Synergy Analysis in Reaction Systems

ZELIMIR KURTANJEK
Faculty of Food Technology and Biotechnology
University of Zagreb
Pierottijeva 6, 10 000 Zagreb
CROATIA
zkurt@pbf.hr http://www.pbf.hr

Abstract: This work introduces the concept of analysis of synergy interactions in chemical (biochemical) reaction systems. The systems are defined by mass balances for reacting species, reaction network is given by a stoichiometric matrix, reaction (kinetic) parameters, and parameters of interaction between the reaction system and it's surrounding. Systems analysis is based on statistical evaluation of numerical results of computer simulation of effects of a set of uncorrelated stochastic parameters, defined by corresponding ranges and probability distributions. From the simulation data are extracted parameter global sensitivities and synergy effects of each parameter in interaction with the whole system. The global parameter sensitivities are defined as ratios of conditional variances of expected values for each parameter and the total dispersion of simulation data. The synergy effects are defined as the relative difference between expected values of the conditional variances with exclusion of each parameter and the variance of the conditional expected value of the parameter. Numerical evaluations are performed by use of sampling the parametric space by Lissajous type of curves and the Fourier Amplitude Sensitivity Test (FAST) algorithm. The method is applied for analysis of propagation of hepatitis B virus (HBV) in a stochastic single cell model, and for metabolic control analysis of a branched network. The results for virus propagation show dominant synergy effect of the system parameters, compared to negligible influence of individual parameters. The metabolic control analysis show that under homeostatic conditions influence of the individual enzymes is dominant and the key enzyme responsible for flux regulation is determined. However, under perturbed homeostatic conditions influence of synergy effects dominates over impact of individual parameters. Applicability of the proposed concept is discussed in view of model improvement and potential control of reaction systems.

Key-Words: reaction system, synergy, global parameter sensitivity, HBV, MCA, FAST

1 Introduction

This work is focused on systems analysis and information extraction on individual parameter importance and their communal interaction in complex reaction systems. The motivation for this analysis is its application for model development and improvement by better selection of experiment design, reduction of model dimension, increased accuracy of parameter estimation, and determination of the key parameters for system design and process control. Schematic presentation of the considered system is depicted in Fig. 1. Here is the input-output concept redefined in terms of parameter interaction. A considered reaction system (chemical or biochemical) is defined by dynamics of state variables **x**, a set of system parameters **p**, and interaction with surroundings expressed by a set of **c** parameters. System dynamics is expressed by a set of ordinary differential equations defined by a vector of nonlinear reaction kinetics **v** and system structure given by a constant stoichiometric matrix **N**

$$\frac{d}{dt}\mathbf{x} = \mathbf{N}\cdot\mathbf{v}(\mathbf{x},\mathbf{p},\mathbf{c}) \qquad (1)$$

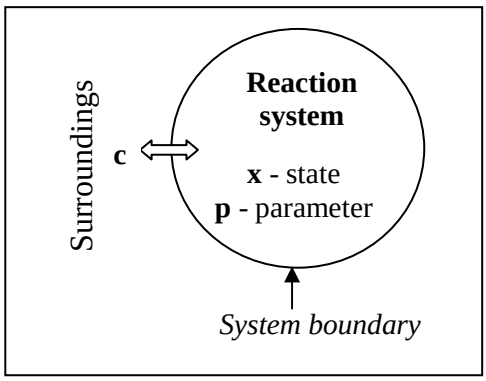

Fig. 1. Systems view of a reaction system defined by concentrations of reacting species **x**, reaction parameters **p**, and interaction parameters ***c*** with surroundings.

The stoichiometric matrix are deterministic integers, usually in the range [-2,2]. Structural properties of the reaction network are considered to be fixed, or deterministic, and can evaluated by the singular value decomposition of the corresponding stoichiometric matrix

$$\mathbf{N} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \tag{2}$$

by which are defined the four fundamental subspaces:
1) the null space contains all the steady state fluxes (corresponding to homeostasis in a living reaction system);
2) the row space contains all the dynamic flux distributions;
3) the left null space contains all the conservation relationships or time invariants;
4) and the column space contains all possible time derivatives of concentration vectors.

The stochastic part of the system is associated with the reaction rates or kinetics **v**. The kinetic models are derived from the law of mass action, or by Langmuir-Hinshelwood kinetics in heterogeneous catalytic systems, or Michaelis-Menten rate expressions when reactions are mediated by enzymes like in living systems. In general, choice of kinetic functions can also be considered to be associated with a certain probability. In this work the main source of stochastic effects is due to uncertain values of kinetic parameters **p** and the interaction with parameters **c**.
Reaction system performance is viewed through mass and energy exchange rates or fluxes **J** between network nodes (reacting species) and its surroundings

$$\mathbf{J} = \mathbf{J}(\mathbf{x}) \tag{3}$$

The usual, classical, approach to analysis of parameter effects on the system performance is evaluated by the local and infinitesimal one parameter at time sensitivities

$$S_{k,l}^{J}(t) = \frac{p_l}{J_k} \cdot \frac{\partial J_k(t)}{\partial p_l} \tag{4}$$

Evaluation of (4) requires simultaneous integration of the sensitivity equations:

$$\frac{d}{dt}\frac{\partial J_k}{\partial p_l} = \frac{\partial J_k}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial p_l} \tag{5}$$

$$\frac{d}{dt}\frac{\partial \mathbf{x}}{\partial p_l} = \mathbf{N} \cdot \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial p_l} + \frac{\partial \mathbf{v}}{\partial p_l} \right) \tag{6}$$

However, evaluation of the effect of perturbation of an individual parameter p_l depends on the rest of the parameters which need to be assumed constant and accurately known.

2 Problem Formulation

In view of significant uncertainties in models of complex reaction systems, here is applied a method of stochastic simulation of the complete set of parameters. Each parameter p is considered as an independent random variable with corresponding probability density distribution function ρ_p in a finite range

$$p \in \rho_p[p_{\min}, p_{\max}] \tag{7}$$

System analysis is based on evaluation of the first and second moments (expected values and variances) from the multidimensional joint probability distributions. The first moment is

$$E(J_k) = \int_{p_{1,\min}}^{p_{1,\max}} \cdots \int_{p_{N,\min}}^{p_{N,\max}} J_k(p_1, \cdots p_N) \cdot \cdot \rho_1(p_1) \cdots \rho_N(p_N) \cdot dp_1 \cdots dp_N \tag{8}$$

The second moment is the variance of the flux

$$\sigma^2(J_k) = E(J_k^2) - E(J_k)^2 \tag{9}$$

The expected value of the ensemble of the square term is given by

$$E(J_k^2) = \int_{p_{1,\min}}^{p_{1,\max}} \cdots \int_{p_{N,\min}}^{p_{N,\max}} J_k^2(p_1, \cdots p_N) \cdot \cdot \rho_1(p_1) \cdots \rho_N(p_N) \cdot dp_1 \cdots dp_N \tag{10}$$

Effect $S_{k,l}^{J}$ of each individual parameter p_l on the flux J_k is evaluated by the ratio of the conditional variance of a given parameter and the total dispersion of the ensemble of parameters

$$S_{k,l}^{J} = \frac{\sigma^2\left(E\left(J_k \middle| p_l\right)\right)}{\sigma^2(J_k)} \tag{11}$$

Synergy effect of each parameter p_l is evaluated by the difference between the expected value of conditional ensemble variance when all parameters are fixed except p_l and the conditional ensemble variance of the expected value when p_l is fixed relative to the total variance:

$$Syn_{k,l}^{J} = \frac{1}{\sigma^2(J_k)} \cdot \left[E\left(\sigma^2\left(J_k|p_{\sim l}\right)\right) - \sigma^2\left(E\left(J_k|p_l\right)\right)\right] \quad (12)$$

The multidimensional integrals can be numerically evaluated by Monte Carlo system simulation based on parameter space sampling by a random generator. In this work is applied numerically more efficient method proposed by Cukier *et al.* [1-2]. It is based on application of Lissajous type of curves for uncorrelated exploration of the parameter space. Each parameter is associated with a frequency and simulation results are expanded into Fourier components. The numerical procedure is accordingly called the Fourier Amplitude Sensitivity Test (FAST) [1-2].

Assumed are uniform probability distributions of each parameter in a predefined vale range. For example, the ranges may be selected from significance intervals obtained after parameter estimation procedure. Alternatively, the ranges may be assumed to cover a whole feasible range of parameter variations. The parameter ranges are scaled to the standard range [-1,1]. The uniform probability distribution is generated by piece wise linear functions *arcsin*(*sin*(*s*)) of a scan variable *s* and two randomly selected parameters, frequency ω_l, and phase φ_l, corresponding to each parameter p_l

$$p_l = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin\left(\sin\left(\pi \cdot \omega_l \cdot \mathrm{s} + \varphi_l\right)\right) \quad (13)$$

Integer frequencies are randomly selected in a preselected range, and the scan variable is incrementally covering the complete range [-1,1]. Random selection of the parameters ensures that the functions (13) are mutually independent, i.e. uncorrelated. For each value of the scan variable s_i the flux $J_k(s_i)$ is evaluated and the resulting data are interpolated to provide a continuous function $J_k(s)$ which is expanded into Fourier series

$$A_\omega = \frac{1}{2 \cdot \pi} \cdot \int_{-\pi}^{\pi} J_k(s) \cdot \cos(\omega \cdot s) \cdot ds \quad (14)$$

$$B_\omega = \frac{1}{2 \cdot \pi} \cdot \int_{-\pi}^{\pi} J_i(s) \cdot \sin(\omega \cdot s) \cdot ds \quad (15)$$

The total dispersion is calculated from the full spectrum (effectively truncated)

$$D_T = 2 \cdot \sum_{\omega=1}^{\infty} \left(A_\omega^2 + B_\omega^2\right) \quad (16)$$

Contribution of each parameter p_l in the total dispersion is calculated from the harmonics of the corresponding fundamental frequency ω_l

$$D_l = 2 \cdot \sum_{\omega=k \cdot \omega_l}^{\infty} \left(A_\omega^2 + B_\omega^2\right) \quad (17)$$

The partial dispersion that measures the effect of all parameters except the parameter p_l is given by

$$D_{\sim l} = 2 \cdot \sum_{\omega=k \cdot \omega_l}^{\infty} \left(A_{k \cdot \sim \omega l}^2 + B_{k \cdot \sim \omega l}^2\right) \quad (18)$$

Global sensitivity factor is determined for each parameter as the ratio of the dispersion corresponding to each parameter and the total dispersion

$$S_{kl}^{J} = \frac{D_l}{D_T} \quad (19)$$

and the synergistic effect of each parameter p_l as the relative difference

$$Syn_{k,l}^{J} = 1 - \frac{1}{D_T} \cdot \left(D_l + D_{\sim l}\right) \quad (20$$

Numerical simulations, random number generation and Fourier analysis can be efficiently and numerically accurately evaluated by Wolfram Research "*Mathematica*" v.6. software by use of RandomReal, NDSolve, and Fourier algorithms [3-4].

3 Problem Solutions

3.1. Hepatitis B virus propagation

The method is applied for analysis of single cell model virus propagation, brought into chemical engineering by Rawlings and Ekerdt [5]. Since the model involves initially a small number of reacting species it is considered as an example of

a stochastic model of a reaction system, depicted in Fig. 2. The authors analyze the systems dynamics by comparing deterministic with a stochastic Gillespie [6] simulations.

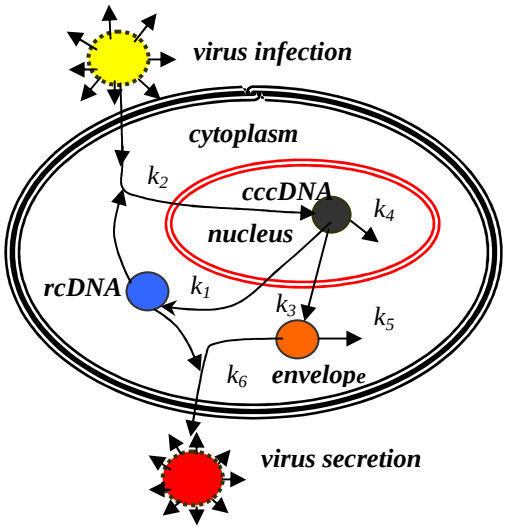

Fig. 2. Schematic presentation of the reaction pathway of the model of hepatitis B virus (HBV) infection. The state variables are: covalently closed circular deoxyribonucleic acid (cccDNA), relaxed circular deoxyribonucleic acid (rcDNA) and envelope proteins.

The key of the model of propagation is a positive feedback loop which results in instability and "explosive" nature of response upon perturbation by transfer of a virus from surrounding into cell interior. The reaction system is given by the first and second order kinetics and the following mechanism:

$$\begin{aligned} &\text{nucleotide} \xrightarrow{cccDNA} \text{rcDNA} \\ &\text{nucleotide} + \text{rcDNA} \rightarrow \text{cccDNA} \\ &\text{amino acids} \xrightarrow{cccDNA} \text{envelope} \\ &\text{cccDNA} \rightarrow \text{degraded} \\ &\text{envelope} \rightarrow \text{degraded} \\ &\text{rcDNA} + \text{envelope} \rightarrow \text{secreted virus} \end{aligned} \tag{21}$$

A typical deterministic behavior is depicted in Fig. 3. Steady state flux J of the virus excretion from the cell into surroundings

$$J = k_6 \cdot c_{rcDNA} \cdot c_{envelope} \tag{22}$$

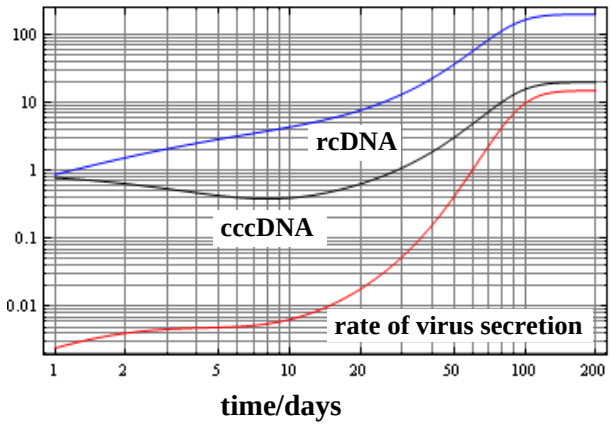

Fig. 3. Dynamics of hepatitis B virus single cell propagation model.

is considered as a measure of system performance and is analyzed for parameter sensitivity and synergy effects in the global stochastic environment. The kinetic parameters are randomly varied in the range of 2 orders of magnitudes around their deterministic values, while the initial concentration of cccDNA(0) between 0 and 2 molecules per cell. Results of statistical evaluation of 2000 simulation cases are depicted as a bi-plot in Fig. 4.

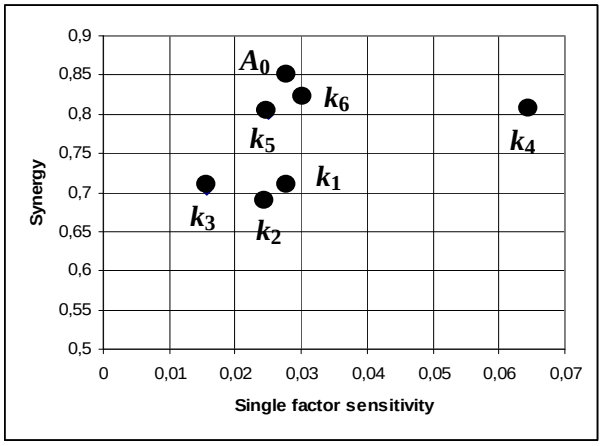

Fig. 4. Bi-plot of single parameter sensitivities and synergies for the model of HBV virus propagation. Specific rate coefficients corresponding to the rates given in eq. (21) are denoted from k_1 to k_6. The initial concentration of cccDNA is labeled as A_0.

The results show a very distinct property of the model parameters. The single most important parameter is k_4 the specific rate of cccDNA degradation. The highest synergy effect is A_0 due to the initial concentration of cccDNA. The total global single parameter sensitivities have profoundly lower

effect compared to their collective synergy. The comparison of the total effects is depicted in Fig. 5.

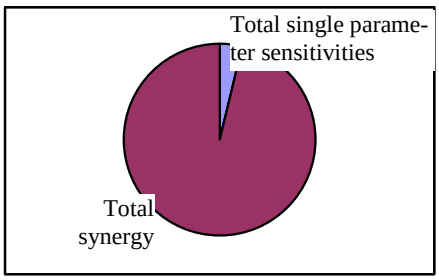

Fig. 5. Ratio of the single parameter sensitivities and the total synergy for the model of HBV virus propagation model.

3.2. Branched metabolic network

The second case of this study is evaluation of synergy effects in a test case of metabolic control analysis (MCA) of a branched pathway.

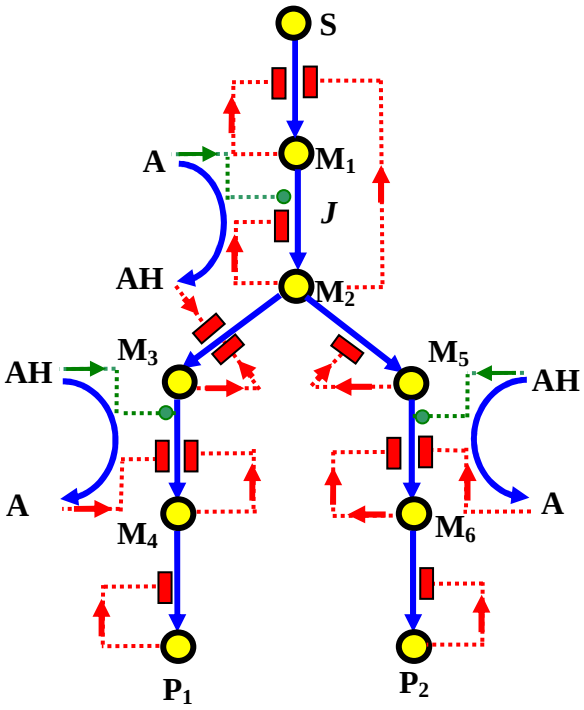

Fig. 6. Scheme of regulation of a branched pathway. Metabolites engaged in the pathway are denoted from M_1 to M_6, and total flux is J.

MCA analysis is a part of the new emerging technologies base on potentials of genetic engineering [7-9]. The model is used as a standard case due to its kinetic complexity and application in analysis of flux redirection at a branch point [10]. The reactions exhibit strong substrate, cofactor and allosteric regulation given by the following kinetics:
Reversible Hill kinetics:

$$v = v_f \cdot \frac{S/S_{0.5} \cdot \left(1 - P/SK_{eq}\right) \cdot \left(S/S_{0.5} + P/P_{0.5}\right)^{h-1}}{\left(S/S_{0.5} + P/P_{0.5}\right)^{h} + \frac{\left(1 + \left(M/M_{0.5}\right)^{h}\right)}{\left(1 + \alpha\left(\frac{M}{M_{0.5}}\right)^{h}\right)}} \tag{23}$$

Ordered bi-bi (reversible) kinetics:

$$v = v_f \cdot \frac{A \cdot B - P \cdot Q / K_{eq}}{A \cdot B \cdot \left(1 + P/K_{iP}\right) + K_{mB} \cdot \left(A + K_{iA}\right) + } \\ + Q \cdot \left(K_{mp} \cdot \left(1 + \frac{K_{mA} \cdot B}{K_{iA} \cdot K_{mB}}\right) + P \cdot \left(1 + \frac{B}{K_{iB}}\right)\right) \tag{24}$$

Uni-uni reversible kinetics:

$$v = v_f \cdot \frac{S - \frac{P}{K_{eq}}}{S + K_{mS} \cdot \left(1 + \frac{P}{K_{mP}}\right)} \tag{25}$$

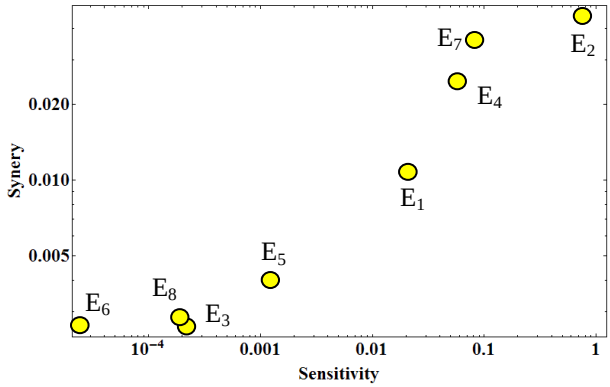

Fig.7. Bi-plot of single parameter sensitivities and synergies under homeostatic conditions.

Statistical MCA results of individual enzymes from E_1 to E_8 and homeostatic parameters are given by Kurtanjek [11]. Here the focus is on comparison of sensitivities and synergy effects in homeostatic (A cofactor, substrate S, and products P_1 and P_2) Fig. 7., and non-homeostatic conditions, Fig. 8. For The total flux J controlling parameter is E_2, the enzyme at the branch point. Under homeostatic conditions it has for an order of magnitude higher sensitivity compared to other enzymes, almost 100 %, and also has the maximum synergy effect. In this case total single parameter sensitivities dominates, 88 %, over the total synergy effects 12%, Fig. 9.

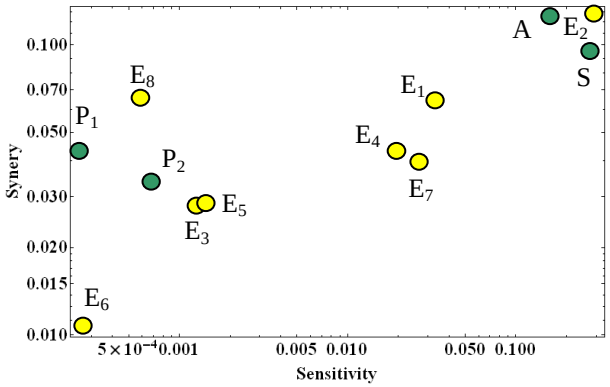

Fig. 8. Bi-plot of sensitivities and synergies obtained under perturbed homeostasis.

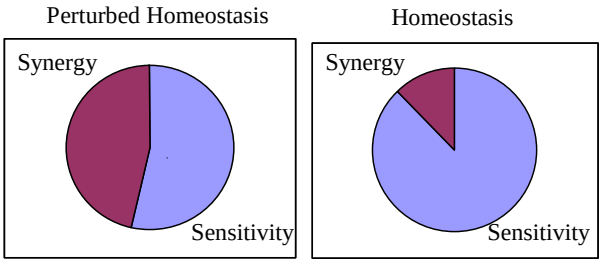

Fig. 9. Parameter sensitivities and synergy effects under steady and perturbed homeostasis.

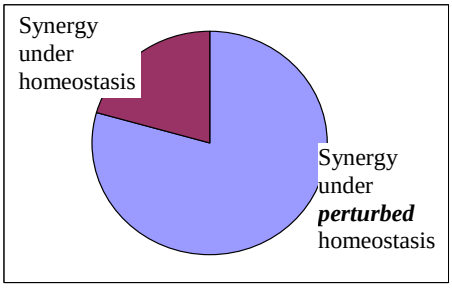

Fig. 10. Synergy effects at homeostatic and perturbed homeostatic conditions.

However, the synergy effects dominate under perturbed homeostasis (Fig. 8-10.). Enzyme E_2 is the key parameter, but of equal importance are substrate concentration consumed by the network and cofactor concentration regulated at a cell level. Opposite to the homeostatic results, here the total synergy effects dominate, 60%, Fig. 10.

4 Conclusions

Synergy analysis of reaction systems provided new information on role of model parameters on systems performance. The case of a stochastic model of HBV virus propagation reveals that synergy effects greatly dominate over individual parameters.

For MCA control of a branched pathway synergy effects become dominant only under assumption of perturbation of homeostasis.

In conclusion, synergy analysis provides a broader horizon for extraction of system information in a parameter space leading to better understanding of model performance, model improvement and reaction system control.

References:

[1] R.I. Cukier, C.M. Fortuin, K.E. Shuler, A.G. Petschek, J.H. Schaibly. Study of the sensitivity of the coupled reaction systems to uncertainties in rate coefficients. I theory, *The Journal of Chemical Physics*, Vol. 53, 1970, pp. 3873-3878.

[2] A. Saltelli, M. Ratto, S. Tarantola, F. Campolongo, Sensitivity analysis of chemical models, *Chemical Reviews*, Vol. 105, No. 7, 2005, pp. 2811-2827.

[3] Wolfram Research "*Mathematica*", v. 6.0, 2008.

[4] A. Saltelli, S. K. Chan, E.M. Scott, "*Sensitivity Analysis*", John Wiley & Sons, Ltd.,2000.

[5] J.B. Rawlings, J.G. Ekerdt, "*Chemical Reactor Analysis and Design Fundamentals*", Nob Hill Publishing, 2002.

[6] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, Vol. 81, 1977, pp. 2340-2361.

[7] G. Stephanopoulos, A. Aristidou, J. Nielsen, *Metabolic Engineering – Principles and Methodologies*, Academic Press, 1998.

[8] G. Stephanopoulos, Metabolic engineering: Enabling technology for biofuels production, *Metabolic Engineering*, Vol. 10., 2008, pp. 293-294.

[9] D. Visser, J.J. Heijnen, Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics, *Metabolic Engineering*, Vol. 5, 2003, pp. 164-176.

[10] P. Mendes, D.B. Kell, Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation, *Bioinformatics*, Vol. 14, No. 10, 1998, pp. 869-883.

[11] Z. Kurtanjek, *Statistical Metabolic Control Analysis*, WSEAS, Recent Advances in Mathematics and Computers in Biology and Chemistry, Prague, 2009, 124-129.