Robustness Analysis of the Heat Shock Response using semiquantitative Reasoning

Robustness Analysis of the Heat Shock Response using semiquantitative Reasoning

H. Koeppl S. Schindler E. Kashefi C. Messina V. Danos

Abstract— **The Heat Shock Response (HSR) is a universal procedure among organisms that repairs protein damage induced by heat and other stresses. It is a simple biological mechanism that is yet rich enough to be analyzed from the perspective of *robustness* and *efficiency*. A series of papers by El-Samad *et al.* [1], [2] have presented a control theoretical approach for studying HSR in *Escherichia coli* bacteria. They argue that the complexity of the HSR control mechanism is necessary to achieve the observed robustness of such biological systems.**

We extend the control theoretical approach, applying interval analysis to Lyapunov's indirect method, and consider the robustness of the HSR with respect to uncertainties in the individual chemical reaction rates. In order to design alternative control mechanisms for the HSR we compute the optimal control to the protein damage-repair cycle. Furthermore we propose a novel reduced order model of the HSR.

I. Introduction

Advances in molecular biology offer a great potential to develop new cures and treatments to human diseases, to provide bioremediation solutions to environmental hazards and to genetically improve living organisms. However, the expectations has not being met, due to the complex nature of most traits. One prevailing problem in the quantitative description of mechanisms in molecular biology is the large measurement uncertainty. Adopting the framework of ordinary differential equations to describe the concentration changes in biochemical reactions networks the uncertain parameters are the reaction rate constants. Thus, mechanism descriptions in molecular biology often boil down to ordinary differential equations with uncertain parameters. An outgrowth of qualitative reasoning [3], an artificial intelligence effort to qualitatively describe physical systems, are semiquantitative differential equations (SQDE) [4]. Their description combines qualitative or uncertain knowledge about a physical system with exact knowledge. Such systems include differential equations, where the parameters are only specified by intervals. Extensions to classical interval analysis enable us to investigate dynamical properties of a SQDE, such as stability [5]. In this work the authors propose the application of these techniques to the description of biochemical reaction networks. The procedure will be exemplified using the ubiquitous HSR mechanism.

The work is organized as follows. In Section II, the HSR is briefly introduced. The main part of the paper, involving stability of differential equations in conjunction with interval analysis, is discussed in Section III. Considerations about the optimal steady state control of the HSR mechanism are given in Section IV. Section V presents a novel reduced order model of the HSR, while Section VI draws the conclusions.

H. Koeppl is with the Department of Electrical Engineering and Computer Science, UC Berkeley, CA, USA (heinz.koeppl@berkeley.edu).

S. Schindler is with the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany (schindle@mis.mpg.de).

E. Kashefi is with the Institute for Quantum Computing, University of Waterloo, Waterloo, Canada (ekashefi@gmail.com).

C. Messina is with Pioneer Hi-Bred International, Johnston IA, USA (charlie.messina@pioneer.com).

V. Danos is with Équipe Preuves, Programmes, Systèmes, CNRS and with Université Paris VII, Paris, France (vincent.danos@gmail.com).

II. The Heat Shock Response

Because several biological processes are highly conserved among organisms and species, the development of methodologies for simple biological systems could be portable to more complex ones. Such a case study is the HSR in *E. Coli* to remedy protein damage due to heat and other stresses. The HSR system is composed by an intricate network of proteases (e.g. Lon and FtsH), the transcription factor (σ^{32}), chaperones (e.g. GroEL/S and DnaK/J) and the machinery involved in DNA transcription (RNA polymerase, herein RNAP), RNA processing and mRNA translation and protein synthesis.

The factor σ^{32} plays a specific role in this system as it recognizes HSR gene promoters up-regulating their transcription. The resulting increase in chaperone and protease concentrations help proper protein folding or their degradation under stress conditions. The regulation of σ^{32} occurs at the translation level. Under low temperatures the translation start site rpoH is occluded impairing the translation of σ^{32}. Changes in mRNA structure under heat stress conditions allows σ^{32} translation, its interaction with RNAP and the transcription of chaperones and proteases. The activity of σ^{32} is also regulated through interactions with DnaK/J. Binding of these chaperones to σ^{32} limits its interaction with RNAP and therefore gene transcription. Another mechanism for the regulation of σ^{32} is through its degradation. During steady state, the degradation of σ^{32} is rapid through the action of proteases. Under heat shock stress, proteases are titrated by misfolded proteins reducing the degradation rate of σ^{32}. Rapid degradation of σ^{32} may require chaperones to recruit FtsH, therefore, under high stress conditions degradation is reduced due to low concentration of free DnaK/J. A detailed description of the system and a differential-algebraic model thereof is given in [1]. It is the basis for the present work, which for the sake of conciseness adopts the mathematical nomenclature of [1].

III. Stability Analysis

We use Lyapunov's indirect method in addition to Kharitonov's theorem as is suggested in [5]. Lyapunov's indirect method is the linearization of a nonlinear differential equation at its equilibrium point with respect to its state variables. This equilibrium point is stable if the real parts of all eigenvalues of the linearization are strictly negative. The eigenvalues are the roots of the characteristic polynomial $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$ where $\mathbf{I}$ is the identity matrix and $\mathbf{A}$ is the Jacobian. Therefore, the test of stability of the nonlinear system becomes a test whether the characteristic polynomial of its Jacobian is Hurwitz, i. e., all roots have strictly negative real parts.

An interval polynomial is a set of polynomials where the coefficients may vary within intervals. Kharitonov's theorem [6] gives a simplified test whether all instances of an interval polynomial are Hurwitz by checking the Hurwitz property of only four polynomials. Given an interval polynomial $\{a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n : a_0 \in [a_0^-, a_0^+], \ldots, a_n \in [a_n^-, a_n^+]\}$ the four Kharitonov polynomials are

$$p^{+-} = a_0^+ + a_1^-\lambda + a_2^-\lambda^2 + a_3^+\lambda^3 + a_4^+\lambda^4 + \cdots + a_n^\pm\lambda^n$$
$$p^{++} = a_0^+ + a_1^+\lambda + a_2^-\lambda^2 + a_3^-\lambda^3 + a_4^+\lambda^4 + \cdots + a_n^\pm\lambda^n$$
$$p^{-+} = a_0^- + a_1^+\lambda + a_2^+\lambda^2 + a_3^-\lambda^3 + a_4^-\lambda^4 + \cdots + a_n^\pm\lambda^n$$
$$p^{--} = a_0^- + a_1^-\lambda + a_2^+\lambda^2 + a_3^+\lambda^3 + a_4^-\lambda^4 + \cdots + a_n^\pm\lambda^n.$$

Every instance of the interval polynomial is Hurwitz if and only if its four Kharitonov polynomials are Hurwitz. In lower dimensions the theorem of Anderson, Jury and Mansour [7] states that one needs to check the Hurwitz property only of p^{+-} for $n = 3$, of p^{+-} and p^{++} for $n = 4$, and of p^{+-}, p^{++} and p^{-+} for $n = 5$. In the case of $n = 2$ the test for stability is simply reduced to checking that the coefficients of p^{--} are positive.

In the following we apply the above framework to the reduced heat shock model in order to perform a stability analysis in the presence of parametric uncertainty. The reduced heat shock model is given in the supplementary information to [1] as

$$\frac{\mathrm{d}[\sigma_t^{32}]}{\mathrm{d}t} = \eta(T) - \alpha_0[\sigma_t^{32}] - \alpha_s\frac{\alpha K_s K_f [DnaK_t]^2}{\Gamma - \kappa[DnaK_t] + K_s[DnaK_t](1 + \alpha K_f[DnaK_t])}[\sigma_t^{32}]$$

$$\frac{\mathrm{d}[DnaK_t]}{\mathrm{d}t} = -\alpha_d[DnaK_t] + K_d\frac{\Gamma - \kappa[DnaK_t]}{\Gamma - \kappa[DnaK_t] + K_s[DnaK_t](1 + \alpha K_f[DnaK_t])}[\sigma_t^{32}], \quad (1)$$

with $\Gamma \equiv 1 + K_u[P_t]$ and $\kappa \equiv K_u(K_T) + K_f)/K_T$. The reduced model consists of two differential equations for the concentration of the σ-factor $[\sigma_t^{32}]$, and the concentration of chaperones $[DnaK_t]$ which depend on 11 parameters and are coupled nonlinearly.

With Lyapunov's indirect method the system of differential equation is linearized with respect to the state variables $[\sigma_t^{32}]$ and $[DnaK_t]$. The resulting Jacobian is analyzed regarding its eigenvalues. The system is asymptotically stable for small perturbations around the equilibrium value if and only if the real part of each eigenvalue is strictly negative. Note that the equilibrium points of the system and the eigenvalues of its Jacobian are effected by the uncertainty in the parameters. Kharitonov's theorem gives a sufficient and necessary condition to determine whether the system is stable for uncertain parameter values. Using interval arithmetic the range of the coefficients of the characteristic polynomial is determined by the possible range of the slope of the components of the Jacobian matrix. These slope intervals are deduced either by measurements or by biological reasoning.

TABLE I

Rounded intervals for real-valued equilibrium point

The value of p gives the percentage with which the parameters are varied both up and down from the given value. The equilibria are calculated before and after heat shock, denoted by $T1$ ($37^\circ C$) and $T2$ ($42^\circ C$), respectively.

p		$[\sigma_t^{32}]$	$[DnaK_t]$
10	T1	[86.00, 150.78]	[5537.95, 7886.97]
	T2	[178.80, 283.43]	[9490.919, 13324.39]
20	T1	[67.28, 209.11]	[4633.66, 9904.34]
	T2	[136.98, 348.60]	[7896.03, 16885.41]
30	T1	[55.10, 284.94]	[3913.12, 11424.47]
	T2	[91.38, 424.22]	[6688.43, 19139.78]
40	T1	[41.61, 451.52]	[3168.20, 20970.58]
	T2	[81.93, 569.69]	[5458.16, 24410.61]
50	T1	[32.42, 465.84]	[1627.48, 15451.44]
	T2	[68.20, 871.76]	[4090.76, 27357.12]
60	T1	[19.86, 1144.45]	[1548.45, 24641.82]
	T2	[50.23, 1340.95]	[2647.19, 80165.40]
70	T1	[14.32, 1267.01]	[1867.39, 26303.98]
	T2	[35.33, 2665.70]	[3191.49, 44515.27]
80	T1	[7.12, 5315.91]	[1460.13, 46885.40]
	T2	[23.90, 5263.83]	[2496.36, 82721.87]
90	T1	[4.72, 23248.36]	[428.56, 133125.58]
	T2	[10.69, 9433.86]	[732.80, 293025.14]

We calculate the equilibrium points of the system and minimize and maximize their positions with respect to the parameters that can vary in the predetermined parameter intervals. The resulting intervals for the position of the considered equilibrium point are given in Table I.

The slope intervals for the entries of the Jacobian matrix can be calculated by the same method. We minimized and maximized them over the equilibrium intervals of S_t and D_t, and the parameter intervals. The Kharitonov's polynomials p^{--} are given in Tab. II. We see that the system described by (1) is stable at least for the parameter perturbations not exceeding 80%.

IV. Optimal Steady State Control

To extract the unique features of biological control mechanisms it is worthwhile to compare them to state-of-the-art control designs from control theory. As discussed in [1] one has multiple strategies for the design of the control signal $[DnaK_f]$ for the biological plant, i.e., the repair-damage-

TABLE II

KHARITONOV'S POLYNOMIALS AND THE REAL PARTS OF ITS ROOTS

Parameters can vary p in [%], the temperature is denoted by $T1$ ($37°C$) before heat shock and $T2$ ($42°C$) after heat shock, respectively.

p		p^{--}
10	T1	$0.12 + 1.26\lambda + \lambda^2$
	T2	$0.31 + 2.26\lambda + \lambda^2$
20	T1	$0.03 + 2.21\lambda + \lambda^2$
	T2	$0.09 + 3.56\lambda + \lambda^2$
30	T1	$0.01 + 3.35\lambda + \lambda^2$
	T2	$0.03 + 4.75\lambda + \lambda^2$
40	T1	$0.002 + 5.83\lambda + \lambda^2$
	T2	$0.008 + 6.27\lambda + \lambda^2$
50	T1	$0.0004 + 6.28\lambda + \lambda^2$
	T2	$0.0018 + 8.11\lambda + \lambda^2$
60	T1	$0.0001 + 10.22\lambda + \lambda^2$
	T2	$0.0002 + 11.33\lambda + \lambda^2$
70	T1	$0.00004 + 12.72\lambda + \lambda^2$
	T2	$0.00007 + 18.84\lambda + \lambda^2$
80	T1	$5.35 \times 10^{-6} + 42.08\lambda + \lambda^2$
	T2	$6.01 \times 10^{-6} + 39.70\lambda + \lambda^2$
90	T1	$-222.00 + 255.24\lambda + \lambda^2$
	T2	$-597.50 + 109.62\lambda + \lambda^2$

cyle

$$[P_{fold}] \xrightarrow{k(T)} [P_{unfold}]$$
$$[P_{unfold}] + [DnaK_f] \rightleftarrows [P_{unfold} : DnaK] \rightarrow [P_{fold}].$$

A useful quantity in designing new control strategies is the minimum number of repair proteins in the plant $[DnaK_p] = [DnaK_f] + [P_{unfold} : DnaK]$, required to keep the protein damage below a predetermined level ρ for constant temperature. We define protein damage as the relative number of unfolded proteins

$$\rho = \frac{[P_t] - [P_{fold}]}{[P_t]}.$$

The optimal steady state level $[DnaK_p]$ turns out to be

$$[DnaK_p] \geq \frac{(k(T) - \nu)(k(T)k' + [P_t]\nu)}{(k(T) + 1)k(T)\nu}, \qquad (2)$$

with $k' = \frac{1}{K_8}$, $k(T) = \frac{K_{fold}}{K(T)}$ and $\nu = \rho(k + 1) - 1$. For $k \gg \nu$ and $kk' \ll [P_t]\nu$, that holds for the considered HSR scenario, we obtain the simple approximate relation

$$[DnaK_p] \gtrapprox \frac{[P_t]}{k(T)}. \qquad (3)$$

With this, the steady values for $[DnaK_p]$ read

$$[DnaK_p] \geq \begin{cases} 1.05 \times 10^4 \,\text{molec/cell} & \text{for} \quad T = 37°\text{C} \\ 1.93 \times 10^4 \,\text{molec/cell} & \text{for} \quad T = 42°\text{C} \end{cases}$$

and

$$[DnaK_p] \gtrapprox \begin{cases} 1 \times 10^4 \,\text{molec/cell} & \text{for} \quad T = 37°\text{C} \\ 2 \times 10^4 \,\text{molec/cell} & \text{for} \quad T = 42°\text{C}, \end{cases}$$

for (2) and its approximation (3), respectively. The results match the observed level of the full-order differential-algebraic model as shown in Fig. 1. The observable offset at the lower temperature is subject to future analysis.

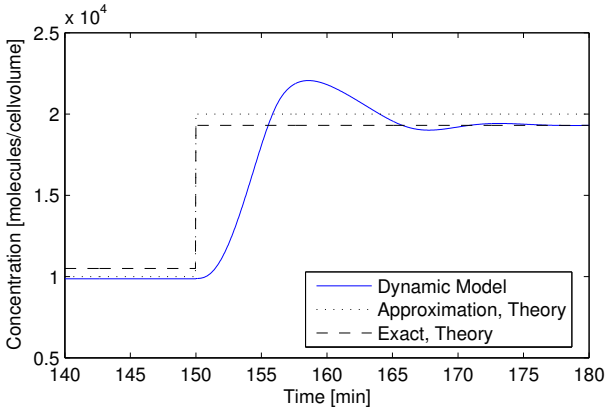

Fig. 1. Theoretical steady state level of the concentration of chaperones in the repair-damage cycle for $T = 37°$C and $T = 42°$C and the time evolution of chaperones obtained by the full-order differential-algebraic model.

V. MODEL REDUCTION

The reduced order, two-dimensional model (1) of the HSR qualitatively reflects the dynamics of the original 31-dimensional differential-algebraic description. While it is able to approximately reproduce the steady state levels of $[\sigma_t^{32}]$ and $[DnaK_t]$, the reduced model does not capture the transients during temperature changes. In order the capture the transient behavior we devised an alternative reduced order, 4-dimensional model. The model has as its additional states, the concentration of folded proteins in the cell $[P_{fold}]$ and the concentration of messenger RNAs for the chaperones $[mRNA(DnaK)]$. The model reads

$$\begin{aligned} \frac{\mathrm{d}[mRNA(DnaK)]}{\mathrm{d}t} =& K_{tr_1}\tilde{s}([\sigma_t^{32}], [DnaK_t], [P_{fold}]) \\ & - \alpha_{mRNA}[mRNA(DnaK)] \\ \frac{\mathrm{d}[DnaK_t]}{\mathrm{d}t} =& K_{tl}[mRNA(DnaK)] \\ & - \alpha_{prot}[DnaK_t] \\ \frac{\mathrm{d}[\sigma_t^{32}]}{\mathrm{d}t} =& K_{tl}\eta(T)[mRNA(\sigma^{32})]_0 \\ & - \alpha_{FtsH}\tilde{f}([\sigma_t^{32}], [DnaK_t], [P_{fold}]) \\ \frac{\mathrm{d}[P_{fold}]}{\mathrm{d}t} =& K_{fold}\,p([P_{fold}], [DnaK_t]) \\ & - K(T)[P_{fold}]. \end{aligned} \qquad (4)$$

The involved functions are defined as follows:

$$\tilde{s}([\sigma_t^{32}], [DnaK_t], [P_{fold}]) \equiv s(\xi([\sigma_t^{32}], [DnaK_t], [P_{fold}]))$$

with

$$s([\sigma_f^{32}]) \equiv \frac{K_2K_9[ph_t][\sigma_f^{32}]\frac{[RNAP_{t2}]}{[D_t](K_3+K_2K_{12}[\sigma_f^{32}])}}{1 + K_2K_9[\sigma_f^{32}]\frac{[RNAP_{t2}]}{[D_t](K_3+K_2K_{12}[\sigma_f^{32}])}}$$

and

$$\xi([\sigma_t^{32}], [DnaK_t], [P_{fold}]) \equiv \frac{1}{2a}\left(-b + \sqrt{b^2 - 4ac}\right),$$

with

$$a \equiv \frac{K_{tr_2}}{K_{tr_1}} K_2 K_4 K_5 K_{12} [D_t][DnaK_t] d([DnaK_t], [P_{fold}])$$
$$b \equiv \frac{K_{tr_2}}{K_{tr_1}} K_3 K_4 K_5 [D_t][DnaK_t] d([DnaK_t], [P_{fold}]) + K_2 K_{12} [D_t][RNAP_{t2}] - K_2 K_{12} [D_t][\sigma_t^{32}]$$
$$c \equiv K_3 [D_t][\sigma_t^{32}].$$

Furthermore

$$\tilde{f}([\sigma_t^{32}], [DnaK_t], [P_{fold}]) \equiv \frac{K_4 K_5 \frac{K_{tr_2}}{K_{tr_1}} [DnaK_t] d([DnaK_t], [P_{fold}])}{1 + K_4 K_5 \xi([\sigma_t^{32}], [DnaK_t], [P_{fold}]) d([DnaK_t], [P_{fold}])} \times \xi([\sigma_t^{32}], [DnaK_t], [P_{fold}])$$

with

$$d([DnaK_t], [P_{fold}]) \equiv \frac{1}{2K_8} \left(-\beta + \sqrt{\beta^2 + 4K_8 [DnaK_t]} \right),$$

where

$$\beta \equiv 1 + K_8([P_t] - [P_{fold}] - [DnaK_t]).$$

Finally we have the synthesis function for the folded proteins

$$p([P_{fold}], [DnaK_t]) \equiv K_8 \frac{[P_t] - [P_{fold}]}{1 + K_8 d([DnaK_t], [P_{fold}])} d([DnaK_t], [P_{fold}]).$$

Besides the constants already defined in the supplementary material of [1], we introduce one additional constant $[RNAP_{t2}] = 1342$ (molec/cell).

To evaluate the novel reduced order model we simulated a temperature up-step from $T = 37°\mathrm{C}$ to $T = 42°\mathrm{C}$ degree Celsius. The time evolution of $[\sigma_t^{32}]$ and $[DnaK_t]$ for the full-order model as well as for both reduced order models are shown in Fig. 3 and Fig. 2, respectively. The model (4) strikingly accurate reproduces the transient behavior of the 31-dimensional differential-algebraic model. The results indicate that the major players for the heat response mechanism have been correctly identified. Because (4) was generated deductively from the full-order model all other state variables present in the full-order model can be computed from the four states variables of (4).

VI. Conclusions

One prevailing problem in the quantitative description of mechanisms in molecular biology is the large measurement uncertainty. We applied methods from the artificial intelligence community, namely semiquantitative reasoning to cope with the parametric uncertainty in biochemical reaction networks. These concept in conjunction with classical nonlinear control theory allowed us to determined the stability margin of the 2-dimensional HSR model. To enhance the accuracy of compact HSR models with respect to the full-order differential-algebraic model we deduced a novel 4-dimensional differential model that supersedes the 2-dimensional model in accuracy. Furthermore, we computed

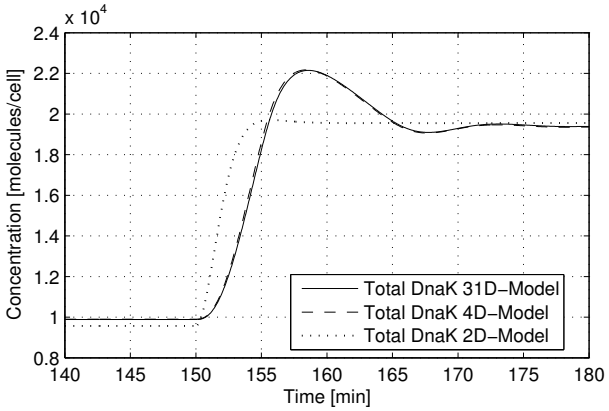

Fig. 2. Time evolution of the chaperones $[DnaK_t]$ for a temperature up-step from $T = 37°\mathrm{C}$ to $T = 42°\mathrm{C}$ at $t = 150\,\mathrm{min}$ for the full-order model, the 4-dimensional model (4) and the 2-dimensional model (1).

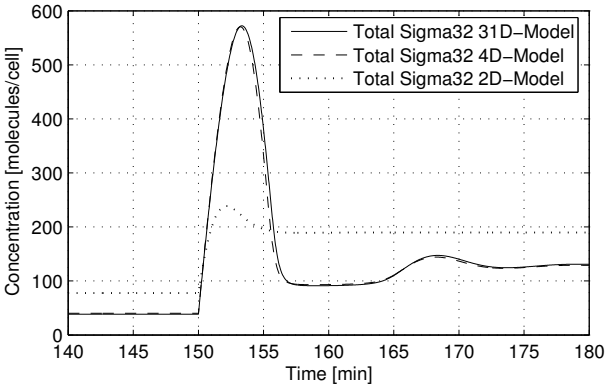

Fig. 3. Time evolution of the σ-factor $[\sigma_t^{32}]$ for a temperature up-step from $T = 37°\mathrm{C}$ to $T = 42°\mathrm{C}$ at $t = 150\,\mathrm{min}$ for the full-order model, the 4-dimensional model (4) and the 2-dimensional model (1).

the optimal steady concentration of repair proteins for a given temperature, such that the protein damage stays below a predetermined threshold. The outlined methodology is very promising for applications in systems biology because of its generality and portability.

References

[1] H. El-Samad, J. C. Kurata H., Doyle, C. A. Gross, and M. Khammash, "Surviving heat shock: Control strategies for robustness and performance," *PNAS*, vol. 102, no. 8, pp. 2736–2741, 2005.

[2] H. El-Samad, S. Prajna, A. Papachristodoulou, J. Doyle, and M. Kammash, "Advanced methods and algorithms for biological network analysis," *Proceedings of the IEEE*, vol. 94, no. 4, pp. 832–853, April 2006.

[3] B. J. Kuiper, *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge*. Cambridge, MA, USA: MIT Press, 1994.

[4] H. Kay, "SQSIM: a simulator for imprecise ODE models," *Computers and Chemical Engineering*, vol. 23, no. 1, pp. 27–46, November 1998.

[5] M. W. Hofbaur, "Lyapunov methods for semiquantitative simulation," Ph.D. dissertation, Graz University of Technology, Graz, Austria, 1999.

[6] V. L. Kharitonov, "Asymptotic stability of an equilibrium position of a family of systems of linear differential equations," *Differential Equations*, vol. 14, pp. 1483–1485, 1979.

[7] B. D. O. Anderson, E. I. Jury, and M. Mansour, "On robust Hurwitz polynomials," *IEEE Transactions on Automatic Control*, vol. 32, no. 10, pp. 909–913, 1987.