SIMULATION OF CHEMICAL PROCESSES WITH UNCERTAIN PARAMETERS

Mirko DOHNAL and †Marie ULMANOVÁ

Department of Mechanical Equipment Chemical and Food Industry, Technical University, 611 00 Brno

Received May 13th, 1980

An integral part of the simulation problem are numerical values of parameters or constants of industrial units. Either these quantities need not be known accurately or it is not possible to control them at the set up value. The result of simulation cannot be then accurate. It is necessary to determine the region in which the flow rates of components through branches could change when the parameters and constants change. For these purposes is applicable the Monte Carlo method. But it is very time-consuming. Here, algorithms are proposed which with the use of the algebraic norm approximately determine the looked-for set of flow rate uncertainty. These algorithms can be applied for solution of a number of practical problems related with optimum sizing of units, sensitivity studies *etc.* An example is given of a cooling cascade cycle.

The steady state simulation of deterministic chemical processes lies in calculation of flow rates of all components and enthalpies in all streams of the process if the following quantities are given[1]

- topology (mutual interconnection of individual units is known),
- numerical values of parameters and constants of mathematical models,
- mass, concentrations and enthalpies of all system inputs.

In solution of actual problems, except of topology, none of the required informations is accurate. Mathematical models of units (unit operations) are not accurate. The same can be said about the values of parameters, constants and properties of input streams.

In majority of cases this situation is simply ignored. In design of dimensions of units are chosen unnecessarily large heat transfer areas, volumes *etc.* The impossibility to describe quantitatively the effects of uncertainty of the studied phenomena results in too expensive units.

There exist several methods for a general estimate of the effect of errors of input data on solution such as is the sensitivity[2] analysis, Monte Carlo[3] method, interval arithmetics[4], *etc.*[5–7]. These methods can supply useful information but only on the account of rather expensive computer time or a complex theoretical analysis.

Here relatively simple algorithms are proposed which overcome the mentioned disadvantages and are sufficiently accurate to supply the designer with reliable design data.

Analysis of the Problem

Mathematical model of the chemical process is a system of non-linear equations. They can be written in the form

$$f(X, K, P) = 0\,, \tag{1}$$

where X represents the vector of variables (flow rates of components in streams), K vector of constants, P vector of parameters. The vector X_r which satisfies Eq. (1) is in general a function of constants K and of parameters P

$$X_r = \Phi(K, P)\,.$$

But the values of constants and parameters in solution of practical problem are not accurately known, *i.e.* they can change in a certain region.

Let us denote as $\mathscr{K}$ the set of all numerical values of constants K and $\mathscr{P}$ the set of all values of parameters P. It implies, that the solution of system (1) is not a single value X_r but a whole set denoted as X_R, which

$$X_R = \{X_r \mid f(X_r, K, P) = 0\,, \quad K \in \mathscr{K}(X_r)\,, \quad P \in \mathscr{P}_r(X_r)\}\,.$$

In general the sets $\mathscr{K}$ and $\mathscr{P}$ depend on values of the vector X_r. But in practical cases it is relatively frequently assumed that these sets are independent of X_r. Let us denote as problem (P) the search for the set X_R.

The mathematical analysis of the problem (P) as has been already stated at the beginning, is relatively complicated and the present theories have nothing to offer in this respect. Very simple is application of the Monte Carlo method. It is based on repeated solution of the system of Eqs (1) while values K and P are randomly generated on sets $\mathscr{K}$ and $\mathscr{P}$. Though this method is not inapplicable from the view of the required computer time it gives useful information on probability distribution of solution as a function of constants and parameters distribution.

In the following part of this study the methods are considered which are not time-consuming and are so simple that they can be relatively easily put into the form of a program.

Linearisation of the Problem

To be able to solve the given problem (P) in an acceptably short time it is necessary to simplify it. The simplification rests in substitution of the nonlinear model, described by the system of Eqs (1) by the linear model

$$f^{L}(X, K, P) = 0\,, \tag{2}$$

which is obtained by linear approximation of Eqs (*1*). In the study of the new problem the estimate of the effect of uncertainties due to linearisation is not considered. Solution of X_r^L of the system (*2*) is again a function of constants and parameters

$$X_r^L = \Phi^L(K, P) .$$

Due to uncertainties of constants and parameters, there exists again a whole set of solutions, which is denoted as X_R^L, and is given by

$$X_R^L = \{X_r^L \mid f(X_r^L, K, P) = 0 , \quad K \in \mathscr{K} , \quad P \in \mathscr{P}\} .$$

The looked for set X_R^L is denoted as the problem (PL).

There are several linearisation methods. As concerns the estimate of the effect of constants and parameter uncertainties on solution it is suitable to use the mathematical model in the form of Nagiev balance equations[8]. The Nagiev balance equations have the form

$$S_i^t + \sum_{\substack{j=1 \\ j \neq i}}^{N} \alpha_{ji}^t \lambda_j^t = \lambda_i^t \qquad (3)$$

$$i = 1, 2, \ldots, N ; \quad t = 1, 2, \ldots, F ,$$

where N is the number of nodes of the system, F number of components, S_i^t is the system input of component t into the i-th node, λ_j^t is the total flow rate of component t through the node j, α_{ji}^t is the splitting ratio for the component t carried from the node j into the node i.

If we assume that the splitting ratios are constant for a certain definite state of the process we obtain F systems of N linear equations. The flow rates are obtained by solution of the system of Eqs (*3*) denoted as λ_r. The effect of uncertainty of both constants and parameters can be expressed by modification of the system of Eqs (*3*) in the following way

$$S_i^t + \delta S_i^t + \sum_{\substack{j=1 \\ j \neq i}}^{N} (\alpha_{ji}^t + \delta\alpha_{ji}^t) \lambda_j^t = \lambda_i^t . \qquad (4)$$

The uncertainties of coefficients can be *e.g.* given by the intervals *i.e.* by

$$\delta\alpha_{ji}^t \in \langle -\Delta\alpha_{ji}^t , \quad \Delta\alpha_{ji}^t \rangle$$

$$\delta S_i^t \in \langle -\Delta S_i^t , \quad \Delta S_i^t \rangle ,$$

where $\Delta\alpha_{ji}^t$, ΔS_i^t are the upper estimates of errors of coefficients α_{ti}^j and S_i^t.

Let us denote as λ_R the set of all solutions of the system (4). The looked for set is the already mentioned problem (PL).

ESTIMATE OF THE EFFECT OF ERRORS

For illustration let us write the set of Eqs (3) for component t in the form

$$A^t\lambda^t = S^t$$

and its solution as λ_r^t. The elements of the matrix of the system A^t and of the vector S^t are not known accurately. Therefore a solution is looked for which fits the system of equations

$$(A^t + \delta A^t)(\lambda_r^t + \delta\lambda^t) = S^t + \delta S^t .$$

The matrix δA^t and the vector δS^t are the errors of coefficients and $\delta\lambda^t$ are the errors of solution.

According to the character of errors δA^t and δS^t the problems are divided into two groups[9]:

- in case of errors of random character with known distribution the problem is called problem under risks,
- in case of unknown distribution of errors, where only their upper and lower estimates are known, *i.e.* where ΔA^t and ΔS^t are known, for which there holds

$$\delta A^t \in \langle -\Delta A^t,\ \Delta A^t \rangle$$

$$\delta S^t \in \langle -\Delta S^t,\ \Delta S^t \rangle ,$$

the problem is called under uncertainties.

When the problem is formulated as the problem under risks it is possible to solve the case when the errors of coefficients δA^t and δS^t have the normal distribution. It is then possible to find the ellipsoid of solution error concentrations and a corresponding probability with which the solutions $\lambda_r^t + \delta\lambda^t$ is situated outside this ellipsoid.

Let us arrange the errors of input data into the form

$$\delta a_{11}^t, \ldots, 1a_{N1}^t, \ldots, \delta a_{1N}^t, \ldots, \delta a_{NN}^t, \ldots, \delta S_1^t, \ldots, \delta S_N^t .$$

Their covariant matrix, which is of the $N^2 + N$ order is denoted as Δ. It can be divided into square submatrices Δ_{ij} of the order N, $i, j\ m\ 1, 2, \ldots, N + 1$. The mean values of elements of the matrix δA^t and of vector δS^t are equal to zero, *i.e.* there holds

$$E(\delta a_{ij}^t) = 0 \quad i, j = 1, \ldots, N .$$

$$E(\delta S_i^t) = 0 .$$

Then the matrix Δ_{ij} is formed by mean values of products of elements of the i-th and j-th column of the matrix δA for $i, j = 1, \ldots, N$. or of the column vector δS^t for $i = N + 1$ or $j = N + 1$, *i.e.* for

$$\Delta_{ij} = \begin{pmatrix} E(\delta a_{i1}^t \,.\, \delta a_{j1}^t) \ldots E(\delta a_{i1}^t \,.\, \delta a_{jN}^t) \\ \vdots \\ E(\delta a_{iN}^t \,.\, \delta a_{j1}^t) \ldots E(\delta a_{iN}^t \,.\, \delta a_{jN}^t) \end{pmatrix} .$$

Then the covariant matrix K of errors of vectors of solution $\delta\lambda^t$ is of the form

$$K = A^{-1} \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} \lambda_{ri}^t \lambda_{rj}^t \Delta_{ij} (A^{-1})^T .$$

The probability that the vector of errors $\delta\lambda^t$ is outside the ellipsoid of concentration is

$$\lambda_r^T K^{-1} \lambda_r = R^2$$

and is given by relation[10-12]

$$\left(\frac{N}{2} - 1\, \Gamma\left(\frac{N}{2}\right)\right)^{-1} \int_R^\infty x^{N-1} \exp\left(-\tfrac{1}{2}x^2\right) \mathrm{d}x .$$

For the estimate of errors in problems under uncertainty it is suitable to use two following methods.

The first method gives the estimate of errors of solution in the vector norm which is in agreement with some multiplicative matrix norm. For definition of needed terms – see[10].

Solution of the system of equations

$$(A^t + \delta A^t)(\lambda_r^t + \delta\lambda^t) = S^t + \delta S^t ,$$

where there are given the upper bounds of errors ΔA^t and ΔS^t *i.e.* for errors δA^t and δS^t there holds

$$\delta A^t \in \langle -\Delta A^t,\ \Delta A^t \rangle$$

$$\delta S^t \in \langle -\Delta S^t,\ \Delta A^t \rangle .$$

When the condition $\|A^t\| \cdot \|A^{-1}\| < 1$ is satisfied for the multiplicative matrix norm corresponding to a norm vector so the norm of solution errors is

$$\|\delta\lambda^t\| < \frac{\|A^{-1}\|}{1 - \|\Delta A^t\| \, \|A^{-1}\|} \left(\|\Delta S^t\| + \|\Delta A^t\| \cdot \|A^{-1}\| \cdot \|S^t\|\right).$$

The set of possible solutions $\lambda_r^t + \delta\lambda^t$ of the given system is inside the convex body whose center is in the point λ_r^t. It is possible to use *e.g.* the following vector norm

$$\|\lambda^t\|_1 = \sum_{i=1}^{N} |\lambda_i^t|$$

$$\|\lambda^t\|_2 = \left[\sum_{i=1}^{N} (\lambda_i^t)^2\right]^{1/2}$$

$$\|\lambda^t\|_\infty = \max_i |\lambda_i^t| .$$

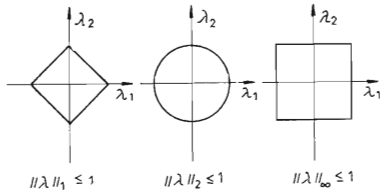

FIG. 1
Geometric shapes of three basical types of norms

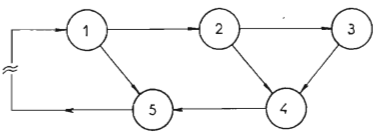

FIG. 2
Computation scheme of the equipment for liquifying of natural gas. Nodes 1—5 are the liquifying stages. From the first and the second stages is withdrawn liquid for cooling in stages 4 and 5

For the two-dimensional case these norms correspond to convex bodies in Fig. 1.

The estimate made according to this method gives relatively pesimistic results as has been obtained by comparison with the Monte Carlo method.

The second method is making use of the elliptic norm. The region of solution $\lambda_r^t + \delta\lambda^t$ is circumscribed by the N-dimensional ellipsoid relatively accurately[11,12] as has again been verified by the Monte Carlo method.

If in the system of linear equations

$$(A^t + \delta A^t)(\lambda_r^t + \delta\lambda^t) = S^t + \delta S^t$$

the upper bounds of errors of coefficients are expressed in the spherical norm *i.e.*

$$\|\Delta A^t\| < \varepsilon_1 \quad \|\Delta S^t\| < \varepsilon_2$$

then the matrix

$$K = 2[\varepsilon_1^2 \|\lambda_r\| + \varepsilon_2^2]\,((A^t)^T A^t)^{-1}$$

determines the looked for ellipsoid. Its center is located in the point λ_r^t, lengths of semiaxes are equal to the second roots of their eigenvalues and semiaxes are located in the directions of eigenvectors.

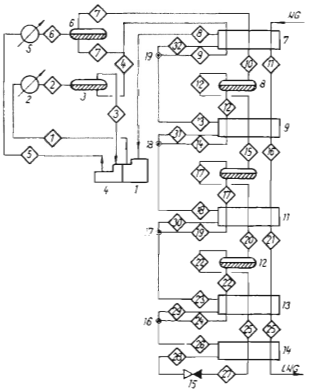

FIG. 3

Processing scheme of the unit for liquifying of natural gas. Nodes *3*, *6*, *8*, *10*, *12* separators, nodes *7*, *9*, *11*, *13*, *14* multistream heat exchangers, nodes *2*, *5* water coolers, nodes *16*, *17*, *18*, *19* valves, node *15* choke valve, nodes *1* and *4* two-stage compressor, NG-natural gas, LNG-liquefied natural gas

Example: For illustration, the application of the described methods has been studied by the modified cascade cooling system[13]. Its simplified flowsheet used for the material balance is in Fig. 2. The original scheme of the process is given in Fig. 3.

The values of splitting ratios and inputs and their errors were given. For determination of the magnitude of errors of splitting ratios and inputs the next assumptions were made:

— the calculated equilibrium concentrations (vapour–liquid) differ from the experimental data either relatively less than by 5% or absolutely by less than for 0·0005 of mole fraction;

— sum of all values of splitting ratio in the corresponding node must according to definition be equal to 1, *i.e.*

$$\sum_{i=1}^{N} \alpha_{ji}^{I} = 1 \quad j = 1, 2, \ldots, N$$

which must hold also when the error is included *i.e.*

$$\sum_{i=1}^{N} (\alpha_{ji}^{I} + \delta\alpha_{ji}^{I}) = 1 .$$

The given problem has been presented as the problem under uncertainty. For information it has been studied also as the problem under risk with normal distribution of coefficients of errors.

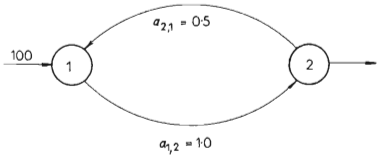

FIG. 4
Testing scheme formed by two nodes

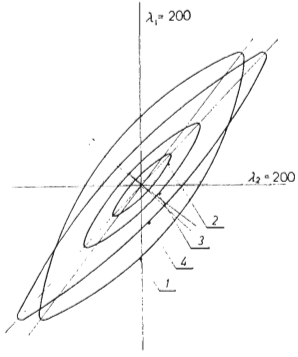

FIG. 5
Uncertainty ellipses — two-dimensional case. Ellipsoid 1, $K_1 = K_2 = 0{\cdot}05$; 2 $K_1 = K_2 = 0{\cdot}00625$; 3 $K_1 = K_2 = 0{\cdot}0125$; 4 $K_1 = K_2 = 0{\cdot}025$

The variations have been chosen as

$$(\tfrac{1}{3}\delta\alpha_{ji}^{t})^2 , \quad \text{or} \quad (\tfrac{1}{3}\delta S_i)^2$$

i.e. the intervals in which the coefficients can be located in the examples at uncertainty have the 99·7% probability (see properties of standard distribution).

The given example is not too illustrative since the result is the ellipsoid in the 5-dimensional space. The situation here is thus similar as in the study of the part of a process for production of low solidifying oils and parafines[14].

For clarification is thus given a simple two-dimensional example which can be demonstrated graphically.

The scheme of this simple system is given in Fig. 4.

This two-dimensional case has been again solved by both methods. For the case at uncertainty the 20% relative errors were considered *i.e.*

$$\Delta\alpha_{ji} = K_1\alpha_{ji} \qquad (5)$$
$$\Delta S_i = K_2 S_i$$

$i, j = 1, 2, \ldots, N$.

The elliptic norm and ellipsoids of concentration with the corresponding probabilities are given in Fig. 5.

REFERENCES

1. Lufcha J., Klemeš J., Vašek V., Dohnal M., Jicha J. in the book: *Modellierung und Optimierung Verfahrenstechnischer Systeme*. Akademie Verlag, Berlin 1977.
2. Chen M. S., Erikson L. E., Fan L.: Ind. Eng. Chem. Proc. Des. Develop. *9*, 514 (1970).
3. Ermakov S. M., Michaikov G. A.: *Course in Statistic Modelling* (in Russian). Nauka, Moscow 1976.
4. Moore R. E.: *Interval Analysis*. Prentice-Hall, Englewoods Cliffs, N. J. 1966.
5. Kulisch U., Bohlender C.: Computing *16*, 239 (1976).
6. Miller W.: SIAM J. Comput. *5*, 204 (1976).
7. Ohnaka K., Yasui H., Kubo T.: *An Automatic Error Estimation Method in Numerical Calculation*. Technology Reports of Osaka University, Vol. 25, No 1254. Osaka 1975.
8. Nagiev M. F.: *Studies on Chemical Systems with Feedback* (in Russian). Nauka, Moscow 1971.
9. Dohnal M.: Ropa Uhlie *20*, 72 (1978).
10. Ralston A.: *A First Course in Numerical Analysis*. McGraw-Hill, New York 1965.
11. Fadeev K. D., Fadeeva N. V.: SIAM J. Numer. Anal. *7*, 526 (1970).
12. Fadeev D. K., Fadeeva V. N.: Zh. Vychislit. Matem. Matem. Fyz. *14*, 439 (1974).
13. Fojtek J., Dohnal M.: Presented at International Cryogenic Congress, Moscow, September 1975.
14. Dohnal M., Ulmanová M., Klemeš J., Lufcha J.: Presented at Symposium on Computers in the Design and Erection of Chemical Plants, Karlovy Vary, September 1975.

Translated by M. Rylek.