MODELS OF STOCHASTIC SYSTEMS

Andrzej W. Ordys
Industrial Control Centre, University of Strathclyde, UK

Joseph Bentsman
Mechanical and Industrial Engineering Department, University of Illinois at Urbana-Champaign, USA

Keywords: System, random process, probability distribution, stochastic process, white noise, Markov process, auto-regressive process, moving-average process, state-space system.

Contents

Summary

Stochastic systems can be defined as systems which are evolving in time and, for each time instant, are characterized by randomly assumed values of their states and, possibly, parameters, dimension, structure, and other dynamic features. This chapter describes different methods of modeling stochastic systems and provides basic properties of such systems and models. The chapter starts with recollection of basic properties of random variables, introducing the probability density function, joint and conditional probabilities and defining basic operators over probability densities such as expectation or covariance. A very important type of probability density functions, the Gaussian function is introduced. Next, the attention moves to stochastic processes with definitions of correlation and crosscorrelation, introduction of white noise and Wiener process as important practical models of stochastic processes. The stationary stochastic processes are explained and the idea of ergodicity is presented. The next section describes linear models often used in analysis of stochastic processes. Those models are based on Markov processes. The spectra of linear stochastic processes are defined and polynomial models of discrete time systems are presented. Finally, the mixed stochastic deterministic description of systems is introduced which is very important from practical point of view, enabling analysis of systems in which there are deterministic inputs along with stochastic inputs.

1. Introduction

Practical experience tells us that all the surrounding reality is not deterministic. It is impossible to tell exactly what will happen in the future. It is equally true in gambling, e.g. playing lottery or betting on horse races as it is in engineering systems. Consider a typical engineering process, for instance a measurement of temperature of exhaust gas from the gas turbine combustion in a gas turbine power plant. Assume that the turbine works in steady-state conditions producing constant power with constant supply of gas and of air. Even though those conditions are constant, the temperature will exhibit oscillations, as shown in Fig .1. Those oscillations are caused by a combination of different factors and influences which are impossible or difficult to fully comprehend, for instance the changes in the ambient temperature, the changes in the calorific value of the fuel, the changes in the ambient temperature, the changes in the calorific value of the fuel, the changes in air pressure or vibrations of the turbine fixings. Faced with this situation, an engineer responsible for the system has two choices. The first is to say that

the fluctuations observed in the measurement are unavoidable and therefore the system should be dealt with as if those did not exist. This would lead to *deterministic* procedures and deterministic models of the system. The second possibility is to admit the existence of uncertainty and try to investigate whether some knowledge of the system and of the system undesirable uncertain effects are minimized. In the gas turbine example, the first approach would concentrate on assuring the desired value of the output gas temperature, in practice this would mean adjusting the *mean value* of this temperature. However, the second approach may lead to minimization of the fluctuations of the temperature, it is to minimization of its *variance*.

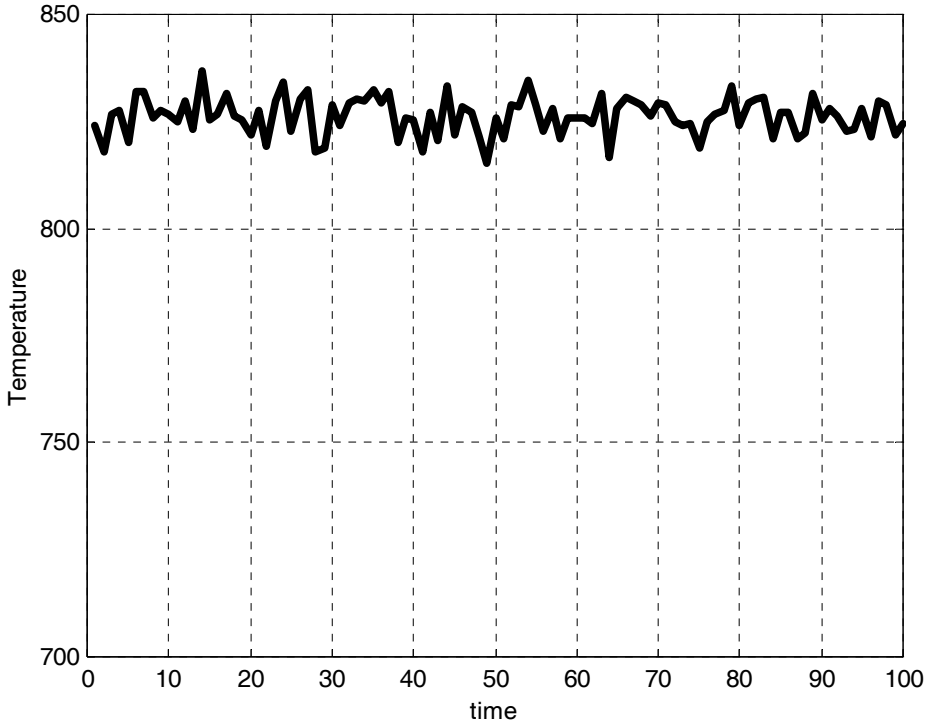

Figure 1: Exhaust gas temperature from the gas turbine

If the presence of uncertainty in a system is acknowledged then a model is needed to show how this uncertainty is transmitted through the system how it influences the system outputs and, possibly, how it changes in time. Among the models of uncertainty which are used in engineering systems the most popular are : *Stochastic models* and *fuzzy models (*see *Fuzzy Control Systems*). Recently, the models resulting from *chaos theory* are gaining increasing publicity, especially for analysis of non-linear and large systems (see *Control of Bifurcations and Control*). This chapter is devoted to stochastic models of uncertain systems.

In the first section we start with a short recollection of main properties and functions of random variables. The Gaussian probability density is introduced and independent (and dependent) random variables are defined.

In the second section, the stochastic processes are introduced, Autocorrelation and crosscorrelation are defined. The idea of ergodicity is introduced which would allow to

average over time rather than over realizations of process.

In the third section we show how the stochastic processes can be generated at the outputs of dynamical systems. Polynomial and state-space models of systems are considered. This section concentrates on linear models. Treatment of non-linear effects is a rather complex matter, extending the scope of this chapter. It is only briefly mentioned within this section.

2. Random Variables

This section starts with recalling the basic properties of random variables. It explains the terms: probability density function, conditional probability, and independent random variables. It also introduces the Gaussian distribution function.

2.1. Probability Density Function

Example 1 *when throwing a dice with differently colored sides, referred to as conducting the basic experiment, there are six possible outcomes, or samples, or elementary events, of having a particular color at the top. Such an elementary event will be further denoted as* ω. *The set of all possible outcomes of a basic experiment is referred to as the universal set or the sample space, further denoted as* Ω. *We can postulate that the way the dice is built and thrown does not result in one color occurring at the top more often than the other ones. This permits us to size up the intuitive likelihood of the occurrence of an event in terms of the probability measure: if we define the probability of any color showing up at the top as* 1, *then the probability of a particular color showing at the top is* $\frac{1}{6}$. *We now define a function, referred to as a random variable, say,* $x \equiv x(\omega)$, *which associates the particular color, i.e., the particular sample* ω *in the sample space* Ω, *with a real number from a set* {1,2,…6}. *Then the probability of a particular value out of this set the random variable takes is* $\frac{1}{6}$, *which, for example for random variable taking the value of* 2 *can be denoted as* $\Pi(x = 2) = \frac{1}{6}$. *The probability that the random variable has a value less than* ζ, *denoted as* $\Pi(x < \zeta)$, *can be presented as a function of* ζ *as in the Table* 1.

ζ	…	-1	0	1	2	3	4	5	6	7	…
$\Pi(x < \zeta)$	0	0	0	0	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$	$\frac{6}{6}$

Table 1: Probability distribution when throwing a dice

Example 2 *When throwing two dices, the probability function is determined over two dimensional space of possible results and has a value* $\frac{1}{36}$ *for each of the* 36 *possible combinations. The probability that the random variable* x_1 *has value less than* ζ_1 *and the variable* x_2 *has value less than* ζ_2 *is given by the Table* 2.

ζ_2 \ ζ_1	…	0	1	2	3	4	5	6	7	…
⋮	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{6}{36}$
3	0	0	0	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{8}{36}$	$\frac{10}{36}$	$\frac{12}{36}$	$\frac{12}{36}$
4	0	0	0	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{9}{36}$	$\frac{12}{36}$	$\frac{15}{36}$	$\frac{18}{36}$	$\frac{18}{36}$
5	0	0	0	$\frac{4}{36}$	$\frac{8}{36}$	$\frac{12}{36}$	$\frac{16}{36}$	$\frac{20}{36}$	$\frac{24}{36}$	$\frac{24}{36}$
6	0	0	0	$\frac{5}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{20}{36}$	$\frac{25}{36}$	$\frac{30}{36}$	$\frac{30}{36}$
7	0	0	0	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{18}{36}$	$\frac{24}{36}$	$\frac{30}{36}$	$\frac{36}{36}$	$\frac{36}{36}$
⋮	0	0	0	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{18}{36}$	$\frac{24}{36}$	$\frac{30}{36}$	$\frac{36}{36}$	$\frac{36}{36}$

Table 2: Probability density function for the experiment with two dices

This relation permits us to introduce function $\Pi(x < \zeta)$ vs. $\zeta \in R^n$ for a random vector variable x of dimension n,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \tag{1}$$

referred to as the distribution function. In general, x can be discontinuous. Introduce the equation

$$\begin{aligned} &\Pi(x < \zeta) = \Pi(x_1 < \zeta_1 \text{ and } x_2 < \zeta_2 \text{ and } \ldots \text{and } x_n < \zeta_n) \\ &= \int_{-\infty}^{\zeta} \varphi(x)dx = \int_{-\infty}^{\zeta_1}\int_{-\infty}^{\zeta_2} \ldots \int_{-\infty}^{\zeta_n} \varphi(x)dx_n \ldots dx_2 dx_1. \end{aligned} \tag{2}$$

This equation defines function $\varphi(x)$, referred to as the probability density function. Function $\Pi(x < \zeta) \equiv \Pi(\zeta)$ will be referred to as the probability distribution. A more rigorous measure-theoretic description of $\Pi(\zeta)$ can be introduced, but will be omitted. Notice that if x is discontinuous random variable (discrete-continuous, or combining discrete and continuous values), $\varphi(x)$ contains delta functions. For the random variables which are discrete in space (e.g. throwing a dice in the example above) the equivalent of the probability density function is probability of occurrence of a given event. Then, to calculate the distribution function for discrete random processes the integral in Eq. 2 is replaced by a summation.

The universal set gives rise to a sigma-algebra: the totality of all possible outcomes of a combined experiment consisting of the union of basic experiments and their complements with respect to the universal set.

2.2. Expectation Operator

If $g(x)$ is a function of a random variable described by the differentiable probability density $\varphi(x)$, the expectation operator of g is defined as:

$$\overline{g} = E\{g(x)\} = \int_{-\infty}^{\infty} (g(x)\varphi(x))\, dx. \tag{3}$$

2.2.1. The Mean Value

The mean value is defined as the expectation operator for the function $g(x) \equiv x$, i.e.:

$$\overline{x} = \int_{-\infty}^{\infty} x\varphi(x)\, dx. \tag{4}$$

2.2.2. The Covariance Matrix

The covariance matrix P will be defined as the expectation operator with the function $g(x) = (x - \overline{x})\ (x - \overline{x})^T$, i.e.:

$$P = E\left\{(x - \overline{x})\ (x - \overline{x})^T\right\} \int_{-\infty}^{\infty} (x - \overline{x})\ (x - \overline{x})^T \varphi(x)\ dx. \tag{5}$$

2.3. The Gaussian Probability Density Function

The Gaussian probability density function of the random vector of dimension n is given by the expression:

$$\varphi(x) = \frac{1}{\sqrt{2(\pi)^n \det(P)}} \exp\left[-\frac{1}{2}(x - \overline{x})^T\ P^{-1}(x - \overline{x})\right]. \tag{6}$$

where $\overline{x}$ is the mean value and P is the covariance matrix. The shape of the Gaussian probability density is presented in Fig.2, for scalar case and in Fig. 3 for the case when x is a vector of size 2. From Eq. 6 it is transparent that the Gaussian probability density is fully determined if the mean value $\overline{x}$ and the covariance P are known. Therefore, if it is known that a random vector variable is Gaussian then knowledge of its mean value and co-variance matrix is sufficient to fully characterize its properties.

This is a very useful feature, which is widely applied in analysis of linear stochastic systems. In such systems, it can be shown that if initial probability density is Gaussian then, with progress of time, the probability densities remain Gaussian and therefore the process could be characterized by the changes to the mean value and to the covariance.

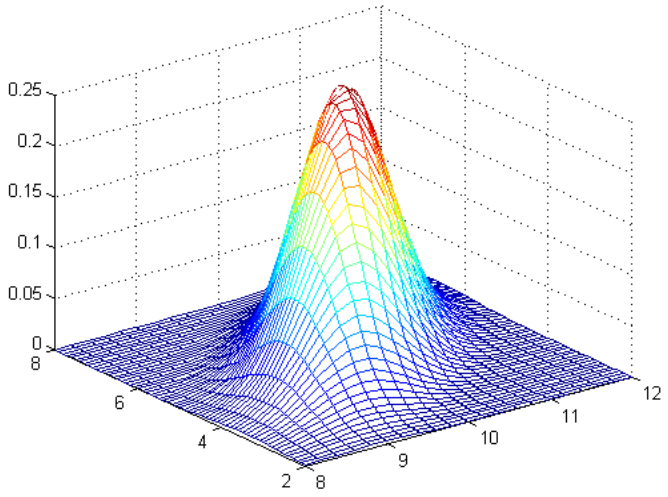

Figure 2: One-dimensional Gaussian probability density function

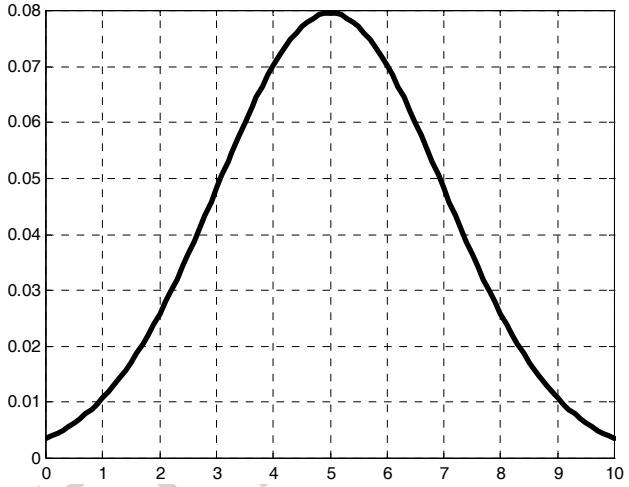

Figure 3: Two-dimensional Gaussian probability density function

2.4. Conditional Probability

Two or more vector random variables, e.g. x_1, x_2, x_3 can be put together in a larger vector

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{7}$$

where $x_1^T = [x_{11} \ldots x_{1n}]^T$, $x_2^T = [x_{21} \ldots x_{2m}]^T$ and $x_3^T = [x_{31} \ldots x_{3q}]^T$. A joint probability density function can be defined:

$$\varphi_{\text{joint}}(x) = \varphi\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right). \tag{8}$$

Then the conditional probability density function determines the probability density function when some of the random variables are known (e.g. have been measured). Assuming that the vector x_3 is known, from Bayes rule one obtains:

$$\varphi_{cond}(x_1, x_2 \,|\, x_3) = \frac{\varphi_{joint}(x)}{\varphi_3(x_3)} \tag{9}$$

where $\varphi_3(x_3)$ denotes the probability density function of a vector random variable x_3 and can be calculated from:

$$\begin{aligned}\varphi_3(x_3) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi_{joint}(x_1, x_2, x_3)\, dx_1\, dx_2 \\ &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \varphi_{joint}(x_1, x_2, x_3)\, dx_{11}\ldots dx_{1n} dx_{21}\ldots dx_{2m}. \end{aligned} \tag{10}$$

2.5. Conditional Expectation Operator

The conditional expectation of function $g(x_1, x_2)$ will be obtained if the probability density function $\varphi(x)$ is replaced by the conditional probability e.g.:

$$E\{g(x_1, x_2)\,|\,x_3\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x_1, x_2)\varphi_{cond}(x_1, x_2\,|\,x_3)\, dx_1 dx_2. \tag{11}$$

Notice that the conditional expected value is a function of x_3.

2.6. Independent Random Vectors

Two random vectors x_1 and x_2 are independent if

$$\varphi_{joint}(x) = \varphi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \varphi_1(x_1)\varphi_2(x_2). \tag{12}$$

As an obvious conclusion for independent random vectors:

$$\varphi_{cond}(x_1\,|\,x_2) = \varphi_1(x_1) \tag{13}$$

and

$$\varphi_{cond}(x_2 \,|\, x_1) = \varphi_2(x_2). \tag{14}$$

2.7. Characteristic Function

For a random vector variable x of dimension n define a vector s of dimension n. The characteristic function of the random variable x is the expectation operation for the function $g(x) \equiv \exp(jx^T s)$, i.e:

$$\Phi(s) = \int_{-\infty}^{+\infty} \exp\left(j\bar{x}^T s\right) Q(x)\, dx. \tag{15}$$

Notice that $\exp(jx^T s)$ is a scalar function of two vector arguments: x and s. Therefore, the expectation operation produces a scalar function of one vector argument: s.

2.7.1. Characteristic Function for Gaussian Probability Density

The characteristic function for Gaussian probability density is given by:

$$\Phi_{Gauss}(s) = \exp\left(j\bar{x}^T s - \frac{1}{2} s^T P s \right). \tag{16}$$

Substituting the Gaussian probability density function into the equation defining the characteristic function yields:

$$\begin{aligned}
&\Phi_{Gauss}(s) = \int_{-\infty}^{\infty} \exp(jx^T s) \frac{1}{\sqrt{(2\pi)^n \det(P)}} \exp\left[-\frac{1}{2}(x-\bar{x})^T P^{-1}(x-\bar{x}) \right] dx = \\
&= \frac{1}{\sqrt{(2\pi)^n \det(P)}} \int_{-\infty}^{\infty} \exp\left(jx^T s - \frac{1}{2}(x-\bar{x})^T P^{-1}(x-\bar{x}) \right) dx = \\
&= \exp\left(j\bar{x}^T s - \frac{1}{2} s^T P s \right).
\end{aligned} \tag{17}$$

2.8. Characteristic Function for Independent Random Vectors

If characteristic functions of random variables x_1 and x_2 are $\Phi_{x_1}(s_1)$ and $\Phi_{x_2}(s_2)$, respectively, and vectors x_1 and x_2 are independent, then:

$$\Phi_x(s) = \Phi_{\left[x_1^T \; x_2^T\right]}\left(\begin{bmatrix} s_1 \\ s_2 \end{bmatrix}\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left[j(x_1^T s_1 + x_2^T s_2) \right] \varphi_1(x_1)\, \varphi_2(x_2)\, dx_1 dx_2 =$$

$$\int_{-\infty}^{+\infty} \exp(jx_1^T s_1)\, \varphi_1(x_1)\, dx_1 \int_{-\infty}^{+\infty} \exp(jx_2^T s_2)\, \varphi_2(x_2)\, dx_2 = \Phi_{x_1}(s_1)\Phi_{x_2}(s_2).$$

-
-
-

Bibliography

Astrom K.J, (1970) *Introduction to Stochastic Control*, Academic Press, New York. [A classical textbook on stochastic control]

Box G.E.P. and Jenkins D.R., (1970) *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco. [A classical textbook on stochastic systems]

Gardiner C.W., (1985) *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd edition, Springer-Verlag [stochastic differential equations, application in physics and chemistry]

Grimble M.J. and Johnson M.A. (1988) *Optimal Control and Stochastic Estimation*, volumes I, II, Wiley. [Frequency domain methods in analysis and control design for stochastic systems]

Jazwinsky A. H., (1970) *Stochastic Processes and Filtering Theory*, Academic Press. [Introduction to stochastic differential equations through the use of mean square convergence and calculus]

Levine W.S., (1996) Linear Systems and White Noise, in: The Control Handbook, ed. W.S. Levine, CRC Press pp575-578.[Linear stochastic systems]

Meditch J.S., (1969) Stochastic Optimal Linear Estimation and Control, McGraw-Hill [Time-domain, state-space description of stochastic systems]

Miller B.M. and Pankov A.P., (2002) *Theory of Random Processes* (in Russian), Physmatlit. [A rigorous, concise and modern exposition of random processes and stochastic differential equations]

Petridis V and Kehagias Ath.,(1998) "A Multi-model Algorithm for Parameter Estimation of Time-varying Nonlinear Systems", Automatica, vol.34, no.4, pp.469-475.

Rabiner L.R.,(1989) "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, vol.77, No.2, pp.257-286. [Hidden Markov models]

Schuss Z., (1980) *Theory and Applications of Stochastic Differential Equations*, Wiley. [Stochastic differential equations and methods of their analysis]

Shwartz A., (1996) Discrete Time Markov Processes, in: *The Control Handbook*, ed. W.S. Levine, CRC Press, pp. 1057-1066[Introduction to Markov processes]

Soderstrom T., (1996) Linear Stochastic Input-Output Models, in: *The Control Handbook*, ed. W.S. Levine, CRC Press, pp. 10791088. [Linear Stochastic models]

Van Kampen N.G., (1981) *Stochastic Processes in Physics and Chemistry*, Elsevier. [Stochastic differential equations, applications in physics and chemistry]

Biographical Sketch

Andrzej Ordys was born in Sosnowiec, Poland. He obtained his Master of Science degree in Electronic Engineering and the Doctorate in Control Engineering, both from the Silesian Technical University in Gliwice Poland. He worked as system engineer in the Industrial Institute of Control Systems in Katowice

and then he joined the Department of Automatic Control at the Silesian Technical University. In 1991 he was awarded the British Council Fellowship at Oxford University and at Strathclyde University. In 1993 Dr. Ordys joined the Department of Electronic and Electrical Engineering, University of Strathclyde. He is presently the British Energy Senior Lecturer in Control Systems and the Technical Manager of the Industrial Control Centre at the University of Strathclyde, Glasgow, UK. He is Member of IEE and Chartered Engineer, and Member of IEEE and of IAPR. Dr. Ordys has been involved in the theoretical development of stochastic and predictive control theory. His current research interests are in optimal predictive control, modeling and simulation of power plants and implementation of advanced process control algorithms. He also works on the theory and algorithms for stochastic benchmarking of performance of controllers. Dr. Ordys recently gave invited plenary lectures on linear and non-linear predictive control. He is author and co-author of over forty publications. He is Editor of Book Reviews in the International Journal of Adaptive Control and Signal Processing.

Joseph Bentsman was born in Minsk, Belarus, in 1952. He received the Electrical Engineering Diploma from Byelorussian Polytechnic Institute, Minsk, in 1979 and the Ph.D. degree in electrical engineering from the Illinois Institute of Technology, Chicago, in 1984. From 1975 to 1980 he worked as an Engineer in the Design Bureau of Broaching Machine Tools, Minsk. In 1985 he was a Lecturer and Postdoctoral Research Fellow in the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor. At present he is an Associate Professor in the Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign. Dr. Bentsman's current research interests are in control of distributed parameter systems, wavelet networks, nonsmooth dynamics, and robust self-tuning control. He is a recipient of the 1989 National Science Foundation Presidential Young Investigator Award in Dynamic Systems and Control. He has served as an associate editor for the ASME Journal of Dynamical Systems, Measurement, and Control and an invited editor for the International Journal of Adaptive Control and Signal Processing. He has authored or co-authored over forty journal publications in control theory and its applications to power systems, electro-mechanical systems, steel casting, bio-medical systems, and chemical reactions.