PERGAMON

Automatica 37 (2001) 1397–1405

automatica

www.elsevier.com/locate/automatica

Brief Paper

Probabilistic solutions to some NP-hard matrix problems[☆]

M. Vidyasagar[a,*], Vincent D. Blondel[b]

[a]Tata Consultancy Services, Coromandel House, 1-2-10 S.P. Road, Hyderabad 500 003, India
[b]Department of Mathematical Engineering, Center for Systems Engineering and Applied Mechanics, Université catholique de Louvain, Avenue George Lemaître, 4 B-1348 Louvain-la-Neuve, Belgium

Received 16 July 1999; revised 24 October 2000; received in final form 4 March 2001

Abstract

During recent years, it has been shown that a number of problems in matrix theory are NP-hard, including robust nonsingularity, robust stability, robust positive semidefiniteness, robust bounded norm, state feedback stabilization with structural and norm constraints, etc. In this paper, we use standard bounds on empirical probabilities as well as recent results from statistical learning theory on the VC-dimension of families of sets defined by a finite number of polynomial inequalities, to show that for each of the above problems, as well as for still more general and more difficult problems, there exists a polynomial-time randomized algorithm that can provide a yes or no answer to arbitrarily small levels of accuracy and confidence. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: NP-hard; Matrix stability; VC-dimension; Interval matrices; Static output feedback

1. Introduction

During recent years, several researchers have explored the computational complexity of various problems arising in robust control theory and in matrix theory. Owing to these efforts, it is now known that several problems in matrix theory are NP-hard.[1] A survey of computational complexity results in systems and control can be found in Blondel and Tsitsiklis (2000). We give below a partial catalog of some such NP-hard problems. These problems can be grouped naturally into two categories: Problems of *analysis*, and problems of *synthesis*. Both types of problems are stated in terms of "interval matrices", which are defined next.

Given an integer n, let Y denote the subset of $\Re^{2n^2}$ defined by

$$Y := \{(\alpha_{ij}, \beta_{ij}),\ i, j = 1, \dots, n\colon\ \alpha_{ij}, \beta_{ij} \in \mathscr{Q}\ \forall i, j\},$$

where $\mathscr{Q}$ denotes the set of rational numbers. Let $\mathbf{y} \in Y$ be a typical element. The corresponding set $\mathbf{A}_{\mathbf{y}}$ is defined by

$$\mathbf{A}_{\mathbf{y}} := \{A \in \Re^{n \times n}\colon\ \alpha_{ij} \leqslant a_{ij} \leqslant \beta_{ij}\ \forall i, j\}.$$

Now let

$$Y_s := \{\mathbf{y} \in Y\colon\ \alpha_{ij} = \alpha_{ji} \text{ and } \beta_{ij} = \beta_{ji}\ \forall i, j\}.$$

For a typical element $\mathbf{y} \in Y_s$, define

$$\mathbf{A}_{s,\mathbf{y}} := \{A \in \Re^{n \times n}\colon\ A \text{ is symmetric and}$$
$$\alpha_{ij} \leqslant a_{ij} \leqslant \beta_{ij}\ \forall i, j\}.$$

The set $\mathbf{A}_{\mathbf{y}}$ is referred to as an "interval matrix" while $\mathbf{A}_{s,\mathbf{y}}$ is a "symmetric interval matrix".

With this notation we can now state several NP-hard problems, all of which pertain to *analysis*.

1. *Robust stability*: Given an element $\mathbf{y} \in Y$, determine whether every matrix in the set $\mathbf{A}_{\mathbf{y}}$ is stable, in the sense that all of its eigenvalues have negative real parts.
2. *Robust positive semidefiniteness*: Given a vector $\mathbf{y} \in Y_s$, determine whether every symmetric matrix in $\mathbf{A}_{s,\mathbf{y}}$ is positive semidefinite.

[☆]This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor S. Hara under the direction of Editor Roberto Tempo. This work was carried out while both authors were in their previous places of employment. M. Vidyasagar was with the Centre for Artificial Intelligence and Robotics, Bangalore, India. Vincent Blondel was with the Université de Liège, Belgium.

[*] Corresponding author.

E-mail addresses: sagar@hydbad.tcs.co.in (M. Vidyasagar), blondel@inma.ucl.ac.be (V. Blondel).

[1] See Garey and Johnson (1979) for a dated but still highly readable account of NP-completeness and NP-hardness, and Papadimitrou (1994) for a more up-to-date treatment.

PII: S0005-1098(01)00089-9

3. *Robust norm boundedness*: Given a vector $\mathbf{y} \in Y$ and a number $\gamma > 0$, determine whether the l_2-induced norm of every matrix in $\mathbf{A}_\mathbf{y}$ is less than or equal to γ; that is, determine whether $\gamma^2 I_n - A^t A$ is positive semidefinite for every $A \in \mathbf{A}_\mathbf{y}$.

4. *Robust nonsingularity*: Given a vector $\mathbf{y} \in Y$, determine whether every matrix in the set $\mathbf{A}_\mathbf{y}$ is nonsingular.

The NP-hardness of each of these problems is demonstrated in Nemirovskii (1993), Poljak and Rohn (1993), and Blondel and Tsitsiklis (1997).

Observe that the problem of robust nonsingularity can be restated in the following equivalent form: Given the element $\mathbf{y}$, the question becomes

$$\exists a_{11} \dots a_{nn} (a_{ij} \in [\alpha_{ij}, \beta_{ij}]) \wedge |A| = 0?$$

where $|A|$ denotes the determinant of the matrix A. In the above formula, the n^2 parameters $a_{11} \dots a_{nn}$ are "modified variables" whereas the $2n^2$ parameters α_{ij}, β_{ij} in $\mathbf{y}$ are "constants". Using standard methods in quantifier elimination theory (see, e.g., Tarski (1951)), it is possible to eliminate sequentially each of the n^2 variables in such a way that the above question eventually becomes equivalent to a *finite set of polynomial inequalities* involving only the $2n^2$ constants α_{ij} and β_{ij}. Then, in principle one would only have to substitute the specific values of the constants into this finite set of inequalities to answer the question of robust nonsingularity. This clearly shows that the problem of robust nonsingularity is decidable. Unfortunately, the difficulty with this approach is that general elimination algorithms take exponential time in the worst case.

The above questions all involve the *analysis* of an interval matrix family. The next two questions involve *synthesis*.

5. *Constant output feedback stabilization with constraints*: An instance of the constant output feedback problem consists of *three* matrices A, B, C, of dimensions $n \times n$, $n \times m$, and $p \times n$, respectively. The *constrained* output feedback question is: Does there exist an $m \times p$ output feedback matrix K such that $\alpha_{ij} \leqslant k_{ij} \leqslant \beta_{ij}\ \forall i,j$, and such that $A + BKC$ is a stable matrix? As shown in Blondel and Tsitsiklis (1997), this problem is NP-hard when C is the identity matrix, and so it certainly remains NP-hard when C is part of the problem instance. It is as yet unknown if there exists a polynomial time algorithm for the problem of knowing whether or not there exists an *unconstrained* matrix K such that $A + BKC$ is stable. This problem is shown to be decidable in Anderson, Bose, and Jury (1975) but the solution procedure given there is based on Tarski's elimination procedure and is not guaranteed to run in polynomial time.

6. *Simultaneous stabilization using constant output feedback*: Suppose one is given, not just one triplet of matrices (A, B, C), but rather a family of such triplets (not necessarily finite), where each matrix A_i has dimension $n \times n$, each matrix in B_i has dimension $n \times m$ and each matrix C_i has dimension $p \times n$. The problem now is to determine whether there exists an $m \times n$ "state feedback" matrix K such that $A_i + B_i K C_i$ is a stable matrix *for each i*. It is shown in Blondel and Tsitsiklis (1997) that this problem is NP-hard.

In the face of these and other negative results, one is forced to make some compromises in the notion of "solving" a problem. An approach that is recently gaining popularity is the use of *randomized* algorithms, which are not required to work "all" of the time, only "most" of the time. Specifically, the probability that the algorithm fails can be made arbitrarily small (but of course not exactly equal to zero). In return for this compromise, one hopes that the algorithm is *efficient*, i.e., runs in polynomial-time.[2] The idea of using randomization to solve control problems is suggested, among other places, in Ray and Stengel (1991) and Marrison and Stengel (1994). In Khargonekar and Tikku (1996) and Tempo, Bai, and Dabbene (1997), randomized algorithms are developed for a general function minimization problem, and these are then applied to a few specific problems such as: (i) determining whether a given controller stabilizes every plant in a structured perturbation model, (ii) determining whether there exists a controller of a specified order that stabilizes a given fixed plant, and so on.

The objective of the present paper is to show that it is possible to develop *polynomial-time* (often abbreviated as "polytime") algorithms for *each* of the above NP-hard problems. In the case of Problems 1–4, which are problems of analysis, the randomized algorithms are based on a well-established classical result known as the Chernoff bound. In the case of Problem 6, which is a problem in synthesis, the randomized algorithm is based on recent results from statistical learning theory on the VC-dimension of a family of sets defined by a finite number of polynomial inequalities. Note that Problem 5 is not studied separately since it is a special case of Problem 6. The present paper actually develops a broad framework for deriving such polytime randomized algorithms for *any problem* where the decision question to be answered yes or no can be posed in terms of a finite number of polynomial inequalities. Hence, the approach is not limited to the specific problems discussed here. No doubt other researchers would be able to apply this approach to other problems as well.

2. Chernoff bounds and Vapnik–Chervonenkis theory

In this section, a very brief overview is given of a powerful theory often referred to as Vapnik–Chervonenkis (VC) theory after its originators. Book-length

[2] See Papadimitrou (1994) for the definition of a polynomial time algorithm.

treatments of VC theory can be found in Vapnik (1982), Vapnik (1995), and Vidyasagar (1997a).

We begin with a classical result that forms the basis of *Monte Carlo simulation*. Suppose X is a set, P is a probability measure on X, and A is a (measurable) subset of X. Suppose it is desired to estimate the measure $P(A)$. A popular method of doing this is to generate independent and identically distributed (i.i.d.) samples $x_1, \dots, x_m \in X$ distributed according to P, and to define

$$\hat{P}(A; \mathbf{x}) := \frac{1}{m} \sum_{j=1}^{m} I_A(x_j), \tag{1}$$

where $\mathbf{x} \in X^m$ denotes the m-tuple $[x_1 \cdots x_m]^t$ and $I_A(\cdot)$ is the *indicator function* of the set A defined by

$$I_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Note that $\hat{P}(A; \mathbf{x})$ is just the fraction of the i.i.d. samples $x_1, \dots, x_m$ that belong to A. The number $\hat{P}(A; \mathbf{x})$ is referred to as the *empirical probability* of the set A based on the multisample $\mathbf{x}$, and is itself a random variable on the product space X^m. A classical result known as the Chernoff bound (see Chernoff (1952)) states that, for each $\varepsilon > 0$,

$$\begin{aligned} &P^m\{\mathbf{x} \in X^m: \hat{P}(A; \mathbf{x}) - P(A) > \varepsilon\} \leqslant \exp(-2m\varepsilon^2), \\ &P^m\{\mathbf{x} \in X^m: P(A) - \hat{P}(A; \mathbf{x}) > \varepsilon\} \leqslant \exp(-2m\varepsilon^2), \\ &P^m\{\mathbf{x} \in X^m: |\hat{P}(A; \mathbf{x}) - P(A)| > \varepsilon\} \leqslant 2\exp(-2m\varepsilon^2). \end{aligned} \tag{2}$$

In other words, after m i.i.d. samples have been drawn, it can be asserted with confidence $1 - 2\mathrm{e}^{-2m\varepsilon^2}$ that the empirical probability $\hat{P}(A; \mathbf{x})$ is within the specified accuracy ε of the true probability $P(A)$. Note that this bound is completely independent of the nature of the underlying set X. In particular, if X is a subset of a Euclidean space $\Re^k$ for some integer k, then the number of samples is completely independent of the dimension k. In contrast, if "gridding" methods are used to estimate $P(A)$, then the number of grid points is exponential in k.

The above method, known as Monte Carlo simulation, is absolutely standard. Now we come to some recent results. Suppose we are given, not a single set $A \subseteq X$, but a *family* $\mathscr{A}$ of (measurable) subsets of X. Given a multisample $x_1, \dots, x_m \in X^m$, let us define the empirical probability $\hat{P}(A; \mathbf{x})$ of each set $A \in \mathscr{A}$ as in (1) above. Next, define

$$q(m, \varepsilon; \mathscr{A}) := P^m\{\mathbf{x} \in X^m: \exists A \in \mathscr{A} \\ \text{s.t. } |\hat{P}(A; \mathbf{x}) - P(A)| > \varepsilon\}. \tag{3}$$

Thus, after m i.i.d. samples have been drawn, it can be asserted with confidence $1 - q(m, \varepsilon; \mathscr{A})$ that *every* empirical probability $\hat{P}(A; \mathbf{x})$ is within ε of the corresponding true probability $P(A)$, for each $A \in \mathscr{A}$.

Definition 1. The family of sets $\mathscr{A}$ is said to have the property of *Uniform Convergence of Empirical Probabilities* (*UCEP*) if $q(m, \varepsilon; \mathscr{A}) \to 0$ as $m \to \infty$ for each $\varepsilon > 0$.

Note that if $\mathscr{A}$ is a finite set, then it follows by repeated application of the Chernoff bound that

$$q(m, \varepsilon; \mathscr{A}) \leqslant 2|\mathscr{A}|\exp(-2m\varepsilon^2).$$

Hence, every finite collection of sets has the UCEP property. However, infinite collections of sets need not have this property. See Vidyasagar (1997a), Section 3.1 for several examples of infinite collections of sets that do not possess the UCEP property.

In a seminal paper, Vapnik and Chervonenkis (1971) gave necessary and sufficient conditions for a given collection of sets to have the UCEP property in terms of the expected value of a combinatorial parameter now known as the Vapnik–Chervonenkis (VC)-dimension, which is defined next.

Definition 2. Let X be a given set and let $\mathscr{A}$ be a collection of subsets of X. A set $S = \{x_1, \dots, x_n\} \subseteq X$ is said to be *shattered* by $\mathscr{A}$ if, for every subset $B \subseteq S$, there exists a set $A \in \mathscr{A}$ such that $S \cap A = B$. The *Vapnik–Chervonenkis dimension* of $\mathscr{A}$, denoted by $VC\text{-dim}(\mathscr{A})$, equals the largest integer n such that there exists a set of cardinality n that is shattered by $\mathscr{A}$.

Thus, if VC-dim$(\mathscr{A}) = d$, then (i) there exists at least one set of cardinality d that is shattered by $\mathscr{A}$, and (ii) *every* set of cardinality greater than d *fails* to be shattered by $\mathscr{A}$. See Vidyasagar (1997a), Section 4.1 for several examples of the computation of the VC-dimension of various collections of sets.

The main theorem proved in Vapnik and Chervonenkis (1971) is the following. See also Vidyasagar (1997a), Theorem 7.2, p. 198, and Theorem 10.2, p. 302.

Theorem 1. 1. *Suppose $\mathscr{A}$ has finite VC-dimension, say $VC\text{-dim}(\mathscr{A}) \leqslant d$. Then*

$$q(m,\varepsilon;\mathscr{A}) \leqslant 4\left(\frac{2\mathrm{e}m}{d}\right)^d \exp(-m\varepsilon^2/8), \\ \forall m \geqslant d, \ \varepsilon > 0, P. \tag{4}$$

Thus, $\mathscr{A}$ has the property of distribution-free uniform convergence of empirical probabilities. Moreover, the inequality

$$q(m, \varepsilon; \mathscr{A}) \leqslant \delta$$

is satisfied provided at least

$$m \geqslant \max\left\{\frac{16}{\varepsilon^2}\ln\frac{4}{\delta}, \frac{32d}{\varepsilon^2}\ln\frac{32\mathrm{e}}{\varepsilon^2}\right\}$$

samples are drawn.

2. *Conversely, suppose $\mathscr{A}$ has the property of distribution-free uniform convergence of empirical probabilities; then the VC-dimension of $\mathscr{A}$ is finite.*

In view of Theorem 1, it is clear that it is of paramount importance (i) to be able to show that a given collection of sets has finite VC-dimension, and (ii) furthermore, to obtain either exact values, or failing that, reasonably good upper bounds for the VC-dimension. During the past few years, several researchers (working mostly though not exclusively in the field of neural networks) have derived several useful results along precisely these lines. Section 10.3 of Vidyasagar (1997a) contains a fairly complete summary of many of the known results. However, there is one result that is particularly appropriate for the class of problems studied here; it is presented next.

Suppose $X \subseteq \Re^k, Z \subseteq \Re^l$ for some integers k, l, respectively, and suppose $\tau_1(x,z), \ldots, \tau_t(x,z)$ are polynomials in x, z. For each $x \in X$, $z \in Z$, each polynomial inequality "$\tau_i(x,z) > 0$" evaluates to either "true" or "false". Now, suppose $\phi(x,z)$ is a *Boolean formula* obtained from the expressions "$\tau_i(x,z) > 0$" using the standard logical connectives $\neg$ (not), $\vee$ (or), $\wedge$ (and) and $\Rightarrow$ (implies).

Let 1 correspond to "true", 0 to "false", and define, for each $z \in Z$,

$$A_z := \{x \in X\colon \phi(x,z) = 1\}, \quad \text{and} \quad \mathscr{A} := \{A_z\colon z \in Y\}.$$

Then, $\mathscr{A}$ is a collection of subsets of X. The objective is to obtain an upper bound for the VC-dimension of $\mathscr{A}$.

The following theorem is a refinement of a result from Karpinski and Macintyre (1995, 1997), and is proved in this refined form in Vidyasagar (1997a), Corollary 10.2, p. 330.

Theorem 2. *With all symbols as above and* $\mathrm{e} = 2.7182\ldots$, *we have*

$$\text{VC-}dim(\mathscr{A}) \leqslant 2l\,\mathrm{lg}(4\mathrm{e}rt), \tag{5}$$

where lg *denotes the logarithm to the base* 2.

3. Randomized algorithms for matrix problems: analysis

In this section, we present some randomized algorithms for Problems 1–4. These algorithms are based on the Chernoff bound, and are in contrast with those in Section 4 which make use of the more advanced VC theory.

Consider first the robust stability problem. Recall that the problem is as follows: Given an element $\mathbf{y} \in Y$, determine whether or not *every* matrix $A \in \mathbf{A}_{\mathrm{y}}$ is stable. For this purpose, we begin by introducing a *probability measure* P_{y} on $\mathbf{A}_{\mathrm{y}}$; the significance of this probability measure becomes clear later. Generate i.i.d. matrices $A_1, \ldots, A_m \in \mathbf{A}_{\mathrm{y}}$, which are distributed according to P_{y}. Test each of these matrices for stability. If any one matrix is unstable, then declare that the answer to the robust stability problem is no; if every matrix is stable, declare that the answer to the robust stability problem is yes.

The above is an example of a *randomized algorithm*, because the outcome of the algorithm depends on the randomly generated matrices $A_1, \ldots, A_m \in \mathbf{A}_{\mathrm{y}}$. Thus, if the algorithm is repeated several times, there is no guarantee that the outcome would be the same each time. Using the one-sided Chernoff bound, it is possible to analyze the error probability of this randomized algorithm. There are two types of errors that need to be analyzed, namely: the false positive and the false negative. A false positive occurs when the correct answer to the decision problem is no, but the algorithm declares that the answer is yes; a false negative is just the opposite.

It is clear that the above randomized algorithm will *never* declare a false negative. Suppose that every matrix in $\mathbf{A}_{\mathrm{y}}$ is indeed stable; then in particular every randomly generated matrix A_i will also be stable. Hence, the algorithm will always declare a yes answer if that is indeed the correct answer. To analyze the probability of a false positive, define

$$\mathscr{S}_{\mathrm{y}} := \{A \in \mathbf{A}_{\mathrm{y}}\colon A \text{ is stable}\}.$$

Note that if every one of the randomly generated matrices A_i is stable, then the *empirical probability* $\hat{P}(\mathscr{S}_{\mathrm{y}})$ is equal to one. Now, it is easy to show that, if the empirical probability of a set is equal to one, then it can be said with confidence

$$q \leqslant (1-\varepsilon)^m$$

that its *true* probability is at least equal to $1-\varepsilon$. To see this, apply Lemma 11.1, p. 357 of Vidyasagar (1997a,b) with $X = \mathbf{A}_{\mathrm{y}}$, $P = P_{\mathrm{y}}$, and let the function f be the indicator function of the complement of $\mathscr{S}_{\mathrm{y}}$; the details are easy and are left to the reader. Hence, in order to bring q below some prespecified confidence threshold δ, it suffices to ensure that

$$(1-\varepsilon)^m \leqslant \delta.$$

To apply this inequality, suppose that one is given specific values for ε and δ, and wishes to determine how many samples m are sufficient to be able to say with confidence $1-\delta$ that $P_{\mathrm{y}}(\mathscr{S}_{\mathrm{y}}) \geqslant 1-\varepsilon$. The above inequality shows that it is enough to draw at least

$$m \geqslant \frac{\ln(1/\delta)}{\ln(1/(1-\varepsilon))}$$

stable matrices.

Now, let us consider a slightly different version of the robust stability problem. Suppose a number $\gamma > 0$ is specified, and it is desired to know whether or not $P_{\mathrm{y}}(\mathscr{S}_{\mathrm{y}}) \geqslant 1-\gamma$. This is also a decision problem for each γ, in the sense that the answer is either yes or no. However, it is not known to which complexity class this problem belongs. Now a randomized algorithm is presented for answering this decision problem.

Select an accuracy parameter $\varepsilon < \gamma$, and generate i.i.d. samples $A_1, \dots, A_m \in A_{\mathrm{y}}$ at random, distributed according to P_{y}. Then proceed as follows:

1. If $\hat{P}(\mathscr{S}_{\mathrm{y}}) > 1 - \gamma + \varepsilon$, declare that the answer to the decision problem is yes.
2. If $\hat{P}(\mathscr{S}_{\mathrm{y}}) < 1 - \gamma - \varepsilon$, declare that the answer to the decision problem is no.
3. If $\hat{P}(\mathscr{S}_{\mathrm{y}}) \in [1 - \gamma - \varepsilon, 1 - \gamma + \varepsilon]$, draw more samples.

The error probability of this algorithm can be analyzed using the one-sided Chernoff bound. If $P_{\mathrm{y}}(\mathscr{S}_{\mathrm{y}}) \leqslant 1 - \gamma$, then the probability that $\hat{P}(\mathscr{S}_{\mathrm{y}}) < 1 - \gamma + \varepsilon$ is at most $\mathrm{e}^{-2m\varepsilon^2}$. Similarly, if $P_{\mathrm{y}}(\mathscr{S}_{\mathrm{y}}) \geqslant 1 - \gamma$, then the probability that $\hat{P}(\mathscr{S}_{\mathrm{y}}) < 1 - \gamma - \varepsilon$ is at most $\mathrm{e}^{-2m\varepsilon^2}$. As these are mutually exclusive events, the probability that either Step 1 or Step 2 will generate an incorrect answer is at most equal to $\mathrm{e}^{-2m\varepsilon^2}$. Now what about Step 3? As $m \to \infty$, $\hat{P}(\mathscr{S}_{\mathrm{y}})$ converges almost surely to $P_{\mathrm{y}}(\mathscr{S}_{\mathrm{y}})$. Hence, unless $P_{\mathrm{y}}(\mathscr{S}_{\mathrm{y}})$ is *exactly* equal to $1 - \gamma$, which is a zero probability event if γ is chosen at random according to some nonatomic measure, it follows that Step 3 will eventually not get invoked for sufficiently large values of m.

A related issue is the amount of computational required to generate these i.i.d. matrices. Note that the interval matrix $\mathbf{A}_{\mathrm{y}}$ is just an n^2-fold Cartesian product of intervals of the form $[\alpha_{ij}, \beta_{ij}]$. Thus, if the measure P_{y} is also an n^2-fold product measure, then the problem naturally decomposes into a much simpler problem of generating n^2 real numbers, distributed in the intervals $[\alpha_{ij}, \beta_{ij}]$. For other choices of P_{y}, the generation of the i.i.d. sample matrices could itself be a difficult task; see Calafiore, Dabbene, and Tempo (1999) for further discussion on this topic.

Using entirely similar reasoning, it is possible to develop efficient algorithms for the problems of robust positive semidefiniteness and robust norm boundedness. In the case of the robust nonsingularity problem, there is an additional feature. Note that the determinant of A is a polynomial of degree n in the elements of the $n \times n$ matrix A. Thus the relation $\det A = 0$ defines a polynomial variety in the set of all matrices. Thus the set of singular matrices *always* has measure zero if P_{y} is the uniform measure, for example. In order for this question to be meaningful, one should instead choose a "tolerance" γ, and study the set of interval matrices

$$\mathscr{N}\mathscr{S}_{\mathrm{y},\gamma} := \{A \in \mathbf{A}_{\mathrm{y}}\colon |\det(A)| \geqslant \gamma\}.$$

Now the question of robust nonsingularity can be modified as follows: Given the element $\mathbf{y}$ and the constant γ, is it true that the magnitude of the determinant of *every matrix* $A \in \mathbf{A}_{\mathrm{y}}$ is at least equal to γ? A randomized algorithm can be developed for this problem along by now familiar lines.

Thus, in summary, it has been shown in this section that, for each of the NP-hard problems Nos. 1–4, it is possible to put forward efficient probabilistic algorithms that never give a false negative, and whose probability of giving a false positive can be analyzed. The chance that the randomized algorithm might fail occasionally is the price we pay for getting an algorithm that runs in polynomial-time.

4. Randomized algorithms for matrix problems: synthesis

In the previous section, we have seen that for several NP-hard *analysis* problems in matrix theory, it is possible to derive efficient randomized algorithms based on nothing more than the simple Chernoff bound. In this section we consider *synthesis* problems and present a randomized algorithm for the more difficult problem of simultaneous stabilization using constant state feedback.

Actually, this problem is a special case of a very general class of controller synthesis problems, including simultaneous stabilization (not necessarily using constant state gain feedback), H_∞- and H_2-optimal control, and so on.

Let us begin by modifying the problem formulation so as to make it amenable to the randomized approach. Recall the problem at hand: There are sets $\mathscr{A}, \mathscr{B}, \mathscr{C}$ and it is desired to know whether or not there exists a matrix K in some interval matrix set such that $A + BKC$ is stable for each triple (A, B, C) belonging to $\mathscr{A} \times \mathscr{B} \times \mathscr{C}$. In the interests of brevity, let us denote the set $\mathscr{A} \times \mathscr{B} \times \mathscr{C}$ by $\mathscr{F}$. Suppose P is a given probability measure on the set $\mathscr{F}$ that reflects our prior belief on how the system triples (A, B, C) are distributed in nature. Also, to avoid confusion, let us denote the interval matrix to which the state gain matrix K must belong by the symbol $\mathscr{K}_{\mathrm{y}}$ (as opposed to $\mathbf{A}_{\mathrm{y}}$ as in earlier sections). In other words, let

$$\mathscr{K}_{\mathrm{y}} := \{K \in \mathfrak{R}^{m \times p}\colon \alpha_{ij} \leqslant k_{ij} \leqslant \beta_{ij},\ \forall i,j\}.$$

In its "pure" form the question becomes: Does there exist a matrix $K \in \mathscr{K}_{\mathrm{y}}$ such that $A + BKC$ is stable for all $(A, B, C) \in \mathscr{F}$? As in the previous section, let us modify this question slightly. Let ε be a given accuracy parameter. For each matrix $K \in \mathscr{K}_{\mathrm{y}}$, define

$$\mathscr{S}(K) := \{(A, B, C) \in \mathscr{F}\colon A + BKC \text{ is stable}\}.$$

Now let us ask: Does there exist a matrix $K \in \mathscr{K}_{\mathrm{y}}$ such that $P(\mathscr{S}(K)) \geqslant 1 - \varepsilon$? In other words, a gain matrix K is considered to be "nearly simultaneously stabilizing" if the matrix $A + BKC$ is stable for all triples except possibly those belonging to a set of volume no larger than ε. To put it another way, if a triple (A, B, C) is chosen at random according to P, and if the gain matrix K stabilizes this triple with probability at least $1 - \varepsilon$, then we are satisfied. This modification is similar to asking in the preceding section whether the volume of the set of stable matrices is at least $1 - \varepsilon$. One way to approach the problem is to maximize the quantity $P(\mathscr{S}(K))$ with respect to K and see if the maximum is at least $1-\varepsilon$.

Now suppose we have a "candidate" gain matrix K. Determining whether or not $P(\mathscr{S}(K)) \geqslant 1-\varepsilon$ for *this particular* matrix K is not an easy task, since in general it is not easy to compute the volume $P(\mathscr{S}(K))$ *exactly*. unless the families $\mathscr{A}$,$\mathscr{B}$,$\mathscr{C}$ are all finite. Instead, we can try to *approximate* this volume by empirical means. Generate random triples (A_i, B_i, C_i), $i=1,\ldots,m$ distributed according to P. Given a gain matrix K, check each triple $A_i + B_i K C_i$ for stability, and then compute what fraction of the m triples are stabilized by K. This leads to an empirical estimate $\hat{P}(\mathscr{S}(K))$. Now we already know from the Chernoff bounds that, *for each fixed gain matrix* K, the quantity $\hat{P}(\mathscr{S}(K))$ converges to the true value $P(\mathscr{S}(K))$ as the number of samples m approaches infinity. The difficulty is that K is now itself a variable of design. Consequently, if we compute the quantity $\hat{P}(\mathscr{S}(K))$ for different matrices K, we cannot be sure that *all* of these empirical estimates are *uniformly close* to the corresponding true values $P(\mathscr{S}(K))$. However, suppose it so happened that the *collection of sets* $\{\mathscr{S}(K),\ K \in \mathscr{K}_y\}$ had the UCEP property, by virtue of having finite VC-dimension. Then we would be able to state that *each* of the empirical estimates $\hat{P}(\mathscr{S}(K))$ is uniformly close to the corresponding correct value $P(\mathscr{S}(K))$. Hence if we were to go ahead and maximize the quantity $\hat{P}(\mathscr{S}(K))$ with respect to K, we would be fairly close to the maximum of $P(\mathscr{S}(K))$ as well. More precisely, let us choose the integer m sufficiently large that $|\hat{P}(\mathscr{S}(K)) - P(\mathscr{S}(K))| \leqslant \varepsilon/2$ for *every* $K \in \mathscr{K}_y$, with some confidence $1-\delta$; this can be done using Theorem 1 once we have an upper bound for the VC-dimension of the family $\{\mathscr{S}(K),\ K \in \mathscr{K}_y\}$. Then, if we maximize $\hat{P}(\mathscr{S}(K))$ with respect to K, and if the maximum turns out to be at least $1-\varepsilon/2$, then it can be said with confidence $1-\delta$ that there exists a gain matrix K such that $P(\mathscr{S}(K)) \geqslant 1-\varepsilon$. On the other hand, if this maximum turns out to be at most $1-3\varepsilon/2$, then it can be said with confidence $1-\delta$ that there does not exist a gain matrix K such that $P(\mathscr{S}(K)) \geqslant 1-\varepsilon$. Thus, it follows that it is worthwhile to know under what conditions the collection of sets $\{\mathscr{S}(K),\ K \in \mathscr{K}_y\}$ has finite VC-dimension, and to derive *explicit* upper bounds for this VC-dimension.

Note that maximizing even the modified objective function $\hat{P}(\mathscr{S}(K))$ with respect to K is not easy in general, since there might not be a closed-form expression for $\hat{P}(\mathscr{S}(K))$, without which it is not possible to use efficient gradient-based optimization methods. It turns out that even this maximization can be carried out using randomized methods, using an approach advocated in Khargonekar and Tikku (1996), Tempo et al. (1997). These papers assume that the underlying probability distribution function is continuous. This assumption is removed in Vidyasagar (1997a), Lemma 11.1, p. 357. See Vidyasagar (1997b) for full details.

5. Bounds on the VC-dimension for specific families

In order for the randomized algorithm suggested in the preceding section to work, it is necessary that the collection of sets $\{\mathscr{S}(K),\ K \in \mathscr{K}_y\}$ have the UCEP property. A sufficient (and necessary, if we insist that P could be *any* probability measure) condition for this collection to have the UCEP property is that the VC-dimension of $\{\mathscr{S}(K),\ K \in \mathscr{K}_y\}$ be finite. Using Theorem 2, we now derive an upper bound for this VC-dimension. It turns out that Theorem 2 is a very versatile tool that can be used to derive such upper bounds for a wide variety of controller synthesis problems, such as simultaneous stabilization (not necessarily using constant gain state feedback), H_∞- and H_2-optimal control, and so on. See Vidyasagar (1997b) for details.

Let us begin by recalling the *Hurwitz stability test* for the stability of a polynomial, which is equivalent to the Routh test that is more familiar to control engineers (and which is almost universally mislabelled as the "Routh–Hurwitz test"). In the case where the coefficients of the polynomial are constants, there is not much difference between the Hurwitz test and the Routh test; however, in the case where the coefficients of the polynomial are functions of some parameter (as is the case here), it turns out that the Hurwitz test leads to more economical estimates. The Hurwitz test as described below can be found in many places, for example, Gantmacher (1959), Chapter XV, pp. 190ff.

Lemma 1. *Let*

$$a(s) := a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n = \sum_{i=0}^{n} a_i s^{n-i}.$$

Suppose without loss of generality that $a_0 > 0$. *Define the Hurwitz determinants*

$$H_i := \begin{vmatrix} a_1 & a_3 & a_5 & \ldots & a_{2i-1} \\ a_0 & a_2 & a_4 & \ldots & a_{2i-2} \\ 0 & a_1 & a_3 & \ldots & a_{2i-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \ldots & a_i \end{vmatrix}, \quad i = 1,\ldots,n,$$

where a_i *is taken as zero if* $i > n$. *Then* $a(\cdot)$ *is a stable polynomial* (i.e., *all of its zeros have negative real parts*) *if and only*

$$H_i > 0, \quad i = 1,\ldots,n.$$

Now suppose that each coefficient a_i is in fact a function of some parameter p, and that $a_i(p)$ is a polynomial in p of degree no larger than k. Then it is easy to see that H_i is also a polynomial in p, and that the degree of $H_i(p)$ is no larger than ki.

In contrast, if one were to form the *Routh array* of the polynomial $a(s)$, then the polynomial is stable if and only if the elements in the first column of the Routh array are all of the same sign. Now it is well-known (see e.g., Gantmacher, 1959) that the first element in the ith row of the Routh array (call it R_i) is the product of the first i Hurwitz determinants. Thus, in the case where the coefficients of the polynomial $a(s)$ are themselves polynomials in an auxiliary parameter p of degree no larger than k, the ith Hurwitz determinant has degree no larger than ki with respect to p, whereas the Routh array element R_i has degree no larger than $ki(i+1)/2$ with respect to p.

Next, consider the case of matrices $F(p) = [f_{ij}(p)] \in \Re^{n\times n}$, where each $f_{ij}(p)$ is a polynomial in p of degree no larger than k. Let

$$\phi(s,p) := \det(sI - F(p)) = s^n + \sum_{i=1}^{n} \phi_i(p)s^{n-i}$$

denote the characteristic polynomial of $F(p)$. Then each coefficient $\phi_i(p)$ is a sum of $i \times i$ minors of $F(p)$, and is thus a polynomial in p of degree no larger than ki. The Hurwitz determinant

$$H_i := \begin{vmatrix} \phi_1 & \phi_3 & \phi_5 & \dots & \phi_{2i-1} \\ \phi_0 & \phi_2 & \phi_4 & \dots & \phi_{2i-2} \\ 0 & \phi_1 & \phi_3 & \dots & \phi_{2i-3} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \phi_i \end{vmatrix}$$

is a polynomial in p of degree no larger than $i(i+1)k/2$, as can easily be shown by induction on i.

Now we state the main result. In the interests of simplicity it is assumed that $p = m = n$; the modifications in case $m < n$ and/or $p < n$ are easy and left to the reader.

Theorem 3 (Simultaneous stabilization using constant output feedback). *Suppose*

$$\mathbf{y} := (\alpha_{ij},\beta_{ij}), \quad 1 \leqslant i,\ j \leqslant n,$$

where α_{ij},β_{ij} *are rational numbers for all i,j. Define*

$$\mathscr{S}(K) := \{(A,B,C) \in \mathscr{A} \times \mathscr{B} \times C : A + BKC \textit{ is stable}\}.$$

Define $\mathscr{K}_{\mathbf{y}}$ *to be the corresponding interval matrix, that is: Let* $\mathscr{S}_{\mathbf{y}}$ *be a shorthand for the collection of sets* $\{\mathscr{S}(K),\ K \in \mathscr{K}_{\mathbf{y}}\}$. *Then*

$$\text{VC-}\mathit{dim}(\mathscr{S}_{\mathbf{y}}) \leqslant 2n^2 \log_2[2en^2(n+1)]. \tag{6}$$

Proof. The theorem is an immediate consequence of Theorem 2 and Lemma 1. Note that, for each $\mathbf{y}$, the interval matrix $\mathscr{K}_{\mathbf{y}}$ is a subset of $\Re^{n\times n}$. Hence, if we can show that the VC-dimension of the collection of sets $\mathscr{S} := \{\mathscr{S}(K),\ K \in \Re^{n\times n}\}$ is bounded as above, then the theorem is proved, since each $\mathscr{S}_{\mathbf{y}}$ is a subcollection of $\mathscr{S}$.

Let $F := A + BKC$, and let $H_i(F)$ denote the ith Hurwitz determinant of its characteristic polynomial. The triplet (A,B,C) belongs to the set $\mathscr{S}(K)$ for a fixed K if and only if $A + BKC$ is stable. We now express this condition as a set of polynomial inequalities and then invoke Theorem 2. The stability of the matrix $F := A + BKC$ is equivalent to $H_i(F) > 0$ for $i = 1, \dots, n$, where $H_i(F)$ denotes the ith Hurwitz determinant of F. Note that

$$f_{ij} = a_{ij} + \sum_{l=1}^{n} \sum_{m=1}^{n} b_{il}k_{lm}c_{mj}.$$

Hence, each f_{ij} is a polynomial of degree one or less in the elements of the matrix K. As a consequence, by the argument preceding the statement of the theorem, each Hurwitz determinant $H_i(F)$ is a polynomial of degree no larger than $i(i+1)/2$ in the elements k_{lm}. Since $i \leqslant n$, the maximum degree of these Hurwitz determinants is $n(n+1)/2$. Now apply Theorem 2 with the elements of A, B, C playing the role of the parameter vector x and the elements of K playing the role of the parameter vector z. Then

- l = the dimension of the parameter space = n^2,
- t = the number of inequalities = n, and
- r = the maximum degree of each inequality = $n(n+1)/2$.

The desired bound now follows from (5). □

In the spirit of the above theorem, we now derive upper bounds for the VC-dimension of various types of sets that arise in connection with interval matrices. Note that, as of now, there is no direct application for these bounds. Nevertheless, it is worthwhile to derive these bounds, in case some applications for them can be found in future. The proofs are omitted, as they follow along the same lines as that of Theorem 3.

Theorem 4 (Stability of interval matrices). *Let* $X = \Re^{n\times n}$, $Y = \Re^{2n^2}$, *where n is a given integer. Suppose*

$$\mathbf{y} := (\alpha_{ij}, \beta_{ij}), \quad 1 \leqslant i,\ j \leqslant n,$$

where $\alpha_{ij} \leqslant \beta_{ij}$ *for all i,j. Define*

$$\mathscr{S}_{\mathbf{y}}:\{A \in \mathscr{A}_{\mathbf{y}} \textit{ and A is stable}\}, \quad \mathscr{S} := \{S_{\mathbf{y}}: \mathbf{y} \in Y\}.$$

Let γ *be a given constant, and define*

$$\mathscr{N}_{\mathbf{y}} := \{A \in \mathbf{A}_{\mathbf{y}}: |\det(A)| > \gamma\}, \quad \mathscr{N} := \{\mathscr{N}_{\mathbf{y}}: \mathbf{y} \in Y\}.$$

Let $\mathbf{A}_{s,\mathbf{y}}$ *be as above, and define*

$$\mathscr{PSD}_{\mathbf{y}} := \{A \in \mathbf{A}_{s,\mathbf{y}}: A \textit{ is positive semidefinite}\},$$

$$\mathscr{PSD} := \{\mathscr{PSD}_{\mathbf{y}}: \mathbf{y} \in Y_s\},$$

$$\mathscr{PD}_{\mathbf{y}} := \{A \in \mathbf{A}_{s,\mathbf{y}}: A \textit{ is positive definite}\},$$

$$\mathscr{PD} := \{\mathscr{PD}_{\mathbf{y}}: \mathbf{y} \in Y_s\}.$$

With these definitions, *we have*

$$VC\text{-}dim(\mathscr{S}) \leqslant 4n^2 \lg[4\mathrm{e}(2n^2 + n)].$$

$$VC\text{-}dim(\mathscr{N}) \leqslant 4n^2 \lg[4\mathrm{e}(2n^2 + 1)].$$

$$VC\text{-}dim(\mathscr{P}\mathscr{S}\mathscr{D}) \leqslant 4n^2 \lg[4\mathrm{e}(2n^2 + 2^n)],$$

$$VC\text{-}dim(\mathscr{P}\mathscr{D}) \leqslant 4n^2 \lg[4\mathrm{e}(2n^2 + n)].$$

6. Conclusions

In this paper, it has been shown that by using standard bounds on empirical probabilities as well as recent results in VC-dimension theory, it is possible to generate polynomial-time *randomized* algorithms for *modified versions of* several NP-hard problems in matrix theory. Actually, the approach described here is widely applicable. By applying these ideas, future researchers can undoubtedly prove many more such results.

The problems studied here have one noteworthy feature in contrast with problems in *computational* learning theory, as studied in Anthony and Biggs (1992), and Kearns and Vazirani (1994). In the type of problems studied in the computer science literature, for each fixed integer n, the number of problem instances is *finite*, for example, the number of 3-conjuctive normal form Boolean formulas in n variables. In such a case, it follows automatically that the VC-dimension of the corresponding problem class is finite for each integer n, and an issue that merits attention is the *rate* at which the VC-dimension grows with respect to n. In contrast, the type of matrix theory problems studied here have the feature that, *even for a fixed finite integer n*, the set of the problem instances is infinite; consequently, proving that the VC-dimension of the set of problem instances is finite is by itself a nontrivial achievement. Moreover, a perusal of Theorem 4 shows that the VC-dimension grows at most polynomially with respect to the “size” parameter n.

References

Anderson, B. D. O., Bose, N. K., & Jury, E. I. (1975). Output feedback stabilization and related problems—Solution via decision methods. *IEEE Transactions on Automatic Control*, *AC-20*, 53–66.

Anthony, M., & Biggs, N. (1992). *Computational learning theory*. Cambridge, UK: Cambridge University Press.

Blondel, V. D., & Tsitsiklis, J. N. (1997). NP-hardness of some linear control design problems. *SIAM Journal on Control and Optimization*, *35*, 2118–2127.

Blondel, V. D., & Tsitsiklis, J. N. (2000). A survey of computational complexity results in systems and control. *Automatica*, *36*(9), 1249–1274.

Calafiore, G., Dabbene, F., & Tempo, R. (1999). Randomized algorithms and probabilistic robustness with real and complex structured uncertainty, preprint.

Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, *23*, 493–507.

Gantmacher, F. R. (1959). *Matrix theory. Vol. II*. New York: Chelsea.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. New York: Freeman.

Karpinski, M., & Macintyre, A. J. (1995). Polynomial bounds for VC dimension of sigmoidal neural networks. *Proceedings of the 27th ACM Symposium Theories of Computing* (pp. 200–208).

Karpinski, M., & Macintyre, A. J. (1997). Polynomial bounds for VC dimension of sigmoidal and general Pfaffian neural networks. *Journal of Computer System and Sciences*, *54*, 169–176.

Kearns, M., & Vazirani, U. (1994). *Introduction to computational learning theory*. Cambridge: MIT Press.

Khargonekar, P. P., & Tikku, A. (1996). Randomized algorithms for robust control analysis have polynomial complexity. *Proceeding Conference on Decision and Control*.

Marrison, C., & Stengel, R. (1994). The use of random search and genetic algorithms to optimize stochastic robustness functions. *Proceedings of the American Control Conference* Baltimore, MD (pp. 1484–1489).

Nemirovskii, A. (1993). Several NP-hard problems arising in robust stability analysis. *Mathematics of Control, Signals, and Systems*, *6*(2), 99–105.

Papadimitrou, C. (1994). *Computational complexity*. Reading, MA, USA: Addison-Wesley.

Poljak, S., & Rohn, J. (1993). Checking robust nonsingularity is NP-hard. *Mathematics of Control, Signals, and Systems*, *6*(1), 1–9.

Ray, L. R., & Stengel, R. F. (1991). Stochastic robustness of linear time-invariant control systems. *IEEE Transactions on Automatic Control*, *36*, 82–87.

Tarski, A. (1951). *A decision method for elementary algebra and geometry*. Berkeley, USA: University of California Press.

Tempo, R. E., Bai, W., & Dabbene, F. (1997). Probabilistic robustness analysis: Explicit bounds for the minimum number of sampling points. *Systems & Control Letters*, *30*, 237–242.

Vapnik, V. N. (1982). *Estimation of dependences based on empirical data*. New York: Springer-Verlag.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.

Vapnik, V. N., & Chervonenkis, A. Ya. (1971). On the uniform convergence of relative frequencies to their probabilities. *Theory of Probability and its Applications*, *16*(2), 264–280.

Vidyasagar, M. (1997a). *A theory of learning and generalization: with applications to neural networks and control systems*. London: Springer-Verlag.

Vidyasagar, M. (1997b). Statistical learning theory and its applications to randomized algorithms for robust controller synthesis. (Semi-Plenary Talk), European Control Conference, Brussels, Belgium.

Vincent D. Blondel received a Bachelor Degree in Engineering in Applied Mathematics from the University of Louvain (Louvain-la-Neuve, Belgium) in 1988, a Master Degree in Pure Mathematics from Imperial College (London, UK) and a Ph.D. in 1992 from the University of Louvain. In 1993, he was a visiting scientist at Oxford University. During the academic year 1993–1994, he was the Göran Gustafsson Fellow at the Royal Institute of Technology (Stockholm, Sweden). In 1993–1994 he was a research fellow at the French National Research Center in Control and Computer Science (INRIA Paris). From 1994 to 1999 he was an associate professor at the Institute of Mathematics of the Université de Liège in Belgium. Since October 1999 he is with the University of Louvain where he is currently professor in the Department of Applied Mathematics. He has been a visitor with the Australian National University (1991), the Massachusetts Institute of Technology (every year since 1994) and the University of California in Berkeley (1998). He has also been an invited professor at the Ecole Normale Supérieure in Lyon (1998) and at the University of Paris VII

(1999 and 2000). Dr Blondel's major current research interests lie in several area of mathematical control theory and theoretical computer science. He has been a recipient of a Grant from the Trustees of the Mathematics Institute of Oxford University, the Prize Agathon De Potter of the Belgian Royal Academy of Science and the Prize Paul Dubois of the Montefiore Institute. He is the coordinator of a NATO collaborative grant with the Massachusetts Institute of Technology (USA) and the Russian Academy of Science, and is a partner in a European TMR network on discrete event systems. He is an associate editor of Systems and Control Letters (Elsevier), of the European Journal of Control (Springer) and of the Journal on Mathematics of Control, Signals, and Systems (Springer).

Mathukumalli Vidyasagar was born in Guntur, Andhra Pradesh, India on 29 September 1947. He received the B.S., M.S., and Ph.D. degrees, all in electrical engineering, from the University of Wisconsin, in 1965, 1967, and 1969, respectively. He has taught at Marquette University, U.S.A. (1969-70), Concordia University, Canada (1970-80), and the University of Waterloo, Canada (1980-89). In 1989 he returned to India as the Director of the Centre for Artificial Intelligence and Robotics, (under the Defence Research and Development Organisation) in Bangalore. In 2000 he took up his current assignment as Executive Vice President (Advanced Technology) in Tata Consultancy Services, which is India's largest IT firm. At present he is based in the city of Hyderabad. In his current position, his responsibilities are to create an Advanced Technology Centre (ATC) within TCS, to develop futuristic technologies of relevance to the IT industry. At the present time, the scope of activities of the ATC includes PKI (Public Key Infrastructure), security in e- and m-commerce, advanced cryptography including elliptic curve cryptography, and neural networks.

In the past, he has held visiting positions at several universities including M.I.T., California (Berkeley), California (Los Angeles), C.N.R.S. Toulouse, France, Indian Institute of Science, University of Minnesota, and Tokyo Institute of Technology. He is the author or co-author of seven books and more than one hundred and twenty papers in archival journals. He has received several honours in recognition of his research activities, including the Distinguished Service Citation from his Alma Mater (The University of Wisconsin). In addition, he is a Fellow of Institute of Electrical and Electronics Engineers (IEEE), the Indian Academy of Sciences, the Indian National Science Academy, the Indian National Academy of Engineering, and the Third World Academy of Sciences. His current research interests are control theory, machine learning, and cryptography.