aligned} \rho c_p F_0(T_0 - T_1) + k_1 \exp(-E_A/RT_1) C_A(-\Delta H_1) V \\ + k_2 \exp(-E_B/RT_1) C_B(-\Delta H_2) V - Q_{HE} = 0 \\ F_0(C_{A0} - C_A) - k_1 \exp(-E_A/RT_1) C_A V = 0 \\ F_0(C_{B0} - C_B) + k_1 \exp(-E_A/RT_1) C_A V \\ - k_2 \exp(-E_B/RT_1) C_B V = 0 \end{aligned} \tag{96}$$

while for the heat exchanger (heat balance and design equations) the process model is:

$$Q_{HE} = F_1 C_p(T_1 - T_2) = C_{pw}(T_{w2} - T_{w1})W \tag{97}$$

$$Q_{HE} = AU\frac{(T_1 - T_{w2}) - (T_2 - T_{w1})}{\ln\{(T_1 - T_{w2})/(T_2 - T_{w1})\}} \tag{98}$$

Here $F_0$, $T_0$, $C_{A0}$, and $C_{B0}$ are the feed flow rate, temperature of the feed, and the concentrations of the reactants A and B in the feed, respectively. $V$, $T_1$, $C_A$, and $C_B$ are the reactor volume, the reactor temperature, and the concentrations of the reactants $A$ and $B$ in the product, respectively; $E_A$ and $E_B$ are activity energies; $\Delta H_1$ and $\Delta H_2$ are the heats of reaction; $F_1$ is the flow rate of the recycle; $T_2$ is the recycle temperature; and $C_p$ and $C_{pw}$ are the heat capacities of the recycle mixture and the cooling water, respectively. $T_{w1}$, $T_{w2}$, and $W$ are the inlet and outlet temperatures and the flow rate of the cooling water, respectively. $A$ is the heat transfer area of the heat exchanger, and $U$ is the overall heat transfer coefficient.

We select two design variables, $V$ and $A$, two control variables ($T_1$ and $T_{w2}$), five state variables $x = (C_A, C_B, T_2, F_1, W)$, and eight uncertain parameters $\theta = (F_0, T_0, T_{w1}, k_1, k_2, U, \theta = (F_0, T_0, T_{w1}, k_1, k_2, U, E_A/R, E_B/R)$. The nominal values of the uncertain parameters are $\theta^N = [45.36, 333, 300, 314, 40, 1635, 560, 500]$ and $\delta\theta_i = [0.08, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1]$. Specific data for the example are in ref. [1]. There are the following constraints in this problem:

| | |
|---|---|
| $0.9 \le conv$ | reactor conversion |
| $311 \le T_1 \le 389$ | reactor temperature |
| $301 \le T_{w2} \le 355$ | heat exchanger |
| $-T_2 + 311 \le 0$ | heat exchanger |
| $-(T_2 - T_{w1}) + 11.1 \le 0$ | minimum temperature approach |
| $T_2 - T_1 \le 0$ | thermodynamic constraint for heat exchanger |
| $-T_{w2} + T_{w1} \le 0$ | thermodynamic constraint for heat exchanger |

(99)

where $conv = (C_{A0} - C_A)/C_{A0}$.

The performance objective (income) $f$ (\$/yr) is given by:

$$f = 100F_0C_B - (345.6V^{0.7} + 436.8A^{0.6} + 0.88W + 3.528F_1 + 12.5(\exp 0.3C_C - 1)) \tag{100}$$

Here the first term is the cost of the useful product, the second and third terms are the capital costs of the reactor and the heat exchanger, the fourth term is the cost of cold water, the fifth term is the cost of the pumping recycle, and the sixth term is the cost of separation.[23]

The goal of the design is to find such values of the design variables $V$ and $A$ that guarantee that at the operation stage: (1) the average value of cost function (100) will be maximal; (2) the chemical process will be flexible, that is, for each value of the vector of the uncertain parameters $\theta = (F_0, T_0, T_{w1}, k_1, k_2, U, U, E_A/R, E_B/R) \in T$ one will be able to find such values of the control variables $T_1$ and $T_{w2}$ that all constraints (Eq. 99) will be met.

For the sake of simplicity, instead of Eq. 70 we use the following approximate model of the heat exchanger:

$$Q_{HE} = AU\frac{(T_1 - T_{w2}) + (T_2 - T_{w1})}{2} \tag{101}$$

In order to obtain reduced objective Eq. 9 and reduced constraints Eq. 5, we have to exclude the state variables $C_A$, $C_BT_2$, $F_1$, $W$. After simple transformations of Eqs. 96, 97, and 101, we can obtain analytical formula Eq. 9 connecting the state variables with the design variables, control variables, and uncertain parameters. Now these formulas have the following form:

$$C_A = \frac{F_0}{F_0 + Vk_1C_{A0}\exp(-E/RT_1)} \tag{102}$$

$$C_B = \frac{F_0C_{B0} + Vk_1\exp(-E_A/RT_1)C_A}{F_0 + Vk_2\exp(-E/RT_1)}$$

$$T_2 = \frac{2Q_{HE}}{AU} - (T_1 - T_{w2}) + T_{w1}$$

$$F_1 = \frac{Q_{HE}}{c_p(T_1 - T_2)} \quad W = \frac{Q_{HE}}{c_{pw}(T_{w2} - T_{w1})}$$

where

$$Q_{HE} = \rho c_p F_0(T_0 - T_1) + k_1\exp(-E_A/RT_1)C_A(-\Delta H_1)V + k_2\exp(-E_B/RT_1)C_B(-\Delta H_2)V$$

Using these formulas we can calculate in series the values of $C_A$, $C_B$, $T_2$, $F_1$, $W$ if we know the values of the design variables, control variables, and uncertain parameters.

Substituting the expressions for $C_A$, $C_B$, $T_2$, $F_1$, and $W$ into Eqs. 99 and 100, we obtain reduced objective Eq. 9 and reduced process constraints Eq. 5.

We consider three cases. In the first case, we suppose that all the uncertain parameters can be determined with sufficient precision at the operation stage. In the second case, we suppose that there are sensors only for the measurement of the parameters $T_0$, $T_{w1}$ and there is no sensor for the measurement of the input flow rate $F_0$. In addition, we take into account that the parameters $k_1$, $k_2$, $U$, $E_A/R$, and $E_B/R$ cannot be measured. In this case, for the optimal design of the chemical process, we should solve TSOP2 in which $\theta^1 = [T_0, T_{w1}]$ and $\theta^2 = [k_1, k_2, F_0, U, E_A/R, E_B/R]$. In the third case, we suppose that at the operation stage we can measure the state variables $T_1$, $C_A$, and $C_B$. It is clear that we cannot determine the uncertain parameters $F_0$, $k_1$, $k_2$, $U$, $E_A$, and $E_B$ using this plant data. Therefore, if we use TSOP2, we will not be able to use this plant data. However, using TSOP4 we can take into account this plant data. In this case, $\bar{x} = (T_1, C_A, C_B)$ and we cannot extract a set of equations that depend only on the measured

**Table 1. Results for Nominal Optimization, TSOP1, and TSOP2 for Example 1**

| | $F$, \$/yr. | $V$, m$^3$ | $A$, m$^2$ | $\Delta V$, % | $\Delta A$, % |
|---|---|---|---|---|---|
| Nominal | 40231 | 5.9 | 32.4 | | |
| TSOP1 | 27512 | 7.3 | 40.3 | 32 | 24 |
| TSOP2 | 26253 | 7.7 | 42.3 | 40 | 31 |
| TSOP4 | 27040 | 7.4 | 41 | 35 | 27 |

state variables $\bar{x}$. Therefore, the region $\bar{T}^2$ is determined by Eq. 45. For the sake of simplicity, we will solve Eq. 38.

We suppose that probability density functions of the uncertain parameters are not known. In this case a user should give the approximation points and weight coefficients. We used the technique from ref. [3] to select the approximation points and weight coefficients. According to it, the nominal point and some corner points of the uncertainty region are used as the approximation points. We used four approximation points for the vector $\theta^1$ : $\theta^{1,1} = N, N$; $\theta^{1,2} = L, L$; $\theta^{1,3} = U, U$; $\theta^{1,4} = L, U$, where $N$, $U$, and $L$ denote the nominal value, and upper and lower bounds, respectively. For each approximation point for $\theta^1$, we employed five approximation points for $\theta^2$ ($\theta^{2,1} = N, N, N, N, N, N$; $\theta^{2,2} = L, U, U, L, L, L$; $\theta^{2,3} = L, L, L, L, L, L$; $\theta^{2,4} = U, L, L, U, U, U$; $\theta^{2,5} = U, U, U, U, U, U$). Thus, in the space of the parameters $\theta$ , we employ 20 approximation points. The following weight coefficients were used: first group $w_1 = 0.7$; $w_2 = w_3 = w_4 = 0.1$; second group $v_1 = 0.6$; $v_2 = v_3 = v_4 = v_5 = 0.1$.

In Table 1, we give the results for the nominal optimization, TSOP1, TSOP2, and TSOP4. According to TSOP1, we must use the following design margins to obtain a flexible CP: $\Delta V = 32\%$, $\Delta A = 24\%$. TSOP2 suggests the following design margins: $\Delta V = 40\%$, $\Delta A = 31\%$. On the other hand according to TSOP4, we must use the following design margins: $\Delta V = 35\%$, $\Delta A = 27\%$. Besides, in comparison with TSOP2, the use of TSOP4 permits increasing the optimal value of the objective function from 26253 to 27040 (3.0%).

### *Example 2: Three-stage flowsheet*

Consider the reactor/heat exchanger network in Figure 2. Each stage has one CSTR of the volume $V_i$ and the heat exchanger with the heat exchange area $A_i$. In each CSTR the reaction is:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

$$B \xrightarrow{k_3} D$$

Here $B \rightarrow D$ is endothermic, while $A \rightarrow B$ and $B \rightarrow C$ are exothermic. Moreover, each reaction step is first order. The process models of the reactor and heat exchanger are given as: CSTR: for the $i$-th stage:

$$\rho c_p F_0(T_0^i - T_1^i) + k_1^i \exp(-E_A^i/RT^i)C_A^i(-\Delta H_1)V_i + k_2^i \exp(-E_B^{1,i}/RT)C_{B_i}^i(-\Delta H_2)V_i + k_3^i \exp(-E_B^{2,i}/RT)C_{B_i}^i(\Delta H_3)V_i - Q_{HE}^i = 0$$

$$F_0(C_{A0}^i - C_A^i) - k_1^i \exp(-E_A^i/RT^i)C_A^i V_i = 0$$

$$F_0(C_{B0}^i - C_{B_i}^i) + k_1^i \exp(-E_A^i/RT^i)C_A^i V_i - k_2^i \exp(-E_B^{1,i}/RT^i)C_B^i V_i - k_3^i \exp(E_B^{2,i}/RT^i)C_B^i V_i = 0$$

$$F_0(C_{D0}^i - C_D^i) + k_3^i \exp(E_B^{2,i}/RT^i)C_{B_i}^i V_i = 0 \qquad (103)$$

The design equations for the heat exchanger are given by Eqs. 97 and 101. Neighboring stages are connected by $T_0 = T_0^1$, $C_{j0} = C_{j0}^1$, $C_j^i = C_{j0}^{i+1}$, $T_0^{i+1} = T_i^1$, $j = A, B, C, D$, $i = 1, 2$ such that $C_{B0} = C_{C0} = C_{D0} = 0$. The cooling water temperature in each heat exchanger is $T_{w1}$. We employ the same notation as in Example 1. There are six design variables ($V_i$, $A_i$, $i = 1, 2, 3$); six control variables ($T_i^1$, $T_{w2}^i$, $i = 1, 2, 3$); 18 state variables $x = (C_A^i, C_{B_i}^i, C_D^i, T_2^i, F_1, F_2, F_3, W_1, W_2, W_3, i = 1, 2, 3)$ and 30 uncertain parameters $\theta = (T_0^1, T_{w1}, F_0, k_1^i, k_2^i, k_3^i, U_1^i, U_2^i, U_3^i, E_A^i/R, E_B^{1,i}/R, E_B^{2,i}/R, i = 1, 2, 3)$. Specific data for this example are given in ref. [1]. The deviations are given by $\delta\theta = [\delta\theta_1 = \delta\theta_2 = 03; .\delta\theta_i = 0.1, i = 2, \ldots, 30]$.

The production of $B$ should be at least 18.4 kmol/mg. Thus,

$$C_b^3 \geq 18.4 \qquad (104)$$

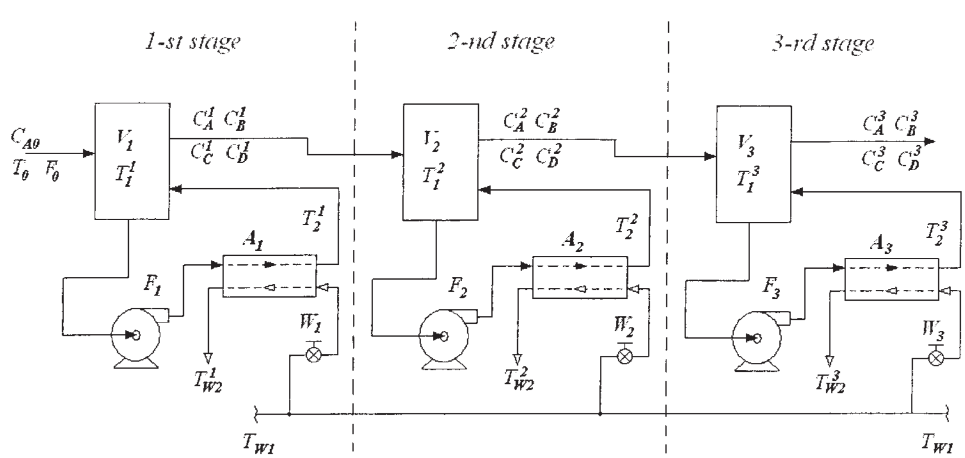


**Figure 2. Three-stage flowsheet.**

Table 2. Results for Nominal Optimization, TSOP1, and TSOP2 for Example 2

| | | $F$, \$/yr. | $V_1$, m$^3$ | $V_2$, m$^3$ | $V_3$, m$^3$ | $A_1$, m$^2$ | $A_2$, m$^2$ | $A_3$, m$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | Nominal | 14100 | 1.6 | 1.7 | 1.6 | 7.2 | 10.6 | 6.0 |
| 2 | TSOP1 | 17400 | 2.0 | 2.1 | 2.1 | 9.6 | 12.0 | 7.4 |
| 3 | Design margins, % | | 25 | 24 | 30 | 35 | 13 | 10 |
| 4 | TSOP2 | 18100 | 2.2 | 2.3 | 2.0 | 12.0 | 13.0 | 4.0 |
| 5 | Design margins, % | | 40 | 35 | 25 | 67 | 23 | −33 |
| 6 | TSOP4 | 17700 | 2.1 | 2.2 | 2.05 | 10.1 | 12.2 | 5.0 |
| 7 | Design margins, % | | 31 | 29 | 28 | 40 | 15 | −16 |

In addition, all the constraints in Eq. 99 (excluding $0.9 \leq conv$) are also enforced here. Thus,

$$-T_2^i + 311 \leq 0, \quad T_{w2}^i + T_{w1} \leq 0 \quad 301 \leq T_{w2}^i \leq 355, \quad i = 1, 2, 3. \quad (105)$$
$$-(T_2^i - T_{w1}) + 11.0 < 0, \quad T_2^i - T_1^i \leq 0, \quad 311 \leq T_1^i \leq 389$$

In this study it is required to minimize the averaged cost such that the production of $B$ should be at least 18.4 and all constraints in Eq. 105 are met for all possible values of the uncertain parameters. The capital cost includes the cost of the reactors and the heat exchangers, while the operating costs include the cost of the cold water, recycles, and separation of the byproducts $C$ and $D$ after the third stage. The performance objective is finally given by ref. [1]:

$$f = 700(V_1^{0.7} + V_2^{0.7} + V_3^{0.7}) + 174.72(A_1^{0.6} + A_2^{0.6} + A_3^{0.6}) + 1.76(W_1 + W_2 + W_3) + 7.056(F_1 + F_2 + F_3) + 100(\exp 0.5(C_C^3 + C_D^3) - 1$$

Similarly to the previous example, we can obtain explicit expressions for the state variables as functions of the design and control variables and the uncertain parameters. We will not bring here the expressions since they are cumbersome enough. Calculating all the stages in series (starting from the first stage), we can obtain the values of all the state variables. After that we can find the value of the reduced objective function and the left-hand sides of the reduced constraints. Thus, we have the case when analytical formula Eq. 8 is absent but we have the algorithm that permits calculating (without iterations) the values of all the state variables if the values of the design and control variables and the uncertain parameters are given. In the case of TSOP4, we supposed that at the operation stage we can measure the state variables:

$$T_1^i,\ C_A^i C_{B_i}^i,\ i = 1, 2, 3$$

We solved TSOP1, TSOP2, and TSOP4. In the TSOP2, $\theta^1 = (T_0, T_{w1})$ and $\theta^2 = (F_0, k_j^i, U_j^i, E_A^i, E_B^{1,i}, E_B^{2,i}, i = 1, 2, 3; j = 1, \ldots, 3)$. For $\theta^1$ we used five approximation points $[\theta^{1,1}, \theta^{1,2}, \theta^{1,3}, \theta^{1,4}, \theta^{1,5}] = [NN, UU, LL, LU, UL]$. Moreover, for each approximation point for $\theta^1$, we used one approximation point: $\theta^{2,1} = [N, \ldots, N]$. The following weight coefficients were used for the first group $w_1 = 0.6$, $w_2 = w_3 = w_4 = w_5 = 0.1$.

The results are given in Table 2. If at the operation stage values of all uncertain parameters can be determined accurately enough, then from TSOP1 we need to use the design margins given in row 3 of Table 2 in order to guarantee flexibility of the flowsheet. If at the operation stage the uncertain parameters $\theta^2$ cannot be corrected, then it is necessary to use the design margins from row 5. Here the total cost increases from \$17,400 to \$18,100 (increase of 4%). If at the operation stage we can measure the state variables $T_1^i$, $C_A^i C_{B_i}^i$, $i = 1, 2, 3$, then it is necessary to use the design margins from row 7. In comparison with TSOP2, the use of TSOP4 permits decreasing the total costs from 18100 to 17700 (decrease of 2.2%). Thus, taking into account the results of measurement of the state variables $T_1^i$, $C_A^i C_{B_i}^i$, $i = 1, 2, 3$ at the operation stage permits creating a less conservative chemical process.

## Conclusions

We have developed the new formulations of the feasibility test and the two-stage optimization problem. They are based on the following suppositions: We assume that the set $\boldsymbol{\theta}$ of the uncertain parameters can be split into two subsets. The first set $\bar{\boldsymbol{\theta}}$ consists of the parameters that can be measured at the operation stage, and the second set $\bar{\bar{\boldsymbol{\theta}}}$ consists of the parameters that cannot be measured at the operation stage. Similarly, the set $x$ of state variables can be split into two subsets. The first set $\bar{\boldsymbol{x}}$ consists of the state variables that can be measured at the operation stage, and the second set $\bar{\bar{\boldsymbol{x}}}$ consists of the parameters that cannot be measured at the operation stage. We suppose that the plant data $(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{x}})$ does not permit determining accurate values of the parameters $\bar{\bar{\boldsymbol{\theta}}}$. Formally in this case we can use TSOP2 in order to solve the optimization problem under uncertainty. However, in this case we do not use the plant data $(\bar{x})$, which may be available at the operation stage. In connection with this, we have developed the new forms of the feasibility test and the two-stage optimization problem, which enable taking into account the plant data. This allows creating a less conservative CP design. For solving the problem we have developed the modification of the split and bound method, originally developed in ref[1] for solving TSOP2.

For comparison we give the formulations of the feasibility test and the two-stage optimization problem for different levels of fullness and accuracy of the plant data available at the operation stage. In all the cases, the feasibility tests are based on the following flexibility condition: the chemical process is flexible if for all possible values of all or a part of the uncertain parameters, one can find corresponding values of the control variables such that all constraints of Eq. 5 are satisfied. For example, in Case 1 such values of the control variables (enabling satisfaction of all constraints in Eq. 5) must exist for each $\theta \in T$ ($\theta^1 \in T^1$ for Case 2). The expressions that are the mathematical equivalents of this condition contain many min-max operators (see Eqs. 4, 14, and 38). Since the goal of the design is to construct the flexible CP, the flexibility condition is used as a constraint in TSOP as well (for example, Eq. 2 in Case 1). The presence of many min-max operators is the reason[3] to use nondifferentiable optimization methods in the case of direct calculation of the flexibility function and of the direct solution of TSOP. These methods are very computationally intensive. Therefore, an important problem is to develop approaches for calculation of the flexibility function and solving TSOP that do not require the use of nondifferentiable optimization methods. In ref[1] we have developed the split and bound method for solving TSOP2. In this article, we have developed the modification of the split and bound method for solving TSOP4.

## Notation

$d$ = vector of the design variables
$E_{\boldsymbol{\theta}\in T}\{f_i^*(\boldsymbol{d},\boldsymbol{\theta})\}$ = mathematical expectation of $f_i^*(\boldsymbol{d},\boldsymbol{\theta})$ over the region $T$
$f(\boldsymbol{d}, z, \boldsymbol{\theta})$ = reduced objective function
$f(\boldsymbol{d}, \boldsymbol{x}, z, \boldsymbol{\theta})$ = original objective function
$f_i$ = optimal value of the objective function of TSOP$i$
$f_i^*(\boldsymbol{d},\boldsymbol{\theta})$ = optimal value of the objective function of the internal optimization problem for the $i$-th case
$\tilde{f}_4$ = optimal value of the objective function of the simplified discrete variant of the TSOP4
$f_4^{U,(k)}$ = upper bound of the objective function of TSOP4 at the $k$-th iteration of the SB method
$f_4^{L,(k)}$ = lower bound of the objective function of TSOP4 at the $k$-th iteration of the SB method
$g_j(\boldsymbol{d}, z, \boldsymbol{\theta})$ = left-hand side of the reduced process constraints
$\bar{g}_j(\boldsymbol{d}, \boldsymbol{x}, z, \boldsymbol{\theta})$ = left-hand side of the original process constraints
$T_i^{1,(k)}$ = the $i$-th subregion of the region $T^1$ at the $k$-th iteration
$v_l$ = weight coefficient
$w_i$ = weight coefficient
$\boldsymbol{x}$ = vector of the state variables
$\bar{x}$ = vector of the measured state variables
$\bar{\bar{\boldsymbol{x}}}$ = vector of the unmeasured state variables
$\boldsymbol{z}$ = vector of the control variables
$\boldsymbol{z}^{il}$ = vector of the control variables corresponding to the approximation point $\theta^{il} = (\theta^{1i}, \theta^{2l})$
$\boldsymbol{z}^{IL}$ = set of the control variables corresponding to all the approximation points $\theta^{il} = (\theta^{1i}, \theta^{2l})$
$z^r$ = control variable vector corresponding to the point $\bar{\boldsymbol{\theta}}^{1,r}$,
$\boldsymbol{z}^R$ = set of the control variables vectors corresponding to all the points $\bar{\boldsymbol{\theta}}^{1,r}$.
$z^s$ = vector of the control variables corresponding to the $s$-th subregion
$\boldsymbol{z}^S$ = the set of the control variables vectors corresponding to all the subregions $T_s^{1,(k)}$ $s = 1, \ldots, N_k$.

### Greek letters

$\chi_i(d)$ = feasibility function for the $i$-th case
$\chi_4^U$ = upper bound of the feasibility function $\chi_4(d)$ for the uncertainty region $T^1$
$\chi_4^{U,i}$ = upper bound of the feasibility function $\chi_4^i(d)$ for the subregion $T_i^1$
$\mu(\boldsymbol{\theta})$ = probability density function (pdf) of the vector of the uncertain parameters $\boldsymbol{\theta}$
$\boldsymbol{\theta}$ = vector of the uncertain parameters
$\bar{\boldsymbol{\theta}}^{1,i}$ = approximation point in the subregion $T^1$
$\bar{\boldsymbol{\theta}}^{2l}$ = approximation point in the subregion $T^2$
$\bar{\theta}$ = vector of the measured uncertain parameters
$\bar{\bar{\theta}}$ = vector of the unmeasured uncertain parameters
$\varphi_j(d, x, z, \theta)$ = left-hand side of the state equations

## Literature Cited


1. Ostrovsky GM, Achenie LEK, Datskov I, Volin YuM Optimization of chemical processes under uncertainty: the case of insufficient process data at the operation stage. *AIChE J.* 2003;49:1216-1240.
2. Rooney WC, Biegler LT. Optimal process design with model parameter uncertainty and process variability. *AIChE J.* 2003;49:438-449.
3. Halemane KP, Grossmann IE. Optimal process design under uncertainty. *AIChE J.* 1983;29:425-433.
4. Swaney RE, Grossmann IE. An index for operational flexibility in chemical process design. *AIChE J.* 1985;31:621.
5. Grossmann IE, Floudas CA. Active constraints strategy for flexibility analysis in chemical processes. *Comput Chem Eng.* 1987;11:675.
6. Pistikopoulos EN, Grossmann IE. Optimal retrofit design for improving process flexibility in nonlinear systems—1: Fixed degree of flexibility. *Comput Chem Eng.* 1989;12:1003-1016.
7. Pistikopoulos EN, Ierapetritou MG. Novel approach for optimal process design under uncertainty. *Comput Chem Eng.* 1995;19:1089-1110.
8. Straub DA, Grossmann IE. Design optimization of stochastic flexibility. *Comput Chem Eng.* 1993;17:339-354.
9. Ostrovsky GM, Volin YuM, Barit EI, Senyavin MM. Flexibility analysis and optimization of chemical plants. *Comput Chem Eng.* 1994; 18:755-767.
10. Ostrovsky GM, Volin YuM, Senyavin MM. An approach to solving a two-stage optimization problem under uncertainty. *Comput Chem Eng.* 1997;21:317-325.
11. Acevedo J, Pistikopoulos EN. Stochastic optimization based algorithms for process synthesis under uncertainty. *Comput Chem Eng.* 1998;22:647-671.
12. Bernardo FP, Pistikopoulos EN, Saraiva PM. Integration and computational issues in stochastic design and planning optimization problems. *Ind Eng Chem Res.* 1999;38:3056.
13. Bansal V, Perkins JD, Pistikopoulos EN. Flexibility analysis and design using parametric programming framework. *AIChE J.* 2002;48: 2851-2868.
14. Ahmad S, Sahinidis NV, Pistikopoulos EN. An improved algorithm for optimization under uncertainty. *Comput Chem Eng.* 2000;23:1589-1604.
15. Raspanti CG, Bandoni JA, Biegler LT. New strategies for flexibility analysis and design under uncertainty. *Comput Chem Eng.* 2000;24: 2193-2209.
16. Ierapetritou MG. New approach for quantifying process feasibility: Convex and 1.D quasi-convex regions. *AIChE J.* 2001;47:1407-1947.
17. Floudas CA, Gumu ZH, Ierapetritou MR. Global optimization in design under uncertainty: feasibility test and feasibility index problems. *Ind Eng Chem.* 2001;40:4267.
18. Carnahan BH, Luther A, Wilkes JO. *Applied Numerical Methods*. New York: Wiley; 1969.
19. Walsh S, Perkins J. "Operability and control in process synthesis and design," *Process Synthesis, 301*, JL Anderson (Ed.), New York: Academic Press; 1996:301-341.
20. Hettich R, Kortanek KO. Semi-infinite programming: theory, methods and applications. *SIAM Review.* 1993;35:380-429.
21. Ostrovsky GM, Achenie LEK, Karalapakkam AM, Volin YuM. Flexibility analysis of chemical processes: selected global optimization sub-problems. *Optimization and Eng.* 2002;3: 31-52.
22. Craig L, Zhou JL, Tits AL. *User's Guide for CFSQP Version 2.5*. College Park: Institute for Systems Research, University of Maryland; 1997.
23. Fogler HS. *Elements of Chemical Reaction Engineering*. Upper Saddle River, NJ: Prentice Hall PTR; 1999.
24. McKinsey JCC. *Introduction to the Theory of Games*. New York: McGraw-Hill; 1952.


## Appendix

### Auxiliary relations

We need the following equivalent relations and inequalities:

$$\max_{x\in X} \varphi(x) \le 0 \Leftrightarrow \varphi(x) \le 0, \quad \forall\, x \in X \tag{A1}$$

$$\exists\, x \quad f(x) \le 0 \Leftrightarrow \min_x f(x) \le 0 \tag{A2}$$

$$\max_x \max_y f(x, y) = \max_y \max_x f(x, y) \tag{A3}$$

$$\min_x \max_y f(x, y) \ge \max_y \min_x f(x, y) \tag{A4}$$

where $x$ is a vector of continuous or discrete variables. The relations in Eqs. A1, A2, and A3 are obvious, while Eq. A4 is proved in ref.[24].

*Theorem A.1.* Consider the problem

$$f = \min_{x \in X} f(x) \tag{A5}$$

$$\min_{y \in Y} \varphi_i(x, y) \le 0, \quad i = 1, \ldots, m$$

Here there are $n_x$ search variables ($n_x$ = dim $x$). Consider the auxiliary problem

$$\bar{f} = \min_{x, y^i} f(x)$$
$$\varphi_i(x, y^i) \le 0, \quad i = 1, \ldots, m \tag{A6}$$

where $y^i,(i = 1,\ldots, m)$ is a vector of new search variables, corresponding to the $i$-th constraint. Thus, in Eq. A6, there are $n_x + mn_y$ variables. It is shown in ref. [10] that if $x^*$, $y^{i*}$ is the solution of Eq. A6, then $x^*$ is the solution of Eq. A5.

### *Theorem 2A*

Let us consider the problem

$$f_1 = \min_x \sum_{i=1}^{N} f_i(x^i)$$
$$g_i(x^i) \le 0, i = 1, \ldots, N.$$

where $x = \{x^i\}\ i = 1,\ldots, N\}$ and $x^i$ is a subvector of $x$. Suppose the subvectors $x^i$ ($i = 1,\ldots, N\}$ do not have common variables. Therefore, we can change the order of the minimization and summation operators since each minimization is carried out with respect to its own search variables. Therefore, we have

$$\min_x \left( \sum_{i=1}^{N} f_i(x^i)/g_i(x^i) \le 0, i = 1, \ldots, N \right\}$$
$$= \sum_{i=1}^{N} \min_{x_i}\{f_i(x^i)/g_i(x^i) \le 0 \tag{A7}$$

$$x^i \cap x^j = 0, \quad \forall\, i, j = 1, \ldots, N \tag{A8}$$

*Manuscript received May 24, 2005, and revision received May 23, 2006.*