

## COMPARATIVE ANALYSIS OF KOLMOGOROV ANN AND PROCESS CHARACTERISTIC INPUT-OUTPUT MODES

Georgi M. DIMIROVSKI<sup>1</sup>

Yuanwei JING<sup>2</sup>

<sup>1</sup> Department of Computer Eng., Dogus University, 34722 – Kadikoy / Istanbul, Turkey, and  
Institute of ASE-FEE, SS Cyril and Methodius University, 1000 – Skopje, R. of Macedonia

<sup>2</sup>School of Information Sci. & Eng., Northeastern Univ., 110006 – Shenyang, P.R. of China

<sup>1</sup>E-mail: [gdimirovski@dogus.edu.tr](mailto:gdimirovski@dogus.edu.tr)

<sup>2</sup>E-mail: [ywjing@etang.com](mailto:ywjing@etang.com)

### ABSTRACT

*In the past decades, representation models of dynamical processes have been developed via both traditional math-analytical and less traditional computational-intelligence approaches. This challenge to system sciences goes on because essentially involves the mathematical approximation theory. A comparison study based on cybernetic input-output view in the time domain on complex dynamical processes has been carried out. An analytical decomposition representation of complex multi-input-multi-output thermal processes is set relative to the neural-network approximation representations, and shown that theoretical background of both emanates from Kolmogorov's theorem. The findings provided a new insight as well as highlighted the efficiency and robustness of fairly simple industrial digital controls, designed and implemented in the past, inherited from input-output decomposition model approximation employed.*

**Keywords:** *Approximation models, characteristic input-output modes, complex systems, infinite matrices, neural networks.*

### I. INTRODUCTION

It may seem a paradox, but all the exact science is dominated by the idea of approximation – Bertrand Russell

This outstanding statement by Russell, one of the two founders of modern mathematical science [50], best supports all endeavors of systems and control science in identification of approximate model representations of dynamical processes, no matter whether the word is about input-output black-box [47], [52] or structural state-space views [38], [53], and whether obtained by math-analytical [37], [38] or computational intelligence [49], [54], [55] methods. This

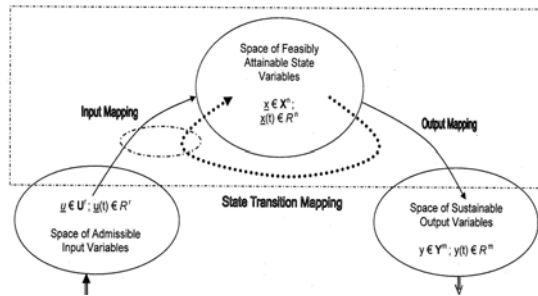
statement may well be readily inferred from Figure 1 on the fundamental concept of a dynamical system in engineering terms [9], [10]. For, indeed the word is about *approximation models* for time-domain functions, functionals or sequences that are theoretically consistent and empirically correct and valid.

By and large, all approaches to approximate system representations involve some kind of decomposition either from input-output or input-state-output points of view [10], [24]. The latter, of course, is associated intrinsically with the math-analytical fundamental approach [38], [53], and not with the one of the

*Received Date : 30.10.2002*

*Accepted Date: 13.12.2002*

computational-intelligence [49], [55]. the input-output (I/O) view the foundations of which were laid down by the founders of Cybernetics [47], [52]., however, generally applies to both of them to the same extent [17]. In addition, in both fundamental approaches the ideas of system decomposition and expansion have been exploited to the full, and both rely essentially on real-world recorded time series possessing all possible peculiarities [37], [41].



**Fig. 1.** The concept of a dynamical system in general, redefined in systems engineering terms according to essentially non-separable interplay of fundamental natural quantities – energy, information, and matter – in dynamical processes.

In the present study, the focus is put on a comparative analysis [17] of certain approximate model representation results via math-analytical and neural-network techniques for the class of complex multi-input-multi-output (MIMO) processes having real-world natural steady-states such as in thermal plants [13]-[16]. Note, in addition, that certain on-site experiments were done before on real-world, high-power, industrial furnaces in addition to a number of experimental simulations [14]-[16], [45]. Here it is impossible, of course, to cite properly all the relevant references in the literature. Nonetheless, it is pointed out that some of them are found in the reference list included, and in the references therein. In particular, works based on the I/O systems view in the math-analytical (MAN) approach, [3], [5], [10]-[12], [21]-[23], [29], and in the computational intelligence approach employing artificial neural networks (ANN), [1], [6], [19], [20], [27], [30], [31]-[36], [40], [43], [44], should be observed.

Next Section II gives a review discussion of some analytical results on I/O dynamical systems

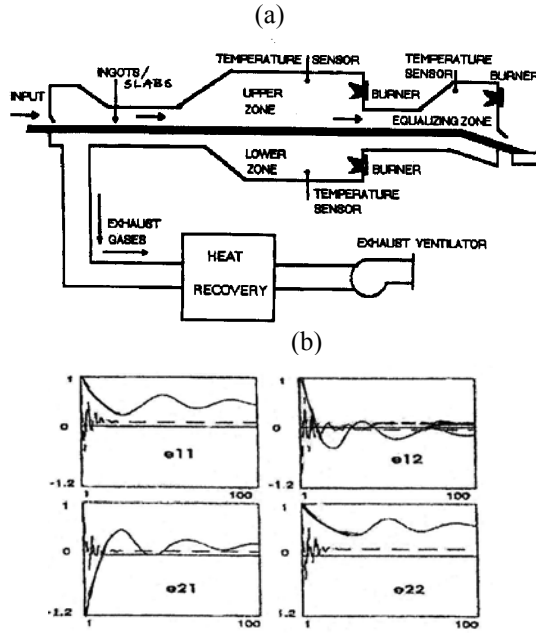
representations via characteristic I/O modes decomposition [10], [11]-[13], [21]-[23]. Similarly, but in the realm of neural-network computing structures [2], [4], [26], [48], [51], some of the most outstanding representation results in the literature, are recalled as appropriate along with a relevant review discussion in Section III, where these are cited and referenced. In Section IV, the comparison findings are summarized. Conclusion, an appendix (on relevant tools from mathematical approximation theory) and references are given thereafter.

## 2. ON MATH-ANALYTICAL BLACK-BOX IDENTIFIED SYSTEM MODELS

Pragmatically, the main goal in real-world system engineering applications is always to drive the plant so as to achieve the desired state(s) out of the all attainable ones controlling by means of the specified task-oriented and set-point controls, i.e. by integrated control and supervision [1], [16]. For this purpose, firstly, a set of families of models is indispensable, at least for operating points at low-, medium-, normal- and maximum-operating loads. And, secondly, ultimately both control and supervision algorithms are implemented in terms of time-domain, discrete-time approximations of the theoretical designs as well as the practical stability of the overall system in all operating regimes must be ensured [14], [23], [45].

In this paper, the proper theoretical treatment of the process model identification, which is found elsewhere (e.g., see [24], [37], [38]), is left aside and reference is made to the typical practice of operating common industrial plants that have finite steady state equilibrium, as the case study of 25 MW reheating furnace RZS (see the schematic in Figure 2) in Skopje Steelworks [14], [16], [23] is. In principle, any identification and signal processing methods may be used. Then, for a complex  $N_{inp} \times N_{out}$  plant ( $N_{inp} = N_i$  number of inputs;  $N_{out} = N_o$  number of outputs) in a given plant environment with specified operating conditions, systems engineering design starts with a designed identification experiment and a study of a family of steady-state, non-linear input/output

models representing “energy/material supply variables,  $m_j$ , to controlled measurable state variables,  $y_i$ ” within respective ranges of admissible inputs and sustainable outputs; moreover, within ranges where these are convex. In addition, a set of empirically (by either observation or experiment) obtained impulse-response representations are made available.



**Fig. 2.** Real-world 25 MW pusher furnace RZS illustrating this comparative study: the schematic (a) of and (b) the reduced errors via learning the process characteristic I/O modes for the identified kTSM model of main 2x2 sub-system ‘upper zone – lower zone’.

In here we make use of the case studies [14], [16], [45] of high-power, large, multi-zone, gas/oil-fired industrial furnaces (like the one in Fig. 2) where the main subset of measurable states are physical temperatures. The I/O approximation models that are operationally identified on the grounds of recorded non-linear time sequences [41] of input and output variables have the form

$$T_i = f_i(m_j, m_l = \text{const}, l \neq j) = y_i, \quad (1)$$

$$i = 1, \dots, N_o$$

at the operating point  $OpPo_\mu$ ,  $\mu = 1, 2, \dots, N_\mu$  and

$$[G(t_k)]_{N_o \times N_i} = [g(t_k)]_{N_o \times N_i \times N_{t_T}}, \quad t_0 \leq t_k \leq t_{t_T} \quad (2a)$$

or

$$[G(t_k)]_{N_o \times N_i} = [g(t_k)]_{N_o \times N_i \times N_{t_{SS}}}, \quad t_0 \leq t_k \leq t_{t_{SS}} \quad (2b)$$

where truncation and steady-state time instants are finite  $t_{t_T} < t_\infty$  and  $t_{t_{SS}} < t_\infty$ . These models are produced by using admissible inputs with regard to the respective magnitude ranges at  $OpPo_\mu$  operating point  $\mu$  (a number of them).

Ultimately, the latter two are in fact pseudo-impulse responses, discretized correctly with respect to the time, known as pseudo-weighting functions [3] or I/O weighting patterns [21], or as k-time sequence matrices [10], kTSM for short. In here, constant  $N_{t_T}$  indicates the three-dimensional matrices of instantaneous values span up to the truncation time  $t_T$ , and  $t_{SS}$  is the time when steady state is reached, while theoretically  $t_\infty \rightarrow +\infty$  [10]. In any case, these are approximated three-dimensional infinite Markov matrices [29] of plant process dynamics that have been conveniently truncated. Note for thermal processes, having natural steady-states, these still can contain information on the local nonlinear distortions [10]-[13]. Hence identified models in all forms, relevant to the approach in the time domain, may readily be available:

$$y(t_k) = [g(t_l)]_{N_o \times N_i \times N_{t_T}} * m(t_l), \quad t_0 \leq t_l \leq t_k < t_{t_{SS}} \quad (3)$$

with \* denoting convolution operation, and

$$[g(t_l)]_{N_o \times N_i \times N_{t_T}} \cong [g(t_k)]_{N_o \times N_i \times N_{t_\infty}} \quad (3a)$$

if and only if

$$\lim_{N \rightarrow N_{t_T}} \left\{ [g(t_l)]_{N_o \times N_i \times N} \right\}_0^{N_{t_T}} =$$

$$\lim_{N \rightarrow N_{t_T}} \left\{ [g(t_l)]_{N_o \times N_i \times N} \right\}_0^{N_{t_T} + 1} \text{ when}$$

$$N_{t_T} \rightarrow N_{t_{SS}}; \quad (3b)$$

$$y(q^{-1}) = [G(q^{-1})]_{N_o \times N_i} m(q^{-1}) + [\Psi]_{N_o \times N_o} e(t_k); \quad (4)$$

$$\begin{aligned} x(t_k) &= [A(T_s)]_{n \times n} x(t_k - 1) + [B(T_s)]_{n \times N_i} m(t_k - 1), \\ y(t_k) &= [C]_{N_o \times n} x(t_k) + [\Psi]_{N_o \times N_o} e(t_k); \end{aligned} \quad (5)$$

$$\begin{aligned} y(t_k) &= f(y(t_k - 1), \dots, y(t_k - N_y), \\ m(t_k - 1), \dots, \\ m(t_k - t_\lambda - N_m), e(t_k - 1), \dots, e(t_k - N_e)) + e(t_k) \end{aligned} \quad (6)$$

where  $f: R^{N_y + N_u} \rightarrow R^{N_o}$  and  $e(t_k)$  is the vector of a zero-mean disturbance term. Of course, by allowing correctly for limited bandwidth via proper time sampling and using anti-aliasing filters, this enables always a convenient use of digital computations. Hence, the following formal definitions of characteristic I/O modes and vectors in the time domain for MIMO dynamical processes result [10], [21].

*Definition II.1.* The scalar Toeplitz operator  $[\lambda_i(t_k)]$  is a *characteristic input-output mode or pattern (CPA)* of a MxM matrix plant convolution (pseudo-impulse response) operator  $[G(t_k)]_{N_o \times N_i} = [g(t_k)]_{N_o \times N_i \times N_{tT}} \hat{=} G$  for  $N_{inp} = N_{out} = M$ ,  $t_0 \leq t_k \leq t_\infty$ , if it is a root of the characteristic polynomial equation  $\det[\lambda_i(t_k)I - [G(t_k)]_{M \times M \times N_{iss}}]$ . The Mx1 vector operator  $w(t_k)$  is a *characteristic vector (CVE)* if it satisfies the equation

$$\{ \lambda_i(t_k)I - [G(t_k)]_{M \times M \times N} \} w(t_k) = 0. \quad (7)$$

Apparently, the above concepts of CPA and the CVE describe the *natural operational input-output dynamic behavior* of plant convolution operator  $G$ . In turn, plant convolution operator  $G$  can be obtained by means of the spectral decomposition

$$[G(t_k)]_{M \times M \times N} = \sum_{i=1}^M \lambda_i(t_k) w_i(t_k) v_i(t_k) \quad (8a)$$

for simplicity

$$G \hat{=} W \Lambda V. \quad (8b)$$

Here,  $\Lambda$  is a diagonal matrix comprising

$\lambda_i(t_k)$ , which represent the CPA's, and the columns of  $W = V^{-1} = [w_1 \ w_2 \ \dots \ w_M]$  are the CVE's. The validity of the approximation can be confirmed by the known Gershgorin eigenvalue theorem for diagonal dominance. A normalization of the CVE is carried out such that each diagonal element of  $W$  is the identity element (1, 0, 0, ...), and also an reordering procedure using the decomposition result in [13] is carried out such that to each  $\{\lambda_i\}$  corresponds a constituent matrix  $C_i$  with maximum norm on the diagonal element 'i' of approximated plant convolution operator  $G$ .

Recall now that for any compatible set of norms there exists a vector norm (of signal variables), denoted by  $|\cdot|$ , of such that  $|Gu| < \|G\| |u|$ .

Then  $\|G\|_1, \|G\|_2, \|G\|_\infty$  are system operator norms induced by signal norms  $|\cdot|_1, |\cdot|_2, |\cdot|_\infty$ , respectively. Operator elements of the kind of  $G$ , together with the zero  $O$  and the identity  $I$  operators, belong to a commutative Banach algebra in which the product corresponds to discrete convolution [8] and may thus be manipulated using block-diagram algebra. The relationship between defined sequences and the corresponding stable or unstable Z-transforms is detailed by Cheng and Desoer in [5], who showed that the discrete time case can be treated by more straightforward methods in contrast to the continuous-time distributed case, which normally is plagued by difficult fine points of analysis. In addition, for Lure-Postnikov class of non-linear systems that are time-discretized, a left-distributive algebra on extended Banach spaces of p-summable ( $p=1, 2, \infty$ ) sequences exists [10]. It is well-known that the elements of  $G$  do not reveal the useful information about the interaction behavior of the plant, while the spectral decomposition (7)-(8) does [3], [10], [21].

In principle, a CPA and associated CVE may be real and distinct, confluent or ones of a complex pair, however, for the well-posed gas furnaces they are real and distinct. When a CPA is real then its steady-state gain (SSG) matrix  $G_{ss}$  is also real, and thus a simple test is to check the

eigenvalues of  $G_{ss}$ . The CPA may be positive or negative, depending on the sign of its SSG; again, a simple test is to check the eigenvalues of  $G_{ss}$ . A  $G$  has a characteristic inverse response pattern if a CPA is positive but the initial transient is negative, and vice versa, which implies that an eigenvalue of the first non-singular Markov matrix  $G(t_k)$  shall be positive if the eigenvalue of  $G_{ss}$  is positive. The CPA type is defined in the usual way for scalar operators, i.e. if the system type is the same for all elements along a row of  $G$  then the CPA will take the same type. Internal process interaction is defined as the effect of one subsystem on the others, and this is reflected by the CVE precisely. If a CVE is aligned with its basis vector, i.e.  $w_i = I$ , then  $G$  is partially or triangular decoupled, and if all CVE are aligned with the basis vectors, i.e.  $W = I$ , then  $G$  is fully decoupled or non-interacting. If a CVE vary with time  $t_k$ , the interaction is dynamic, otherwise it is static implying these vectors are zero for all  $t_k > 0$ . Static characteristic vectors imply that the constituents are constant real matrices.

In real-world processes, the propagation information takes place along with flow and processing of energy and mass. Hence their behavior depend essentially on the magnitude of manipulated variables (which induces all sorts of nonlinear distortions that may be present in the plant), and hence the concept of energy and power balance within control loops has got a considerable impact [10], [17]. Moreover, from the energy balance point of view, the power contained in a kTSM operator of  $G$  can be well-defined via a decomposition theorem [13]. For the purpose of the present comparative analysis its updated version [14] is stated below.

*I/O Mode Decomposition Theorem II.1.* The ordered characteristic input-output modes of well-posed MIMO processes for  $M = N_{N_o} = N_{N_i}$  and  $N = N_{t_r}$ , represented by three-dimensional k-time sequence matrices, have their natural decomposition in terms of characteristic signal power distribution

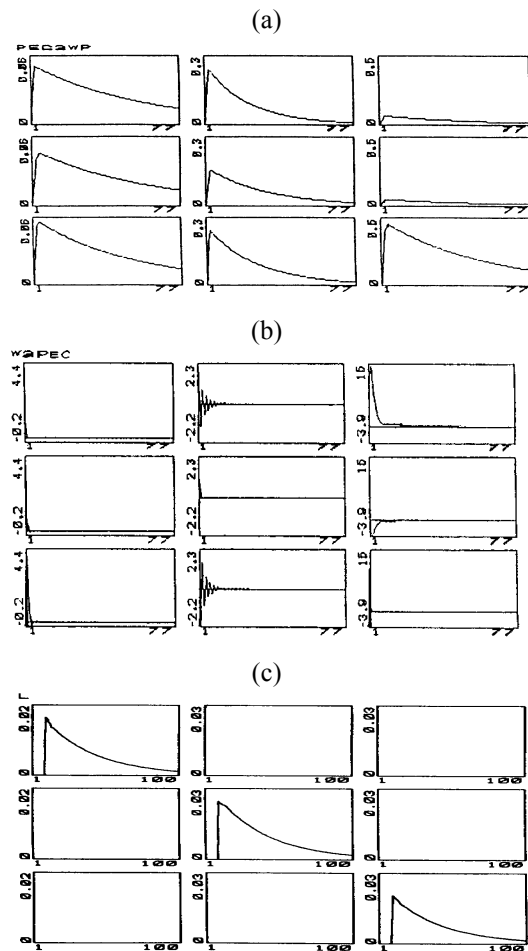
$$P = \sum_{i=1}^M \left( \sum_{j=1}^M \lambda_j^2(t_k) \right) C_i = \sum_{i=1}^M \pi_i C_i, \quad (9a)$$

$$\sum_{i=1}^M C_i = I, \quad (9b)$$

$$V^{-1}LV = WLV = G = [g(t_k)]_{M \times M \times N}, \quad (10a)$$

$$C_i = w_i v_i^T, \quad (10b)$$

with the constituent matrices  $C_i$  (CMA) of the characteristic I/O process modes  $L = [\lambda_i(t_k)]_{N_o \times N_i \times N}$  and characteristic vectors  $w = [w(t_k)]_{N_o \times N}^T$ , which only in special cases may be constant.



**Fig. 2.** Sample results for furnace RZS: (a) the identified pseudo-impulse response matrix; (b) the interaction characteristic vectors with time delays removed; (c) the ordered I/O

characteristic modes extended with time delays.

The set of constituent matrices, with their essential property  $\sum_{i=1}^M C_i = I$  (also see Figure 3 c), act as a modulating distributor of generalized power weights.

It is apparent that the diagonal of  $C_i$  reveals how generalized powers  $\pi_i$  are being associated with a particular diagonal element of  $P$ . In turn,  $P$  may be

obtained as a sum of convolutions between  $\pi_i$  and  $C_i^{ii}$ , the array  $C_i^{ii}$  represents the power distribution weight in  $P_{ii}$  associated with the respective input-output process channel. Hence, the indicator of the proper position of  $\lambda_i$  in the matrix structure is given by means of  $Max\{\sum_{i=1}^M |C_j^{ii}| = I\}$ ,  $i = 1, 2, \dots, M$ ,  $j = 1, 2, \dots, N$ . This indicator also represents the rule to resolve characteristic I/O modes ordering and, consequently, the input-output pairing of controlling and controlled process variables. It is normally expected that the strongest interactions (represented by the elements of CVE and CMA) appear in the main diagonal of  $W$ . That is, a certain sub-system within MIMO systems ‘interacts mainly’ with itself validating the diagonal dominance in the process characteristic I/O modes has been attained.

Figures 3 a, b, c for the real-world RZS pusher furnace provide a good physical illustration of the above presented I/O mode decomposition theorem as well as of the entire discussion in this section. Nonetheless, its is pointed out the word is about approximate representation models of operational processes in the plant object, where precisely the energy-information-matter inter-play goes on.

### 3. ON APPROXIMATION MODELS BY ANN AND KOLMOGOROV’S SECOND THEOREM

Within the context of artificial neural networks (ANN) in this paper, we are interested solely in approximation of functions applied to process model emulation identification. So, we confine

ourselves to a certain narrow area of approximation modelling of dynamical processes primarily using feed-forward ANN’s and solely to networks that emanate directly from Kolmogorov’s seminal results [31]-[32]. The emphasis is on Kolmogorov’s second theorem, which subsequently has been reformulated by Kurkova [34], [35] in more practical ways for ANN applications with somewhat relaxed requirements on approximating functions employed. This discussion follows the main streamline of mathematical theory of function approximation (for in depth study, see [33], [39], [42]) that is relevant for ANN based applications to generate approximate models, which represent stable but complex dynamical processes having natural steady-state equilibrium (e.g., industrial thermal systems) of interest in this comparative study. It should be noted, in addition, that computing ANN system structures are also readily modeled [2], [26], [51] by using theory of directed graphs [7], [25].

Now let here recall (6), which represents a fairly general input-output representation of time-discretized system models, where by assumption the underlying mapping  $f: R^{N_y+N_u} \rightarrow R^{N_o}$  is continuous. In its very essence, the neural modeling approach makes use of ANNs for approximate representation of this function. Hence the neural modeling, in its very essence, is a problem of mathematical approximation theory.

Formal mathematical formulation of the problem of neural modeling using feedforward neural networks can be stated as follows. Consider a continuous mapping  $f: S_C \rightarrow R^{N_o}$  where  $S_C$  is an uncountable compact subset of  $R^{N_o}$ . Compactness of  $S_C$  implies that it is closed and bounded, which is a realistic assumption given the fact that operating conditions are specified by means of *admissible* inputs and *sustainable* outputs, i.e. represented by some bounded functions of time. Of course, the mapping  $f$  is not given explicitly, but rather by a finite number of pairs  $\{U(t_k), Y(t_k)\} \in S_C \times R^{N_o}$ ,  $k = 1, \dots, N_{t_r}$ , now  $N_{t_r}$  representing the number of observed

input-output pairs, namely,  $U(t_k) = [y(t_k - 1), \dots, y(t_k - N_y); u(t_k - 1), \dots, u(t_k - N_u)]^T$  is an  $N_o \times (N_i N_u + N_o N_y)$  vector, and  $Y(t_k) = y(t_k)$  is an  $N_o \times 1$  vector. Then, seek to find a representation of the mapping  $f$  by means of known functions and a finite number of real parameters, such that the representation yields uniform approximation of it over set  $S_C$ , which should be well-posed interpolation hyper-surface  $f(S_C)$ . Hence, in essence, this is an existence problem independent of the form in which the mapping  $f$  may be given, and therefore approximating known functions implies qualitative property such as continuity of the mapping  $f$ . On the other hand, by its very essence, finding the interpolation of the continuum  $f(S_C)$  intrinsically implies constructing it by a finite number of argument-value pairs, which may be done by appropriate learning/training method and algorithms [2], [4], [26], [48], [51].

In the Appendix, there are found further points of argument in terms of a brief outline of the known abstract mathematical schema, compiled from the literature, via which main theorems of mathematical approximation theory have been derived along with these theorems.

At this point, let briefly relate the discussion in this section to the one in Section II. It may well be noticed in the conceptual definition and the respective math-analytical description of process characteristic I/O modes via k-time sequence matrices, (7)-(8), that these are existent per-se because of the physical nature of the class of real-world plants considered. And, finding their constructive digital computation is dealt with in the associated decomposition theorem, (9)-(10), which is equivalent to learning/training in the neural modeling approach. So, for the time being, one can notice that indeed between these two I/O modeling approaches some kind of reminiscent relationship exists.

In order to proceed further, here we recall and place into proper perspective a subset of the main results involving Stone-Weierstrass [46]

and Kolmogorov's second [32] theorems, and the polemical discussions related to the latter in [19], [20], [27], [30], [34]-[36], [40], [43], [44]. These two theorems constitute the fundamental toolbox of the approximation theory. The Stone-Weierstrass theorem has been identified in the literature as a prime candidate for establishing the property of the existence of approximation representation  $f(S_C)$ , employing ANN's, of the continuous mapping  $f : S_C \rightarrow R^{N_o}$ , and more detail is found in the appendix. In the sequel we focus on Kolmogorov's second theorem and the related polemics.

It is well known that artificial neural networks, because of the biological inspiration and greater transparency of the analysis, do have layered structures. The sigmoidal model is described by

$$y_i = \sum_{j=1}^N \alpha_{ij} \sigma(b_{ij}^T U + d_{ij}), \quad i = 1, \dots, N_Y, \quad (11)$$

where  $\alpha_{ij}, d_{ij} \in R$ ,  $b_{ij} \in R^{N_i \times N_o}$  are parameters, and  $U = U(t_k)$  is the input vector, described at the beginning of this section. This representation equation possess the identity of that of an ANN having input layer, sigmoidal hidden layer, and linear output layer. Similarly the radial-basis-function model has the structure

$$y_i = \sum_{j=1}^N \alpha_{ij} g_{c_{ij}, m_{ij}}(U), \quad i = 1, \dots, N_Y, \quad (12)$$

$$y_i = \sum_{j=1}^N \alpha_{ij} g_{c_{ij}, m_{ij}}(\|U - c_{ij}\| / m_{ij}), \quad i = 1, \dots, N_Y \quad (13)$$

Of course, in both cases, respectively, it is necessary and sufficient to show that the set of all finite linear combinations of SG and RB functions constitute a non-vanishing algebra separating points on a compact  $S_C \subset R^{N_i \times N_o}$ .

This has been done, but the same has been Gaussian-function networks only after introducing the additional requirement of convexity of  $S_C \subset R^{N_i \times N_o}$ .

On the grounds of Stone-Weierstrass theorem it has been found out that the use of both the SG and the RB functions is suitable for uniform approximation of an arbitrary continuous mapping, their interpolation properties,

however, are different; in a sense the RB functions, unlike the SG, are designed for interpolation. On the other hand, the issue on finding the best approximation remains much more involved. Nonetheless, the way deep into the existence problem has been highlighted when Kolmogorov's second theorem, which already had a decisive impact on approximation theory [39], [43], has been brought into ANN prospective [27].

During a decade or so Kolmogorov's theorem, resolving the representation of continuous functions defined on an  $n$ -dimensional cube by sums and superposition of continuous functions of one variable, has been one of the focuses attracting attention in the theory and applications of neural modeling approach. Hence the main result of Kolmogorov [32] as well as the subsequent reformulation by Kurkova [34], [35], who disproved the criticism by Girosi and Poggio [20] and verified its essential relevance for the neural modeling approach, are recalled next.

*Theorem III.1 (Kolmogorov's Second).* Any function continuous on the  $n$ -dimensional unit cube  $E^n$ ,  $E = [0, 1]$ , can be represented in the form

$$f(x_1, \dots, x_n) = \sum_{i=1}^{2n+1} \psi_i \left( \sum_{j=1}^n \varphi_{ij}(x_j) \right), \quad (14)$$

where  $\psi_i$  and  $\varphi_{ij}$  are real continuous functions of one variable, and the functions  $\varphi_{ij}$  are independent of the given function  $f$  while only the functions  $\psi_i$  are specific for the given function.

Kolmogorov has made clear that on the grounds of this theorem it is possible to represent exactly every continuous function of many variables as a superposition of a finite number of continuous functions of one variable and of a single particular function of two variables, viz. addition. Considerably later, Lorentz [39] has proved a simplification of Kolmogorov's theorem in which the functions  $\psi_i$  may be replaced by only one function  $\psi$ , and Sprecher [43] has shown that the functions  $\varphi_{ij}$  by  $\alpha^{ij} \phi_j$  with  $\alpha^{ij} = \text{const}$  and some monotonic

increasing functions  $\phi_j$ . Only in the eighties, Hecht-Nielsen [27] has reformulated Sprecher's version of Kolmogorov's representation theorem and applied to neural network modeling, and so did Funahashi, thus the use of ANN's was made plausible entirely. More recently Lin and Unbehauen [36], Katsura and Sprecher [30], Sprecher [44] made yet other realizations based on Kolmogorov's second theorem.

*Theorem III.2 (Hecht-Nielsen form of Sprecher's version).* Any continuous function defined on the  $n$ -dimensional cube  $E^n$ , where  $E = [0, 1]$ , can be implemented exactly by a three-layered network having  $2n+1$  units in the hidden layer with transfer functions  $\alpha^{ij} \phi_j$  from the input to the hidden layer and  $\psi$  from all of the hidden units to the output layer.

In their work Girosi and Poggio, [20], have pointed to the two drawbacks of this theorem:  $\alpha^{ij} \phi_j$  are highly non-smooth, and  $\phi_j$  on the specific function  $f$  are not representable in a parameterized form. In turn, Kurkova [34], [35] eliminated both these difficulties by using staircase-like functions of a sigmoidal type in a SG feedforward neural network. It is well known that highly non-smooth functions encountered in mathematics are mostly constructed as limits or sums of infinite series of smooth functions (see the Appendix). Kurkova showed that all of the single-variable functions in Kolmogorov's theorem are limits of sequences of smooth functions when staircase-like functions of sigmoidal type are used in the neural network. In fact, this type of function has the property that it can approximate any continuous function on any closed interval with an arbitrary accuracy. Note that, in fact, these functions are directly employed in all discrete-time system representation models and kinds of digital computer generated controls [14], [16], [45].

In the course of deriving her arguments, Kurkova also contributed an appropriate reformulation of Kolmogorov's second theorem within the realm of artificial neural networks. Hence, this reformulation of the celebrated Kolmogorov's representation theorem is presented next.



*Theorem III.3 (Kolmogorov-Kurkova).* Let  $n \in \mathbb{N}$  with  $n \geq 3$ ,  $\sigma: R \rightarrow E$  be a sigmoidal function,  $f$  be a function of class  $C^0(E^n)$ ,  $f \in C^0(E^n)$ , and  $\varepsilon$  be a positive real number. Then there exist  $k \in \mathbb{N}$  and staircase-like functions  $\psi_i, \phi_{ij} \in S(\sigma)$  such that for every  $x = (x_1, \dots, x_n) \in E^n$  holds true

$$\left| f(x_1, \dots, x_n) - \sum_{i=1}^k \psi_i \left( \sum_{j=1}^n \phi_{ij}(x_j) \right) \right| < \varepsilon, \quad (15)$$

where  $S(\sigma)$  is the set of all staircase-like functions of the form

$$f_{SC\sigma}(x) = \sum_{j=1}^k \alpha_j \sigma(b_j x + c_j). \quad (16)$$

This theorem implies that any continuous function can be approximated arbitrarily well by a four-layer sigmoidal feedforward neural network. It should be noted that already has been established (albeit not by Kolmogorov arguments) that even three layers are sufficient for approximation of general continuous functions [6], [19]. In general, approximate implementation of  $\psi_i$  does not guarantee an approximate implementation of the original the specific function  $f$  [36], implying that the former must be exactly realized.

Clearly, also there are limitations of the applicability of Kolmogorov's theorem to neural networks for approximation of mappings  $f: S_C \rightarrow R^{N_o}$ . For instance, in applications, one has to answer the question whether an arbitrary given multivariate function can be represented by an approximate realization of the corresponding functions  $\psi_i$  of one variable. However, the efforts towards applying this theorem are fruitful and important, as some useful neural networks like the sigmoidal feedforward network are correctly and properly described by Kolmogorov's representation theorem [17]. Moreover, it is readily applicable to real-world complex plants such as industrial furnaces [15], [16]. And, so has been the case when applying CVE-CPA-CMA decomposition theorem that worked always for multi-zone furnaces [14]-[16], [23].

#### 4. A SUMMARY OF COMPARATIVE FINDINGS

A closer comparison analysis of the essential substances presented in Sections II and III reveals fairly well the existent of similarities and distinctive differences between the respective input-output time-domain model representations [17]. The ones discussed in the previous section, the foundation of which is Kolmogorov's theorem and which make use of ANNs (e.g., see [16], [17]), of course, are the more 'general' ones, because these are fully consistent with the mathematical approximation theory [33], [39], [42] precisely. These are learned from the I/O time sequences as they are recorded during plant operation.

It should be noted, however, the kTSM representation model of characteristic I/O modes decomposition [10], [13] in Section II also are based on operational real-world I/O time sequences carrying on most essential features of the process [14]. This, in turn, sheds new light on the similarities and differences between both these approximation model representations. The main difference appears in the fact that Kolmogorov ANN do imply decomposition and generalization in the time domain, while I/O k-time sequences CVE-CPA-CMA imply solely decomposition [14], [17].

By referring more closely to characteristic I/O mode decomposition formulae, (9)-(10), and to Kurkova's formulae of Komogorov's representation theorem, (15)-(16), despite the considerable difference a common underlying background is evident. Although derived by means of different arguments – the former have also exploited the engineering physics – clearly both have been derived within the context of approximation theory of multivariate functions [17].

Now, let consider comparing (14) in the original Kolmogorov representation theorem and the approximation inequality (15) of Kurkova, which implies replacing the exact expression representation with an approximate one. Also, for a convenient parallel between neural-network approximation in this Section III and convolution approximation in Section II, and recall the earlier comparison setting for expressions (7)-(8) and (9)-(10). In turn, it may

well be seen that a noticeable reminiscence becomes quite apparent [17].

In the final consequence, from the points of view in the present discussion, it has become evident above that ultimately the characteristic I/O mode decomposition also has its support in the general validity of Kolmogorov's theorem. And, this provides the explanation why fairly simple designs of PI partially-decoupling digital controls of industrial furnaces [14], [23], [45] turned out to work in robustness way and considerably better than initially expected.

## 5. CONCLUSION

In the literature, there have been developed system representation models via both math-analytical and computational-intelligence approaches. However, both essentially involve mathematical approximation theory. In this work, a comparison study of approximate representation models of a class of complex stable MIMO processes through the time domain input-output view of math-analytical time-sequence and of the artificial neural networks has been elaborated.

An analytical decomposition representation of complex MIMO dynamical processes having natural steady-state equilibrium (such as in thermal systems) has been set relative to the ANN approximation representations based on Kolmogorov's theorem and comparison analysis carried out. The main findings resulting out of this study have been presented, which show that both emanate from the theoretical background of Kolmogorov's second representation theorem. These provide a new insight into the actual energy-information-matter inter-play in real-world dynamical processes. In addition, these highlighted the efficiency and robustness of fairly simple industrial digital controls, designed and implemented in the past, which in fact is inherited from model I/O decomposition approximation employed.

Lastly but not least, it may well be argued that Kolmogorov's theorem is much more relevant to all engineering disciplines and physical science than insofar recognized [17]. For, as emphasized by Bertrand Russell, all the exact science is dominated by the idea of

approximation and engineering sciences even more so.

## APPENDIX

The original Weierstrass theorem is the well known first approximation theory result [33], [39], [42]. This theorem has showed that an arbitrary continuous function  $f:[a,b] \rightarrow R$  (the relevance is to single-output plants) can be uniformly approximated by a sequence of polynomials  $\{p(x)\}$  to within a desired degree of accuracy. Thus, given  $\varepsilon > 0$ , it is possible always to find an integer  $N_n \in N$ , such that for any  $n > N_n$  the boundedness  $|f(x) - p_n(x)| < \varepsilon$  is fulfilled uniformly on an arbitrary interval  $[a, b]$ . Apparently, this is the 'archetype' of the problem considered involving a number of real parameters (the coefficients  $p_{N_n+1}(x)$  that is finite for a given  $\varepsilon$ , degree of accuracy. In the work of Weierstrass, there is present a restriction to the class of polynomials.

In [46] Stone has studied Weierstrass theorem on a general abstract level with the purpose of finding the general properties of approximating functions, which may not be intrinsic to polynomials. The model scheme of his abstract approach provide the proper way of tackling the approximation I/O modeling of dynamical processes no matter whether by sequence or neural-net structures, and requires a proper outline. Notice that his mathematically constructive conceptualization schema also supports Theorem II.1 on characteristic I/O modes decomposition.

All sets of real numbers  $R$  are composed of rational,  $R_r$ , and irrational,  $R_i = R - R_r$ , numbers. In essence, all computations, either by hand or by machine, are performed using the rationals because of the fact that any real number can be approximated to a desired accuracy by a sequence in  $R_r$ . Once it is well established what natural numbers are and arithmetic operations are defined, the formal construction of the (sub)set of rational numbers  $R_r$  is simple. It should be noted, however, that the (sub)set of irrationals  $R_i = R - R_r$  is indispensable for theoretical investigations

because set  $R$  is complete while neither of sets  $R_r$  and  $R_i = R - R_r$  are. The emphasized remarkable property of rationals is formalized by stating that set  $R_r$  is dense in  $R$ , or, equivalently, that  $R$  is the *closure of  $R_r$* . That is to say  $R$  is the smallest set in which all rational Cauchy sequences have limits. Hence, any number which can be approximated by a sequence with terms in set  $R_r$  is a real number.

Following Stone, consider the problem (a converse to that of Weierstrass): given the set,  $F_B$ , of all continuous functions from a compact subset,  $S_C \subset R^{N_i \times N_o}$ , to the set of reals,  $R$ , then find a proper subset of it,  $F_A \subset F_B$ , such that  $F_B$  is the closure of  $F_A$ . Note that when considering the approximation of functions of a real argument rather than approximation of real numbers, the set  $F_B$  is playing the role of the set  $R$  of reals and the set  $F_A$  is playing the role of set  $R_r$  of irrationals. Also note that when the word is about function approximation, it is desirable to perform simple algebraic operations on  $F_A$  (and  $F_B$ ), e.g. forming of linear combinations and compositions, which to certain extent is equivalent to arithmetic operations on  $R_r$  (and  $R$ ). Similarly to convergency and uniform convergency of sequences of real numbers in  $R_r$  and  $R$ , convergency and in particular uniform convergency of sequences of functions is desirable [33], [39], [42]. Thus, if  $\{f_n\}$  is a sequence of functions in  $F_A$  such that  $f_n \rightarrow f$ , it is desirable the limit  $f$  of  $\{f_n\}$  to be in  $F_B$ . For issues investigated in this paper, sequences of functions with terms in  $F_A$  are required to *converge uniformly* on  $S_C \subset R^{N_i \times N_o}$ , which is guaranteed if set  $F_A$  is composed of continuous real functions on a compact set  $S_C \subset R^{N_i \times N_o}$ .

The above constructive, abstract, mathematical schemata via which Stone has extended Weierstrass theorem, cited below, remains valid

for any functions and not solely for the continuous ones.

*Theorem A.1 (Stone-Weierstrass).* Let  $F_A$  be an algebra of some continuous functions from a compact  $S_C \subset R^{N_i \times N_o}$  to  $R$  such that  $F_A$  separates points on  $S_C$  and vanishes at no point of  $S_C$ . Then the uniform closure  $F_B$  of  $F_A$  is consisted of all continuous functions from  $S_C$  to  $R$ .

It is emphasized further that Stone-Weierstrass theorem remains valid for multivariate continuous functions too,  $S_C \subset R^{N_i \times N_o}$  to  $R^{N_o}$ ,  $f: S_C \rightarrow R^{N_o}$ , because the co-domain of a vector-valued function is Cartesian product of its components due to conditional property of an algebra of functions [33], [39], [42]. That is, it is valid for multi-output plants defined by (6) in Section II that are mappings  $f: R^{N_y + N_u} \rightarrow R^{N_o}$  provided an appropriate compact subset of  $R^{N_y + N_u}$  be constructed. With regard to neural modeling approach, however, Stone-Weierstrass theorem has to be viewed in the light of employing sigmoidal (SG; roughly speaking almost-smooth saturation nonlinearity) and radial-basis (RB; roughly speaking, almost smooth semi-sphere nonlinearity) functions. Their precise formal definitions are found in the literature cited.

## REFERENCES

- [1] P. Albertos, A. Sala, and A. Dourado, "Learning Control of Complex Systems", *In Control of Complex Systems*, K.J. Aström, P. Albertos, M. Blanke, A. Isidori, W. Schaufelberger, and R. Sanz, Eds., Ch. 6, pp. 123-141. London, UK: Springer-Verlag, 2001.
- [2] J. A. Anderson, *Introduction to Neural Networks*. Cambridge, MA: MIT Press, 1995.
- [3] A. C. P. M. Backx, "Identification of an Industrial Process: A Markov Parameter Approach" *Doctoral Thesis.*, Eindhoven, NL: Eindhoven University of Technology, 1987.
- [4] C. Bishop, *Neural Networks for Pattern Recognition*. Oxford, UK: Clarendon, 1995.

- [5] V. H. Cheng and C. A. Desoer, "Discrete time convolution control systems." *Intl. J. Control*, vol. 36, no. 3, pp. 367-407, 1982.
- [6] G. Cybenko, "Approximation by superposition of a sigmoidal function." *Maths. Control, Signals & Systems*, vol. 2, pp. 303-314, 1989.
- [7] N. Deo, *Graph Theory with Applications to Engineering and Computer Science*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
- [8] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York, NY: Academic Press, 1975.
- [9] G. M. Dimirovski, N. E. Gough, and S. Barnett, "Categories in systems and control theory." *Intl. J. Systems Science*, vol. 8, no. 9, pp. 1081-1090.
- [10] G. M. Dimirovski, S. Barnett, D. N. Klefouri, and N. E. Gough, "An input-output package for MIMO non-linear control systems" (cited in M.G. Singh's [Editor-in-Chief] *Systems and Control Encyclopedia: Theory, Technology, Applications*, Oxford, Pergamon Press, 1987; see vol. 5, pp. 3382-83). In *Software for Computer Control*, M. Novak, Ed., pp. 265-273. Oxford, UK: Pergamon Press, 1979.
- [11] G. M. Dimirovski and N. E. Gough, "Digital modeling and simulation of technological systems by means of k-time sequence matrices." In *Proceedings of the 6<sup>th</sup> International Symposium on Computer at the University*, Paper 605.(1-10). Zagreb, HR: The SRCE and University of Zagreb, 1984.
- [12] G. M. Dimirovski and N. E. Gough, "On a structural duality and input-output properties of a class of non-linear multivariable control systems." *Facta Universitatis Series EE*, vol. 3, no. 1, pp. 1-9, Jan. 1990.
- [13] G. M. Dimirovski, V. P. Deskov, and N. E. Gough, "On the ordering of characteristic input-output modes in MIMO discrete-time systems." In *Mutual Impact of Computing Power and Control Theory*, M. Karny and K. Warwick, Eds., pp. 227-234. London, UK: Plenum Publishing, 1993.
- [14] G. M. Dimirovski, "Learning identification and design of industrial furnace control systems", in *ESF-COSY Lecture Notes on Iterative Identification and Control Design*, P. Albertos and A. Sala, Eds., Ch. III.3, pp. 259-287. Valencia, ES: European Science Foundation and DISA of Universidad Politecnica de Valencia, 2000.
- [15] G. M. Dimirovski, A. Dourado, N. E. Gough, B. Ribeiro, M. J. Stankovski, I. H. Ting and E. Tulunay, "On learning control in industrial furnaces and boilers," in *Proceedings of the 2000 IEEE International Symposium on Intelligent Control*, P. P. Grumpos, N. T. Kousoulas and M. Polycarpou, Eds., pp. 67-72. Piscataway, NJ: The IEEE and University of Patras, 2000.
- [16] G. M. Dimirovski, A. Dourado, E. Ikonen, U. Koretela, J. Pico, B. Ribeiro, M. J. Stankovski and E. Tulunay, "Learning Control of Thermal Systems", In *Control of Complex Systems*, K.J. Aström, P. Albertos, M. Blanke, A. Isidori, W. Schaufelberger, and R. Sanz, Eds., Ch. 14, pp. 317-337. London, UK: Springer-Verlag, 2001.
- [17] G.M. Dimirovski and Yuanwei Jing, "Parallels of Kolmogorov Neural Networks and Characteristic Input-Output Modes Decomposition", *DCE-DUI and ICT-NEU Techn. Report IOD-KNN-01/2001*, Dogus University Istanbul, TR, and Northeastern University, Shenyang, CN, 2001 (unpublished).
- [18] R. C. Dorf, "Analysis and design of control systems by means of time-domain matrices," *Proceedings of Institution of Electrical Engineers*, vol. 109 C, pp. 616-626, 1962.
- [19] K. Funahashi, "On approximate realization of continuous mappings by neural networks," *Neural Networks*, vol. 2, pp. 183-192, 1989.
- [20] F. Girosi and T. Poggio, "Representation properties of networks: Kolmogorov's theorem is irrelevant", *Neural Computation*, vol. 1, pp. 465-469, 1989.
- [21] N. E. Gough and R. S. Al-Thiga, "Characteristic patterns and vectors of discrete multivariable control systems," *Arabian J. Science & Engineering*, vol. 10, no. 3, pp. 253-264, 1985.
- [22] N. E. Gough, and M. A. Mirza, "Computation of characteristic weighting patterns of discrete MIMO control systems", *Int. J. Systems Science*, vol. 18, no. 10, pp. 1799-1814, Oct. 1987.
- [23] N.E. Gough, G. M. Dimirovski, I. H. Ting, and N. Sadaoui, "Robust multivariable control system design for a furnace based on

- characteristic patterns,” In *Application of Multivariable System Techniques*, R. Whalley, Ed., pp. 145-152. London, UK: The IMeasC and Elsevier Applied Science, 1990.
- [24] R. Haber and H. Unbehauen, “Structure identification of nonlinear dynamic systems – a survey of input-output approaches,” *Automatica*, vol. 26, no. 6, pp. 651-677, Jun. 1990.
- [25] F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs*. New York, NY: J. Wiley, 1965.
- [26] S. Haykin, *Neural Networks: A Comprehensive Foundation (2<sup>nd</sup> ed.)*. New York, NY: Macmillan, 1999.
- [27] R. Heht-Nielsen, “Kolmogorov’s mapping neural network existence theorem.” In *Proceedings of the IEEE Intl. Joint Conference on Neural Networks*, vol. 3, pp. 11-14. New York, NY: The IEEE, 1987.
- [28] K. J. Hunt, D. Sbarbaro, R. Zbikowski and P. J. Gawthrop, “Neural networks for control systems – A survey.” *Automatica*, vol.28, no.6, pp. 1083-1112, 1992.
- [29] I. S. Iohvidov, *Hankel and Toeplitz Matrices and Applications*. Basel, CH: Birkhauser, 1982.
- [30] H. Katsuura and D. A. Sprecher, “Computational aspects of Kolmogorov’s theorem.” *Neural Networks*, vol. 7, pp. 455-461, 1994.
- [31] A. N. Kolmogorov, “On the representation of continuous functions of several variables by superposition of continuous functions of a smaller number of variables” (in Russian). *Dokladi Akademii Nauk SSSR*, vol. 108, pp. 358-359, 1956.
- [32] A. N. Kolmogorov, “On the representation of continuous functions of several variables by superposition of continuous functions of one variable and the addition” (in Russian). *Dokladi Akademii Nauk SSSR*, vol. 114, pp. 953-956, 1957.
- [33] A. N. Kolmogorov and S.V. Fomin, *Elements of the Theory of Functions and Functional Analysis (in Russian)*. Moscow, RUS: Nauka, 1968.
- [34] V. Kurkova, “Komlogorov’s theorem is relevant.” *Neural Computation*, vol. 3, pp. 617-622, 1991.
- [35] V. Kurkova, “Komlogorov’s theorem and multi-layer neural networks.” *Neural Networks*, vol. 5, pp. 501-506, 1992.
- [36] J. N. Lin and R. Unbehauen, “On the realization of Kolmogorov’s theorem.” *Neural Computation*, vol. 5, pp.18-20, 1993.
- [37] L. Ljung and T. Soederstroerm, *Theory and Practice of Recursive Identification*. Cambridge, MA: The MIT Press, 1993.
- [38] L. Ljung, *System Identification: Theory for the User (2<sup>nd</sup> ed.)*. Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [39] G. G. Lorentz, *Approximation of Functions*. New York, NY: Holt, Reinhart & Winston, 1966.
- [40] T. Poggio and F. Girrosi, “Networks for approximation and learning.” *Proceedings of the IEEE*, vol. 78, pp. 1481-1497, 1990.
- [41] M. B. Priestly, *Non-linear and Non-stationary Time Series Analysis*. London, UK: Academic Press, 1988.
- [42] W. Rudin, *Principles of Mathematical Analysis (3<sup>rd</sup> ed.)*. Auckland: McGraw- Hill, 1976.
- [43] D. A. Sprecher, “On the structure of continuous functions of several variables.” *Trans. American Mathematical Society*, vol. 115, pp. 340-355, 1965.
- [44] D. A. Sprecher, “A numerical implementation of Kolmogorov’s superposition II.” *Neural Networks*, vol. 10, pp. 447-457, 1997.
- [45] M. J. Stankovski, Non-Conventional Control of Industrial Energy Conversion Processes in Complex Heating Furnaces (*Doctoral Thesis*; Supervisor G.M. Dimirovski). Skopje, MK: SS Cyril and Methodius University, 1997.
- [46] M. H. Stone, “The generalized Weierstrass approximation theorem.” *Mathematics Magazine*, vol. 21, pp. 167-184, 237-254, 1948.
- [47] H. S. Tsien, *Engineering Cybernetics*. New York, NY: Mc-Graw-Hill, 1950.
- [48] P. J. Werbos, “Backpropagation through time: What it does and how to do it.” *IEEE Proceedings*, vol. 78, pp. 1550-1560, 1990.
- [49] B. Widrow and M. A. Lehr, “Thirty years of adaptive neural networks: Perceptron, madaline, and back-propagation.” *Proceedings of the IEEE*, vol. 78, pp. 1415-1442, 1990.

- [50] A. N. Whitehead and B. Russell, *Principia Mathematica* (2<sup>nd</sup> ed.). Cambridge, UK: Cambridge University Press, 1927.
- [51] H. White, *Artificial Neural Networks*. Cambridge, MA: Blackwell, 1992.
- [52] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine* (2<sup>nd</sup> ed.), New York, NY: J. Wiley, 1961.
- [53] L. A. Zadeh and E. Polak *Systems Theory*, New York, NY: Academic Press, 1969.
- [54] L. A. Zadeh, "Fuzzy sets and systems". In *Proceedings of the Symposium on System Theory*, pp. 29-37. Brooklyn, NY: Polytechnic Institute, 1965.
- [55] L. A. Zadeh, "Fuzzy logic, neural networks and soft computing," *Comm. ACM*, vol. 37, pp. 77-84, Mar. 1994.



**Dr. Georgi M. Dimirovski** (IEEE M'86, SM'97) was born on 20.12.1941 in v. Nestram, Greece. He has obtained his degrees Dipl.-Ing. in EE from (his home) SS Cyril & Methodius University of Skopje, Republic of Macedonia, then M.Sc. in EEE from University of Belgrade, Republic of Serbia – F.R. of Yugoslavia, and Ph.D. in ACC from University of Bradford, England - UK, in 1966, 1974 and 1977, respectively. Subsequently, he has had a postdoctoral position at University of Bradford in 1979. Since, he has had visiting academic positions in Bradford, Brussels, Coimbra, Duisburg, Istanbul, Linz, Valencia, Wolverhampton and Zagreb, and also participated to ESF Scientific Programme on Control of Complex Systems (COSY 1995-99) of European Science Foundation under the leadership of Prof. K. J. Astrom. He was editor of two volumes of IFAC Series, and has published more than 30 journal and 150 conference papers in renown IFAC and IEEE proceedings, and successfully supervised 2 postdoctoral, 14 PhD and 25 MSc as well as more than 250 graduation students' projects. He served on the Editorial Board of Proceedings Instn. Mech. Engrs. J. of Systems & Control Engineering (UK) and was Editor-in-Chief of J. of Engineering (MK), and currently is serving on the Editorial Board of Facta Universitatis J. of Electronics & Energetics (YU). He has developed a number of undergraduate and/or graduate courses at his home university and at universities in Bradford, Linz, Istanbul and Zagreb. He has served on IPCs for many IFAC, IEEE, ECC conferences and others.

**Dr. Yuanwei Jing**, has been born in province Liaoning of P.R. of China in 1957. He obtained his graduation Diploma degree in Mathematics, and MSc and Doctoral degrees in Systems & Control from his home Northeastern University, Shenyang, P.R. of China. He has had a position at Institute of Control Theory within the School of Information Science & Northeastern University of Shenyang, consecutively being promoted to Assistant Lecturer, Assistant and Associated Professor, and to Professor of Control & Information Science, after completing his 1996/97 post-doctoral specialization in Automation & Control of Interconnected Systems at SS Cyril & Methodius University of Skopje, Macedonia, with Professor G.M. Dimirovski. During his post-doctoral position at SS Cyril & Methodius University of Skopje, he has also been a visiting Associate Professor of Systems Theory with the Institute of Automation & Systems Engineering at the Faculty of Electrical Engineering. During the years 1998 and 1999, Professor Jing has been a Visiting Research Fellow at University of Missouri – Kansas City, working along with Professor K. Sohrawy on incentive control and optimization of QoS network systems using game theory. He has published more than 20 journal papers and more than 100 conference papers in renown IFAC, Asian CC, and IEEE proceedings. He has successfully supervised 5 doctoral and 15 master student's projects.. Currently, Professor Jing is Vice-Director of the Institute of Control Theory at Northeastern University of Shenyang.