Proceedings of the American Control Conference
Anchorage, AK May 8-10, 2002

Prediction uncertainty from models and data

Michael Frenklach, Andrew Packard, Pete Seiler
Mechanical Engineering, University of California, Berkeley 94720
{myf,pack}@me.berkeley.edu guiness@jagger.me.berkeley.edu

Abstract

We present an approach to uncertainty propagation in dynamic systems, exploiting information provided by related experimental results along with their models. Our computational procedure draws from ideas and tools that are now common in robust control theory. A case study on a well-known database of methane combustion experiments and models demonstrates the viability of our proposed method.

1 Introduction

In this paper, we study the problem of accurately predicting the range of possible outcomes of a modeled physical process, knowing, to some accuracy, the outcomes of several related, but different processes. We are motivated by the current state of model-based and experimental research in chemical reaction networks. This field is characterized by several interrelated, relevant facts:

- Processes are complex, though physics-based governing equations are widely accepted;
- Uncertainty in process behavior exists, but much is known regarding "where" the uncertainty lies in the governing equations;
- Semi-isolated aspects of processes are studied experimentally;
- Numerical simulations of processes, with uncertain parameters "fixed" to certain values, may be performed reliably.

Each experimental result contains information about the uncertainty in the process model. Yet, there is no systematic manner in which to combine this information, and no coordination among researchers as to the form by which information and data should be shared. Most of the time, the information transfer occurs at a highly processed level, in the familiar "read my paper" mode. As a result of individual data processing, the critical correlations within the data and compound uncertainty of the overall model are simply lost. To our knowledge, the first attempt to address this problem was the GRI-Mech project [10], [17] (a brief description is given in section 4). This project was the first to systematically collect, combine, process, and retrieve "raw" experimental and theoretical information – all inter-related, often in different patterns, through their common dependence on a single multivariable model. The essence of the paradigm and its final product was a "living model": not the ultimately right one — possibly an utopic goal of the present science — but the best possible *today*, consistent with separate research results, and quantified against *all* available data.

Here, we go a step further, to numerical procedures which address an even broader issue: modeling realistic uncertainty in the prediction of a diversely-compound model's behavior. We reexamine the concept "model" and then describe numerical procedures we employ, and finally present some preliminary results. Although the case study in this paper is in the area of chemical reaction networks, the methods apply equally well to any scientific field characterized by the bulleted list above.

There are some similarities with the method presented here and those put forth by robust control theory (RCT, [16], [6]) over the past 20 years. In particular,

1. high-order, uncertain differential equation models;
2. $\mathcal{S}$-procedure to yield "outer bounds" and heuristic searches to get "inner bounds"
3. use of experimental data to refine bounds;
4. evidence of what appears to be good performance on a real problem.

Nevertheless, the method presented in this paper departs in many ways from the RCT approach. For one thing, the questions posed here are different, most notably the central role played by (3), and the assumptions that lead to its relevance. Other differences are

- systems considered are (at least as modeled) highly nonlinear, and medium-to-large sized;
- unconventional model transformations that in RCT, with linear fractional uncertainty in linear models, and $\mathcal{H}_\infty$ criterion are more rigorous.

2 Nature of a Predictive Model

Development of predictive models marks human civilization. Ancient models were philosophical in nature and relied entirely on logical deduction; one example

of this can be the concept of corpuscular (i.e., atomic) nature of matter arrived to by the Greeks two millennia ago. The invention of geometry and algebra in the Middle East was motivated by and applied to agricultural problems. Early scientific discoveries were expressed in terms of simple phenomenological laws, like Fick's Law of diffusion and Boyle's Law of gases. Development of calculus expressed these laws in terms of differential equations and formulated solutions to their linearizations. The advent of computer provided means to solve these differential equations in their complete nonlinear formulation, such as the Navier Stokes equation of fluid mechanics or Schrodinger equation of quantum mechanics.

Thus, the concept of a model and the "physical" form it assumes have changed with time: from conceptual statements to simple algebraic relationships to differential equations to numerical algorithms to computer programs and files. The present problems facing our society—such as global warming, earthquake preparedness, safety of transport of nuclear waste, and pollutant emission from automobile engines—call for integration of a variety of computer programs, each solving numerical problems in a different discipline. The overriding concern for such complex models is their reliability: predictability, authenticity, uncertainty, etc. In other words, one has to have a certain degree of confidence to apply these models for political decisions, economic forecast, or design and manufacturing of automotive engines.

The primary difficulty, common to essentially all fields of science and engineering, is the fact that more authentic models introduce larger numbers of parameters. The general expectation is that advancement in science should provide means to establish these parameters with required accuracy. The experience both supports and disproves this expectation. Indeed, the unprecedented advance in scientific instrumentation (e.g., laser spectroscopy) and computer technology (the increase in speed and memory and the reduction in cost) provides powerful means to determine the parameter values on sound experimental and theoretical grounds. At the same time, however, this often (if not always) involves another model, either instrumental or/and theoretical, which in turn introduces additional parameters. The end result is that even most accurately determined model parameters have uncertainties.

The current practice of model development/use is based on the following paradigm: (a) A model is postulated introducing a set of parameters; (b) A "unique" set of parameter values are determined from experiment or/and theory, and (ideally) supplied along with a corresponding set of individual uncertainties; and (c) the model is then applied to conditions of interest employing the unique set of parameter values. The natural uncertainties of the underlying experiment and theory must somehow be transferred into the final prediction uncertainty using the uncertainties that were assigned to the model parameters.

Experience shows that this conventional paradigm does not lead to a desirable quality of prediction. Each model parameter has associated with it an interval of uncertainty. Taken together, the uncertainties of all parameters form a hypercube in the parameter space. Each point of this hypercube is consistent with accepted experimental results since each coordinate (parameter) individually belongs to it's corresponding interval of uncertainty. But, some parts of the hypercube fit the experimental data base better than others. Typically, the central point of the hypercube, whose coordinates correspond to the individual best-fit values of the individual parameters, is not necessarily the best-fit point for the combined set of parameters. So, the better-fit parts form a low-dimensional manifold, which is the result of correlations in our experimental data and hence our knowledge of the parameter vector. A methodology that implicitly or explicitly samples preferentially the manifold and not the entire hypercube volume should provide a more realistic estimate of model uncertainty than superposition of individual parameter errors.

The specific focus of this paper is representation and propagation of uncertainty of dynamic models. We suggest that the combination of solution mapping and polynomial optimization provide a foundation for evaluation of realistic prediction uncertainties of large-scale dynamic models. Our suggested approach lends itself naturally towards a more data-centric approach to model updating.

3 Formulation

Throughout this section, we refer to *physical processes, models, experiments, data*, etc. Clarification of meanings and interpretation, and notational choices are necessary. Associated with a physical process P, consider other entities, listed below.

Sym	Meaning
Y_P	Outcome variable (scalar) of the process P
M_P	Math model of Y_P (usually physics-based)
x_P	Uncertain parameters in M_P
S_P	Surrogate model of Y_P
E_P	Experimental realization of P
D_P	Exp. data (ie., *measured* outcome) from E_P
ϵ_P	Error tolerance associated with D_P

Together, the set $\{Y_P, M_P, x_P, S_P; E_P, D_P, \epsilon_P\}$ constitutes a *model of process* P. For a given process, some of these items may be missing. As an example, for one such process, we may have a mathematical model, but no available experimental data. When there are many

processes under consideration, we usually index them not by name (ie., P, Q, etc.) but by an integer index (process $1, 2, \ldots, n$).

Our prediction uncertainty concepts involve $n+1$ processes, ($P_1, P_2, \ldots, P_n$ and P_0) coupled together by the common dependence that their mathematical models have on m uncertain parameters $\{x_1, x_2, \ldots, x_m\}$. Any of the mathematical models may actually depend only on a subset of these parameters. The processes $\{P_i\}_{i=1}^n$ have experimental data D_i, and are usually associated with specialized laboratory experiments. The *surrogate models* S_i are algebraic functions, reduced from the mathematical model M_i (surrogate models, and the reductions which map differential equations into algebraic functions are described in Section 3.1). P_0 is a more complex scenario, without experimental data, and our goal is to reliably predict its outcome. P_0 is referred to as "the predicted process," and the P_i are "experiment processes." As with S_i, S_0 is the reduction of M_0 into a surrogate, algebraic function.

Each of the uncertain quantities x_j are known a priori to lie in sets X_j, reflecting prior information about the uncertainties. In many cases, the prior information (compactly written as $x \in X$) is not rich enough to bound the possible values of $P_0(x)$ to a range accurately reflecting the compound uncertainty of the entire knowledge base. Usually, the experimental results must be used in addition to this prior information. Through the surrogate model S_i, information about x is obtained implicitly via the constraint

$$\|S_i(x) - D_i\| \leq \epsilon_i \tag{1}$$

Any $x \in X$ which is consistent with **all** of the experiments (ie., satisfying the constraint in (1) for all $1 \leq i \leq n$) yields a possible outcome of P_0, namely $S_0(x)$. All such possible outcomes constitute the predicted outcome set of P_0. Our goal is to understand the "extremes" of this predicted outcome set:

$$L^* := \min_{\substack{x \in X \\ |S_i(x) - D_i| \leq \epsilon_i}} S_0(x) \tag{2}$$

$$H^* := \max_{\substack{x \in X \\ |S_i(x) - D_i| \leq \epsilon_i}} S_0(x) \tag{3}$$

The optimizations posed in (2) and (3) are typically difficult to solve. A more modest analysis problem is to compute outer bounds, $\bar{H}$ and $\underline{L}$, and inner bounds $\underline{H}$ and $\bar{L}$ satisfying: $\underline{L} \leq L^* \leq \bar{L}$ and $\underline{H} \leq H^* \leq \bar{H}$. Computation of these bounds answers the question "What are the extreme values of the prediction model which are consistent with results from collections of related experiments?" There are additional interesting queries: Will a particular experiment (either new, or the refinement of an existing one) add to the accuracy of the predictive model?, Which subset of experiments (if redone) offer the most potential for improving the predictive model?, How well does an experiment need to be carried out to improve the accuracy of the predictive model by a fixed percent? All of these are variants on the original question, and improvements in computing bounds for (2) and (3) will facilitate answering them.

Optimization issues that arise in the outer bounds $\bar{H}$ and $\underline{L}$ are outlined in Section 3.2 and require specific forms of the S_i. Transforming the large scale, complex differential equation models M_i and M_0 into the reduced functions S_i and S_0 is discussed in section 3.1.

3.1 Building Surrogate Models

The general optimizations posed in Section 3 are computationally intensive even for well-behaved, well-parametrized algebraic functions S_i and S_0. The situation is further complicated by the fact these functions are not immediately obtained from the governing equations M_i and M_0. Consider a dynamic model, M, described by a set of first-order nonlinear ODEs,

$$\dot{y}(t) = f(y(t), t, x), \qquad y(0) = y_0 \tag{4}$$

where $x \in \mathbf{R}^m$ is a parameter vector. For most functions f, this equation does not possess a closed-form solution, and expressing some property of the solution, ϕ_y, called *a response* (e.g., the steady-state value $\bar{y}$, the peak excursion from $y_0, \max_{t \geq 0} \|y(t) - y_0\|$) as a function of the parameter vector x is only possible in a tabular/numerical sense.

The essence of the *solution mapping* technique, [8], [12], [9], is approximation of responses ϕ_y by simple algebraic expressions, i.e., $S_y(x) \approx \phi_y(x)$ within a subset X of parameter space $\mathbf{R}^m$, referred to hereafter as the *active* parameter subset. The approximating functions S_y are obtained using the methodology of the response surface technique [3], [4], by means of a relatively small number of computer simulations, referred to as *computer experiments*. They are performed at preselected combinations of the parameter values and the entire set of these combinations is called a *design* of computer experiments. The computer experiments are performed using the complete dynamic model (4) and the functions S_y obtained in this manner are referred to as *surrogate* models.

Once developed, the surrogate model S_y replaces ϕ_y, the solution of the original dynamic model, whenever evaluation of ϕ_y at a given $x \in X$ is required. In our work we have used quadratic functions for S_y, with coefficients determined via computer experiments arranged in a special order, called *factorial design*. These designs originate from analysis of variance, with the objective of minimizing the number of computer experiments to be performed to gain the information required [3], [4]. Factorial designs have found extensive use in

experimental and process development work, and have begun recently being applied to computer experiments (e.g., [15] and [13]). The use of a polynomial form for the surrogate model also ties in closely with our optimization formalism put forth next in section 3.2.

3.2 Computing the Outer and Inner Bounds

If all of the surrogate models are affine, then the optimizations in (2) and (3) are linear programs. However, it is often the case that an affine function is not adequate to represent the dependence. Hence, a natural, nontrivial, and useful extension is to consider quadratic surrogate models. These enlarge the applicable domain of a surrogate model but introduce nonconvexity and computational complexity into subsequent predictions.

The main computational subproblem is an indefinite quadratic program: for $x \in \mathbf{R}^m$, bound

$$\begin{bmatrix} 1 \\ x \end{bmatrix}^T M_0 \begin{bmatrix} 1 \\ x \end{bmatrix}, \text{ subject to } \begin{bmatrix} 1 \\ x \end{bmatrix}^T M_i \begin{bmatrix} 1 \\ x \end{bmatrix} \in [\alpha_i, \beta_i] \tag{5}$$

for $i = 1, 2, \ldots, N$. Note that $N = n + m$ because (2) and (3) have n constraints due to our a priori knowledge of x and an additional m constraints due to the experimental results. Here, all M_i are symmetric $(1+m) \times (1+m)$ real matrices. Upper and lower bounds to this are the outer bounds $\bar{H}$ and $\underline{L}$ defined following equations (2) and (3). Inner bounds are computed using standard constrained optimization.

The convex relaxation we use for the upper and lower (outer) bounds is known as the $\mathcal{S}$-procedure, and is well-known in RCT, with numerous applications summarized in [11], [5].

Given $M_i, \alpha_i, \beta_i, \gamma \in \mathbf{R}$, define $M_i^\beta := \begin{bmatrix} \beta_i & 0 \\ 0 & 0 \end{bmatrix} - M_i$, $M_i^\alpha := M_i - \begin{bmatrix} \alpha_i & 0 \\ 0 & 0 \end{bmatrix}$, $M_0^\gamma := \begin{bmatrix} \gamma & 0 \\ 0 & 0 \end{bmatrix} - M_0$. A specific convex optimization which yields an upper bound to the question in (5) is

$$\min_{\gamma,\lambda,\tau} \gamma \quad \text{subject to} \quad \begin{array}{l} \lambda_i \geq 0, \tau_i \geq 0 \\ M_0^\gamma - \sum_{i=1}^N (\lambda_i M_i^\beta + \tau_i M_i^\alpha) \succeq 0 \end{array}$$

An analogous expression is derivable for the lower bound. Here, $A \succeq 0$ means A is a symmetric, positive semidefinite matrix. In this problem, it is easy to prove that for each i, only one of λ_i and τ_i will be nonzero. Apriori, it is not known which are active, and it is easiest just to solve the optimization directly over these $2N + 1$ variables. However, the fact shows that the $\mathcal{S}$-procedure is unable to take advantage of both constraints implied by M_i. The optimization involves convex constraints and a convex (linear, actually) objective, and is easily solved. There is a great deal of current research (Semi-Definite Programming (SDP) or Linear Matrix Inequalities (LMI)) in developing and improving polynomial-time algorithms to find feasible points, and convergence to optimizers [2], [7], [5].

We recognize that the problem in (5) includes (as a special case) maximization of an indefinite quadratic function subject to indefinite quadratic constraints, and as such is computationally complex, [19]. We do not address or attempt to solve this complexity issue. Rather, we will simply aim for algorithms which do well bounding the optimizations for problems we encounter. Improvements on the $\mathcal{S}$-procedure, most notably the sum-of-squares (SOS) techniques, may prove useful in general, though on the example in section 5 were not significantly better than the $\mathcal{S}$-procedure. The SOS methodology can be traced back to [1], [20] for nonquadratic Lyapunov function generation, and [14] for feasibility calculations involving polynomial constraints. In the case of quadratic polynomials, the SOS procedure has been shown to often be superior to the simple-minded $\mathcal{S}$-procedure.

4 GRI-Mech: Model & Data Set

In this section, we describe a high-fidelity methane combustion chemistry model in widespread use for prediction of pollutant formation in natural gas combustion, ([10], [17] and `www.me.berkeley.edu/gri_mech`). The chemical reaction model, known as *GRI-Mech*, consists of 53 species (and hence 53 coupled 1st-order ODEs) and 300+ reactions. The model entails a total of 102 uncertain parameters which characterize a large number of uncertain, temperature-dependent rate constants. We assume that each uncertain parameter is known to lie in the interval $[-1, 1]$. This hypercube (in $\mathbf{R}^{102}$) of uncertain parameters is denoted $\mathcal{H}$.

A collection of 77 experiments, all involving methane combustion, but under different physical manifestations, and different conditions (such as temperature, pressure and reactor configuration) have been carried out by researchers around the world. The experiments are well-documented, have relatively simple geometry, and considered repeatable and modelable. In fact, in these experiments, the prevailing consensus is that only the rate-constant uncertainty prevents one from accurate numerical simulation. All other relevant physical aspects are understood, accounted for, and well modeled. Typical experiments involve flow-tube reactors, stirred reactors, shock tubes, and laminar premixed flames, with outcomes such as ignition delay, flame speed, and various species concentration properties (time of peak, peak value, relative peaks).

In light of the complete understanding of the experiment, it is possible, apriori, to run computer "experiments," mapping the modeled relation between the rate constants in *GRI-Mech* and the outcome, effectively replacing the seemingly complex dynamic model with an algebraic function. The feasibility of this step requires

that the modeled outcome be dependent on the uncertain parameters in a fairly simple manner. Given this, one obtains simpler models which capture the behavior of (numerically obtained) solutions to the governing equations.

A feature of this specific set of experiments, revealed by numerical simulation, is that each simulated outcome is well-fit (a few percent error over the relevant range) with a quadratic function of a small (< 15) subset of the uncertain parameters. This desirable situation is probably due to a few different reasons:

- the experiments were designed with simple geometry and phenomenon to be strongly dependent on some of the parameters (those which the experimenters intended to learn about) and hopefully nearly unaffected by most.
- the state of knowledge in reaction chemistry is such that the size of the hypercube $\mathcal{H}$ is "small" relative to the outcome quantities of interest. (well-founded by physics, theory and consensus of community).

The *GRI-Mech Data Set* specifically refers to this collection of 77 experiments, surrogate models, and experimental results. They are linked through the methane combustion that is common to all.

This wealth of information about several inter-related complex processes provides an excellent opportunity to test and validate the bounding techniques described in section 3.2.

5 Bound quality on GRI-Mech Dataset

The viability of the method is evident from a small, initial study we have conducted on the natural-gas combustion models which form the *GRI-Mech Data Set*. Here we describe a modest, preliminary study we have carried out in this vein. *Our specific goal is:* predict the possible range of the outcome Y_{10} of process P_{10} using only the model S_{10}, and the information implicit in the models and experimental results associated with processes $P_1, P_2, \ldots, P_9$. So, here, P_{10} is playing the role of P_0 as described in section 3.

Surrogate models $S_1, S_2, \ldots, S_9$ together involve 17 of the 102 uncertain parameters (ie., 85 of the uncertain parameters play no appreciable role in the outcome of processes $1, 2, \ldots, 9$). Surrogate model S_{10} depends on 11 of these 17 parameters, and none of the other 85.

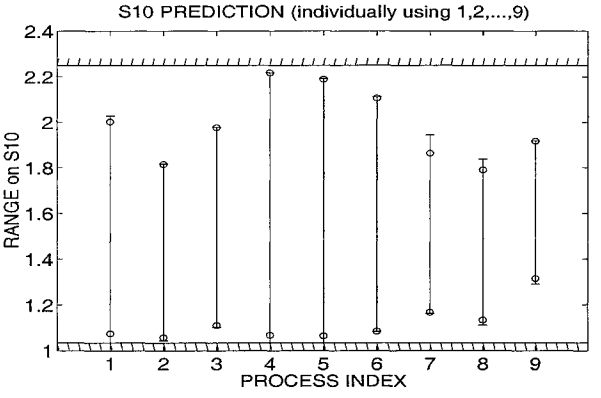
Using *only* the unit-interval bound (ie., $\mathcal{H}$) on the uncertain parameters, the $\mathcal{S}$-procedure bounds the values of Y_{10} (ie., values taken on by surrogate model S_{10}) to lie in [1.03 2.26]. The accepted experimental value is actually about 1.51, consistent with this simple prediction on the range.

Next, we consider how process #1 (and its model and its reported experimental result) can be used to reduce the feasible uncertain parameter set, from the apriori $\mathcal{H}$, to a smaller cube. At 1% (for instance) level of experiment uncertainty, the feasible uncertain parameter set is

$$\{x \in \mathbf{R}^{102} : |x_i| \leq 1, |S_1(x) - D_1| \leq 0.01\,|D_1|\}$$

Using simple constrained optimization, it is easy to verify the smallest cube (aligned with the coordinates), containing the feasible set is indeed still $\mathcal{H}$. So, this experiment, on its own, *is not able to reduce the "cube" description of the uncertainty.* The exact same conclusions are reached when considering processes $2, \ldots, 9$.

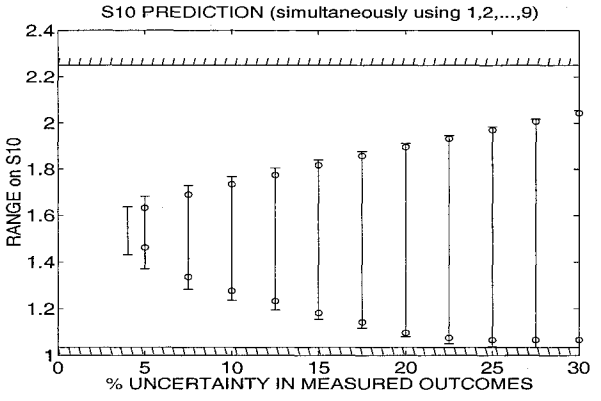
On the other hand, if we use the result of experiment i (though this time at a much coarser 5% experiment uncertainty) to bound the range of S_{10}, using the constraints as they are, following the prescription in section 3.2, we obtain the figure below:



The hashed-lines are the bounds on S_{10}, only using the original hypercube bound on x. The vertical "errorbar"-like lines represent the outer bounds. These are guaranteed to contain the range taken on by S_{10}, given that S_i is within 5% of the measured value, D_i. The circles are the results from the inner bounds, attempting to solve the constrained optimization directly. The gaps between the ends of the vertical bars, and the circles are the ambiguity we have in knowing the actual possible range of Y_{10}.

Finally, we simultaneously use the information in processes 1-9, to refine the range of surrogate model S_{10}. For simple illustration, we assumed a common percentage error tolerance on all experimental results of processes 1-9, and used a single $\mathcal{S}$-procedure calculation to yield the outer bounds on the value that could be taken on by S_{10}, consistent with the information regarding 1-9. Using the *SeDuMi* package, [18], running on a Pentium II, 500 MHz., a upper and lower outer bound pair take about 1 second to compute. The results for percentage uncertainty in measured outcome ranging from 4 to 30% are discussed next. At nearly

50% uncertainty level (not shown), the experimental data from processes 1-9 offers little additional information in bounding the response of S_{10} beyond the apriori [1.03 2.26]. At 4% though, the possible range of S_{10} has been narrowed to [1.43 1.64].



Inner bounds are also shown, which are values of S_{10} taken on at feasible x. These illustrate the potential gaps between our outer bound of the range and its actual spread. At 4%, the constrained optimization is not able to find any feasible points, so circles are not shown. At 3% (not shown), the $\mathcal{S}$-procedure detects an inconsistency between the surrogate models 1-9, experimental data, and experimental tolerance, meaning that no parameter in $\mathcal{H}$ exists that can simultaneously explain the data in processes 1-9 at this small level of measurement uncertainty.

This small example demonstrates the power of the ideas underlying our approach and the feasibility of its practical implementation.

6 Acknowledgements

We gratefully acknowledge support from NSF CTS-0113985 and helpful discussions with Laurent El-Ghaoui, Bernd Sturmfels, Pablo Parrilo and Eric Wemhoff. Eric also assisted with the examples.

7 Bibliography

1. A. Barkin and A. Zelentsovsky, "Method of power transformations for analysis and stability of nonlinear control systems," *Systems and Control Letters*, vol. 3, pp. 303-310, 1983.
2. A. Ben-Tal, L. El Ghaoui, A. Nemirovski. Robust convex programming. in *Semidefinite Programming and Applications,* R. Saigal, L. Vandenberghe, H. Wolkowicz, Eds, Kluwer Academic.
3. G. E. P. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building*, Wiley, 1978.
4. G. E. P. Box and N. R. Draper, *Empirical Model-Building and Response Surfaces*, Wiley, 1987.
5. S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*, SIAM, 1994.
6. J.C. Doyle, "Analysis of feedback systems with structured uncertainties", *IEE Proceedings*, **129**, Part D, pp. 242–250, (1982).
7. L. El Ghaoui and S.I. Niculescu. Robust decision problems in engineering: An LMI approach. Chapter 1 of *Recent Advances on Linear Matrix Inequality Methods in Control*, L. El Ghaoui and S.I. Niculescu, editors, SIAM, 1999.
8. M. Frenklach, "Modeling," In *Combustion Chemistry*, W. C. Gardiner, Jr. (Ed.), Springer-Verlag, New York, 1984, Chapter 7.
9. M. Frenklach and D. Miller, "Statistically Rigorous Parameter Estimation in Dynamic Modeling Using Approximate Empirical Models," *AIChE J.*, vol. 31, pp. 498-500 (1985).
10. M. Frenklach, H. Wang and M. J. Rabinowitz, "Optimization and Analysis of Large Chemical Kinetic Mechanisms Using the Solution Mapping Method—Combustion of Methane," *Prog. Energy Combust. Sci.*, vol. 18, pp. 47-73, 1992.
11. Alexandre Megretski and A. Rantzer, "System analysis via integral quadratic constraints," *IEEE Trans. Autom. Control*, v.42, n.6, pp. 819-830, June 1997.
12. D. Miller and M. Frenklach, "Sensitivity Analysis and Parameter Estimation in Dynamic Modeling of Chemical Kinetics," *Int. J. Chem. Kinet.*, vol. 15, pp. 677-696, 1983.
13. M. D. Morris, T. J. Mitchell, and D. Ylvisaker, "Bayesian Design and Analysis of Computer Experiments: Use of Derivatives in Surface Prediction," *Technometrics*, vol. 35, pp. 243-255, 1993.
14. P. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," PhD. thesis, Caltech, Control and Dynamical Systems, `www.cds.caltech.edu`, May 2000.
15. J. Sacks, W. J. Welch, T. J. Mitchell, and H. Wynn, "Design and Analysis of Computer Experiments," *Stat. Sci.*, vol. 4, pp. 409-435, 1989.
16. M.G. Safonov, "Stability and Robustness of Multivariable Feedback Systems," *MIT Press*, 1980.
17. G. P. Smith, D. M. Golden, M. Frenklach, N. W. Moriarty, B. Eiteneer, M. Goldenberg, C. T. Bowman, R. K. Hanson, S. Song, W. C. Gardiner, Jr., V. V. Lissianski, and Z. Qin, `www.me.berkeley.edu/gri_mech/`.
18. J. Sturm, *SeDuMi* optimization software, from `www.unimass.nl/~sturm/software/sedumi.html`
19. J. Tierno and J. Doyle, "Control problems and the polynomial time hierarchy," *Proceedings 34th IEEE Conference on Decision and Control*, New Orleans, 13-15 Dec. 1995, pp. 2927-31.
20. A. Zelentsovsky, "Nonquadratic lyapunov functions for robust stability analysis of linear uncertain systems," *IEEE Trans. Auto. Control*, vol. 39, no. 1, pp. 135-138, January 1994.