Appl. Math. and Comp. Sci., 1996, vol.6, No.3, 463–483

A NEW APPROACH TO NEUROCONTROL BASED ON FOURIER ANALYSIS AND NONUNIFORM MULTI-DIMENSIONAL SAMPLING†

ANDRZEJ DZIELIŃSKI*, RAFAŁ ŻBIKOWSKI*

This paper presents a new approach to neurocontrol of deterministic, discrete-time non-linear (NARMA) dynamic systems, given only input-output data of finite length. The demanding setting of scarce knowledge about the plant is motivated by practice of data acquisition and the complexity of discretised dynamics.

The essence of the method is a novel modelling technique based on nonuniform multi-dimensional (N-D) sampling and Fourier Analysis. The right-hand side (RHS) of the NARMA model is reconstructed from nonuniformly spaced N-D samples. This is done by approximating the Fourier transform of the RHS. To this end a feedforward neural network is applied as an implementation of a multi-dimensional interpolating filter in the N-D frequency domain. The neural model obtained in this way is smooth and suitable for the purposes of non-linear control. In order to deal with the modelling error, a new technique of BIBO redesign of the closed-loop system is introduced.

The modelling method is inspired by the ideas of Sanner and Slotine (1992), but it goes far beyond them, resulting in a novel approach. The main advantages of the new algorithm are: the realistic engineering setting, computational simplicity, applicability (mild assumptions), flexible neural implementation and relevance for control.

1. Introduction

Throughout the paper we denote by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ natural, integer, real and complex numbers, respectively. Also, $\mathbb{Z}_+$, $\mathbb{R}_+$ denote the set of non-negative integers and reals, respectively.

In this paper, we introduce and analyse a new neurocontrol strategy for the deterministic, non-linear, single-input single-output (SISO) system given by the discrete-time, $t \in \mathbb{Z}_+$, input-output NARMA (Chen and Billings, 1989) model

$$y(t+1) = f\Big(y(t), \ldots, y(t-n+1), u(t), \ldots, u(t-m+1)\Big) \tag{1}$$

† A preliminary version of this paper was presented at the 2nd International Symposium on *Methods and Models in Automation and Robotics*, 30 August–2 September 1995, Międzyzdroje, Poland

* Department of Mechanical Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland, UK, e-mail: andrzej@mech.gla.ac.uk

with $y \in [a,b] \subset \mathbb{R}$, $u \in [c,d] \subset \mathbb{R}$ and $f : D \to [a,b]$ with the domain of definition $D = [a,b]^n \times [c,d]^m$. It is physically natural that output y and input u assume only finite values on a connected set and can attain their bounds. In fact, D is a compact, connected and convex subset of $\mathbb{R}^{n+m}$.

We assume that all we know about (1) are measurements of current and past values of y and u with very little *a priori* knowledge of f. The demanding setting is motivated by the engineering problems in which we are given raw input-output data only, acquired at discrete-time instants. Also, the *a priori* knowledge of the underlying continuous-time dynamics is usually lost in the process of discretisation leading to the NARMA representation (1).

As mentioned above, model (1) is obtained by discretisation (Kalkkuhl and Hunt,1996) of a deterministic non-linear (Lipschitz) continuous-time, $t \in \mathbb{R}_+$, SISO control system

$$\begin{aligned} \dot{x} &= f_1(x,u) \\ y &= h(x) \end{aligned} \tag{2}$$

with $x \in X \subset \mathbb{R}^{n_1}$ and $y \in [a,b] \subset \mathbb{R}$, $u \in [c,d] \subset \mathbb{R}$ and an initial condition $x(t_0) = x_0$. This, in general, is an approximation process (Kalkkuhl and Hunt, 1996), and we cannot expect any mathematically nice properties of f, even if the underlying continuous-time model possesses them. Moreover, the input-output model (1), obtained from the discrete-time state-space description (Chen and Billings, 1989; Leontaritis and Billings, 1985), is valid only locally and therefore it is not the 'ultimate black-box'.

This paper proposes a new neurocontrol framework for modelling of systems (1). It is based on the application of multi-dimensional (N-D) Digital Signal Processing and Sampling Theory. The method substantially expands and improves the basic idea introduced by Sanner and Slotine (1992) and goes beyond it, resulting in a novel approach.

The paper is organised as follows. The new algorithm is described in Section 4.2 and the preceding sections explain the motivation and necessary background material. Thus, Section 2 presents the fundamental underlying problems arising in realistic modelling of (1) from real-world data. Particular attention is paid to engineering and computational considerations and their mathematical consequences. Section 3 deals with one of the central themes (and novelty) of the paper: nonuniform multi-dimensional sampling. Sections 2 and 3 set the stage for Section 4, where the new neural method is put in the context and its features explained. The use of the novel modelling approach for stable control is described in Section 5. The paper ends with conclusions.

2. Underlying Problems

In this section we explain the interplay between real-world engineering and computational constraints and mathematical techniques involved in modelling of (1). We

discuss fundamental problems which arise when Fourier Analysis is applied in this context. We identify function extension as the central issue and conclude that the space-limited extension seems to be the most natural solution.

2.1. Extension Problem

As we argue in Section 3.1, a natural approach to modelling unknown f in (1) is via multi-dimensional sampling and thus a relevant mathematical tool is N-D Fourier analysis. This involves f defined on the *whole* of $\mathbb{R}^{m+n}$, $N = m + n$, while the engineering problem (described in Section 1) defines f on the bounded domain $D = [a, b]^n \times [c, d]^m$ only. Thus, before proceeding to analyse the process of multi-dimensional sampling (Section 3), we have to address the *extension problem*. That is: how should $f : D \to [a, b]$ be extended to $\mathbb{R}^{m+n}$ or what values should be assigned to f for its arguments beyond D?

Because of the multi-dimensional sampling involved, the extension should be Fourier transformable and therefore the minimal requirement is that it belongs to $L^1(\mathbb{R}^{m+n}) \cap L^2(\mathbb{R}^{m+n})$ (Stein and Weiss, 1971), i.e., is absolutely and square integrable. Thus the minimal assumption for f in (1) is $f \in L^1(D) \cap L^2(D)$, a very mild condition, especially in the view of boundedness of f, i.e., $f(D) \subseteq [a, b] \subset \mathbb{R}$.

In what follows, we discuss three approaches: band-limited extension, Fourier series extension and space-limited extension. We conclude that the best method, both from theoretical *and* practical viewpoint, is the space-limited extension (see also (Dzieliński and Żbikowski, 1995a; 1995b).

Band-Limited Extension. The most obvious approach (Sanner and Slotine, 1992) is to extend f as a multi-dimensional band-limited function, i.e., having Fourier transform non-zero only on a bounded set. This, however, implies that the extended f, say $\tilde{f}$, would be an entire function,[1] so, in particular, analytic in the whole of $\mathbb{C}^{m+n}$. Thus $\tilde{f}(x) = \sum_{k \in \mathbb{Z}_+^{m+n}} a_k x^k$ everywhere with $k = (k_1, \ldots, k_{m+n}) \in \mathbb{Z}_+^{m+n}$ and $x^k = x_1^{k_1} \cdots x_{m+n}^{k_{m+n}}$, where $x \in \mathbb{R}^{m+n}$. It is also true that $\tilde{f}(z) = \sum_{k \in \mathbb{Z}_+^{m+n}} a_k z^k$ everywhere, where $z \in \mathbb{C}^{m+n}$. In particular, f would have to be complex analytic on D.

This strong requirement of smoothness of f considerably limits the class of models (1) which could be analysed, themselves already locally valid only (see Section 1).

It cannot be guaranteed (Havin and Jöricke, 1994) that small changes in the Fourier transform result in small changes in its inverse, $\tilde{f}$. This is an example of an ill-posed problem and as such requires regularisation (Tikhonov and Arsenin, 1977).

Fourier Series Extension. Since f is naturally defined on a (multi-dimensional) rectangle, an immediate thought is to expand it into multiple Fourier series (Ash, 1976; Stein and Weiss, 1971). However, the N-D Fourier series do not possess many useful features of the one-dimensional counterpart (Ash, 1976). First of all, there are

[1] In fact, it would be a Paley-Wiener function; see (Żbikowski and Dzieliński, 1996a).

various ways of forming partial sums (summation on rectangles, circles etc. of the lattice $\mathbb{Z}^{m+n}$). They profoundly affect convergence and different summing schemes are useful for different functions, with no simple rules.

Unless f is constant on the boundary of D, the extension $\tilde{f}$ will be discontinuous on the boundary. This results in the multi-dimensional Gibbs phenomenon on the boundary ∂D, i.e., oscillation of the partial sums independent of the number of terms in the sums. The value of the Fourier series on ∂D will not necessarily (Pinsky *et al.*, 1993) tend to the average of the function values $(f(\partial D^-) + f(\partial D^+))/2$, as is in the 1-D case.

One of the useful features of one-dimensional Fourier series is the localisation property. This means that convergence of the series at a point depends only on the behaviour of the expanded function in a neighbourhood of the point. This is not necessarily true in N-D (Pinsky *et al.*, 1993) and even for innocently looking functions their values on the boundary may profoundly affect convergence in the interior. Smoothness helps (at least for the radially symmetric f), but again limits the class of functions f in (1).

Space-Limited Extension. The band-limited approach requires a strong smoothness of f. However, the ideas can be exploited due to dualism of Fourier Analysis. If we reverse the roles of the function and its Fourier transform, then we naturally come to the idea of a space-limited function, whose transform is complex analytic. Then the sampling is done in the (multi-dimensional) frequency domain and the Fourier transform is reconstructed. The only (very mild) requirement for the space-limited function is that it has a unique and invertible Fourier transform.

Let us make these ideas precise. A function $\tilde{f} : \mathbb{R}^{m+n} \to [a, b]$ is called space-limited if it is non-zero only in a bounded subset D of $\mathbb{R}^{m+n}$. Thus, the extension is simply

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

It follows from the Plancherel-Pólya theorem (Ronkin, 1974) that the Fourier transform $\tilde{F}$ of $\tilde{f}$ is entire (Żbikowski and Dzieliński, 1996a). Because of duality of Fourier analysis, we may as well say that $\tilde{F}$ is band-limited, its Fourier transform being $\tilde{f}$ of (3) and the bandwidth being D. However, to avoid the confusion with the usual meaning of the term 'band-limited', we prefer calling $\tilde{F}$ *harmonically-limited* (Kim and Bose, 1990).

Thus, the strong smoothness is a feature of the harmonically-limited Fourier transform $\tilde{F}$ of $\tilde{f}$, not $\tilde{f}$ itself. It is important that the only requirement on f in (1) is that $f \in L^1(D) \cap L^2(D)$. This, of course, ensures $\tilde{f} \in L^1(\mathbb{R}^{m+n}) \cap L^2(\mathbb{R}^{m+n})$.

Hence the problem of reconstruction of f is reduced to the reconstruction of the harmonically-limited Fourier transform $\tilde{F}$ of the space-limited extension (3). This is aided by analyticity of $\tilde{F}$ (Fiddy, 1987), but primarily by the availability of N-D efficient Fourier transform approximations (see Section 2.2) generating samples of $\tilde{F}$ from the samples of $\tilde{f}$. Note that the samples of $\tilde{f}$ will contain only the values of f

on D, which would not necessarily be the case for the methods of Sections 2.1 and 2.2 (there $\tilde{f}$ *may* be non-zero outside D).

Finally, it should be noted that this approach has been successfully applied in reconstruction of (2-D) images and there exists a wealth of practical, numerically sound and tested algorithms. The main emphasis in image reconstruction is, however, on uniform sampling, but there have emerged nonuniform methods very recently (Feichtinger and Gröchenig, 1993).

The space-limited extension, dual to the band-limited one, seems to possess some desirable features of which the mild assumptions on f are of paramount importance. Therefore further developments are devoted to this approach.

2.2. Fourier Transform Approximation

In this section we briefly summarise the simplifications resulting from using a computationally acceptable version of the Fourier transform, i.e., Discrete Fourier Transform (or rather Fast Fourier Transform) instead of the transform itself.

The Fourier transform is defined as an integral of a function of continuously varying arguments as follows

$$F(\omega_c) = \int_{\mathbb{R}^N} f(x) \mathrm{e}^{-j\omega_c \cdot x} \, \mathrm{d}x \tag{4}$$

where $\omega_c = (\omega_{c1}, \ldots, \omega_{cN}) \in \mathbb{R}^N$ and $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and $\cdot$ denotes the Euclidean inner product $\omega_c \cdot x = \sum_{k=1}^{N} \omega_{ck} x_k$.

Let us note that in the multi-dimensional case we deal with variables of different nature and neither the original domain has anything to do with time, nor its transformed counterpart is an analogue of frequency. This is a purely *formal* application of the Fourier integral to multi-dimensional functions. The variables of the NARMA model (1), which in our approach play the role of x in (4), are the values of system inputs and outputs in consecutive time instants. Their nature may be diverse and interpretation of the variables of the transform is not as clear as for the time signals. Therefore we shall refer to (4) as the Continuous Fourier Transform (CFT) to emphasise that both the original and transformed variables vary continuously.

The consecutive steps (Żbikowski and Dzieliński, 1996a) of necessary simplifications and the resulting approximations are as follows. Starting from CFT we pass to the Discrete Continuous Fourier Transform (DCFT) which is CFT applied to the function f taken at discrete values of x. Then, as the data are of finite length in any practical application, we introduce the Truncated DCFT, i.e., DCFT defined for the function of arguments restricted to a finite set only. Note that both DCFT and the Truncated DCFT, while defined for a discrete function, are themselves functions of continuous arguments ω_c. If the Truncated DCFT is to be of any practical value, a systematic procedure of its *numerical evaluation* must be developed. The *Discrete Fourier Transform (DFT)* provides such a mechanism, whereby sampled values of the Truncated DCFT are computed on a prescribed *finite* set of discrete values of transformed arguments. The values of DFT on the set are found through a finite number of

(possibly complex) additions and multiplications and making the algorithm amenable to digital implementation. Furthermore, the Fast Fourier Transform algorithm can be used to obtain a computationally effective means for obtaining the sampled values.

In the sequel, we call the arguments of f in (4) the *space* variables and the arguments of F the *frequencies*. As mentioned before, even in the one-dimensional case the analogy with the terminology of time signals is false, but we nevertheless use the names in order to avoid excessive wording.

3. Multi-Dimensional Sampling

This section describes the fundamental questions of multi-dimensional sampling and their relevance for the reconstruction of the space-limited extension $\tilde{f}$ of f from (1) for the purposes of neurocontrol. As noted in Section 2.1, this amounts to the reconstruction of the harmonically-limited Fourier transform $\tilde{F}$ of $\tilde{f}$. Section 3.1 describes how nonuniform sampling naturally arises for models (1). Section 3.2 gives a brief background for the N-D sampling theory and its nonuniform aspects. The results serve a starting point for Section 4 on practical algorithms for the N-D functions reconstruction.

3.1. Motivation and Background

If f in (1) is unknown, then modelling can be based on the knowledge of the pairs of multi-dimensional samples $((y_t, \ldots, y_{t-n+1}, u_t, \ldots, u_{t-m+1}), y_{t+1})$, where we put $y_t = y(t)$ etc. for brevity ($t \in \mathbb{Z}_+$).

Modelling of f_1 in (2) based on the pairs $((x, u), \dot{x})$ is not feasible (even if we could measure x), due to the need of inherently noisy estimate of $\dot{x}$. Note, however, that y is lowpass filtered before discretisation to make it band-limited and thus avoid aliasing (Åström and Wittenmark, 1990). Filtering of y and its observation at discrete time instants free model (1) from this problem.

Put $N = m + n$. Given the samples

$$\lambda_k = (\lambda_{1,k_1}, \ldots, \lambda_{N,k_N}) = (y_t, \ldots, y_{t-n+1}, u_t, \ldots, u_{t-m+1}) \quad \text{and} \quad f(\lambda_k)$$

the issue is to reconstruct the multivariable function f, a problem from the multi-dimensional (N-D) Signal Processing (note that it is completely separate from the question of band-limiting of y). The approach was introduced by Sanner and Slotine (1992), but they assumed that the multi-dimensional samples are uniform, i.e., regularly distributed in the domain D of f. This seems to be a simplification, as the dynamics of (1) manifest themselves through irregular samples (see (Żbikowski and Dzieliński, 1996a). The notation used there may be not entirely consistent with that used throughout this paper). For example, if f is linear, i.e., $y_{t+1} = a_0 y_t + \ldots + a_{n-1} y_{t-n+1} + b_0 u_t + \ldots + b_{m-1} u_{t-m+1}$, then even for a constant input the output will not take values in constant increments, but according to the slope of the hyperplane determined by f. Thus even uniformity of u cannot, in

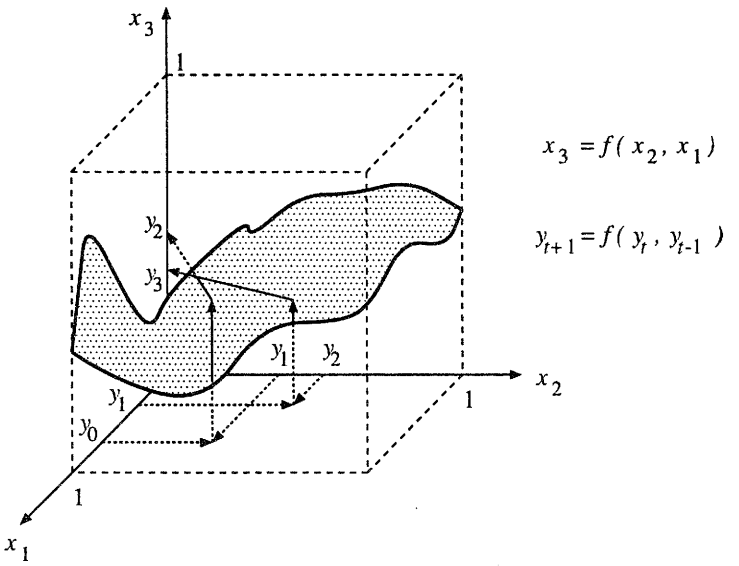

Fig. 1. Iterative map $f : [0,1] \times [0,1] \to [0,1]$ defined by $y_{t+1} = f(y_t, y_{t-1})$.

general, ensure a regular distribution of values of y, because the irregularity of the distribution represents f.

We now examine in detail the nature of this process in low dimensions (as this can be illustrated graphically).

We look at the way nonuniform samples, i.e., λ_k in the pairs $(\lambda_k, f(\lambda_k))$, are generated when f is the right-hand side (RHS) of a dynamic system. For simplicity, instead of dealing with a controlled system of type (1), we concentrate on the low-order autonomous case:

$$y_{t+1} = f(y_t, y_{t-1}) \tag{5}$$

with $y \in [0,1]$. Thus for (5) $\lambda_k = \lambda_{(k_1,k_2)} = (\lambda_{1,k_1}, \lambda_{2,k_2}) = (y_t, y_{t-1})$. We also assume, for illustration purposes only, that f is continuous. Its domain $D = [0,1]^2$ is compact (and connected). The sample points $\lambda_k = \lambda_{(k_1,k_2)} = (\lambda_{1,k_1}, \lambda_{2,k_2}) = (y_t, y_{t-1})$ appear in the x_1x_2 plane, according to the iterative process (5); see Fig. 1. Since system (5) is causal, the samples λ_k arrive in a definite order, i.e., $(k_1, k_2) = (t, t-1)$ with $t \in \mathbb{Z}_+$. In an *off-line* setting the numbering scheme for λ_k can be arbitrary.

If we start with $\lambda_{(1,0)} = (\lambda_{1,1}, \lambda_{2,0}) = (y_1, y_0)$, where $y_1 \in Ox_2$ and $y_0 \in Ox_1$, then we can read out y_2 from the surface representing f. Then y_2 is reflected through $x_3 = x_2$ on the x_2x_3 plane, becoming a point on the Ox_2 axis. In the same time y_1 is reflected through $x_2 = x_1$ on the x_1x_2 plane, becoming a point on the Ox_1 axis. This results in the point (y_1, y_2) in the x_1x_2 plane, corresponding to the sample $\lambda_{(2,1)} = (y_2, y_1)$. We can now read out y_3 from the surface representing f

and repeat the process for $t = 3$. Now y_3 'migrates' from Ox_3 to Ox_2 and y_2 from Ox_2 to Ox_1 generating the point (y_2, y_3) on the x_1x_2 plane corresponding to the sample $\lambda_{(3,2)} = (y_3, y_2)$ etc.

The essential observation is that y_t, i.e., λ_k of the pairs $(\lambda_k, f(\lambda_k))$ appear on Ox_1 and Ox_2 in a *nonuniform* (irregular) way. They will be, in general, unevenly spaced and their pattern of appearance will depend on the dynamics of (5), or the shape of f.

3.2. Nonuniform Sampling in N Dimensions

The main result of multi-dimensional uniform sampling theory is a generalisation of the Shannon Sampling Theorem (Shannon, 1949) to many variables. This may be summarised as follows (Petersen and Middleton, 1962).

Theorem 1 (N-Dimensional Sampling Theorem). *Let $h : \mathbb{R}^N \to \mathbb{R}$ be such that both its N-dimensional direct, H, and inverse Fourier transforms are well-defined. If the spectrum $H(\omega_1, \ldots, \omega_N)$ vanishes outside a bounded subset of N-dimensional space, then h can be everywhere reconstructed from its samples $h(x_k) = h(x_{1,k_1}, \ldots, x_{N,k_N})$, taken over a lattice of points $\{k_1v_1 + k_2v_2 + \ldots + k_Nv_N\}, k_i \in \mathbb{Z}, i = 1, \ldots, N$, provided that the vectors $\{v_1, v_2, \ldots, v_N\}$, where $\{v_i\} \in \mathbb{R}^N, i = 1, \ldots, N$, are small enough to ensure non-overlapping of the spectrum $H(\omega_1, \ldots, \omega_N)$ with its periodic images on the lattice defined by the vectors $\{w_1, w_2, \ldots, w_N\}$, where $\{w_j\} \in \mathbb{R}^N, j = 1, \ldots, N$, with*

$$v_i \cdot w_j = 2\pi\delta_{ij} \tag{6}$$

where δ_{ij} is Kronecker's symbol and $\cdot$ the inner product.

Condition (6) is a multi-dimensional generalisation of the Shannon condition (Shannon, 1949).

Theorem (1) allows finding an N-D interpolation 'filter' with the multi-dimensional impulse response g such that we may reconstruct h from its samples, i.e.,

$$h(x) = \sum_{k \in \mathbb{Z}^N} h(x_k) g(x - x_k), \qquad x \in \mathbb{R}^N \tag{7}$$

If we try to apply the multi-dimensional sampling theorem to the modelling of non-linear systems, we have to address the issue of *nonuniform sampling*. This is because the samples of u and y do not appear in (1) in a regular manner (see Section 3.1).

This problem has been attacked recently (Feichtinger and Gröchenig, 1993; Marvasti, 1993; Sandberg, 1994), but still remains to be solved satisfactorily in general case. Especially in the N-D case there is a need of results similar to 1-D case. However, it is possible to provide at least the existence conditions for function recovery with any degree of accuracy from a sufficiently large *finite* number of irregular samples making in this way the application of neural networks plausible.

Our problem is how to reconstruct the multi-dimensional function $\tilde{f}(\lambda)$ from its nonuniformly spread samples $\{\tilde{f}(\lambda_k)\}$. A basic difference from the uniform sampling

is that there is no guarantee of the uniqueness of a band-limited signal reconstruction from arbitrary nonuniform samples, even if the average sampling rate is equal to the Nyquist rate. Therefore, there arises an issue of choosing the set $\{\lambda_k\}$ in such a way that the existence of a unique solution is guaranteed. A set of sampling instances that assures unique reconstruction is called a *sampling set*. The conditions for a set of samples to be a sampling set usually depend on the method used. In our method (Section 4.2) it is possible to establish appropriate conditions which are closely related to the persistency of excitation issue (see (Dzieliński *et al.*, 1996) for detailed explanation).

If $\{\lambda_k\}$ is a sampling set, we may now come to the problem of the actual function reconstruction. There are a few methods for the nonuniformly sampled N-D function reconstruction; however, most of them are tailored to specific tasks for which they had been designed. The common methodology for many of them is the Lagrange Interpolation Theory.

4. Practical Methods of N-D Function Reconstruction

The Lagrange interpolation provides a general interpolation framework for any sampling scheme. General as it is, the Lagrange formula does not, however, equip us with practical tools for the reconstruction of nonuniformly sampled N-D functions. In this section, we shall present two more practically orientated approaches. First, we shall discuss a method due to Kim and Bose (Kim and Bose, 1990), which provided us with the preliminary idea of transformation between the space and frequency domains. Then we present a sketch of our new algorithm.

4.1. Kim and Bose Method

The method introduced by Kim and Bose (Kim and Bose, 1990) allows us to interpolate over uniformly-spaced DFT points given nonuniformly sampled values of the function. Exploiting the aliasing relationship in the DFT (sampled frequency) domain rather than in the continuous Fourier transform domain, we can exactly reconstruct the discrete function subject to a band-limited discrete spectrum (harmonically-limited function). The essence of the approach is interpolation over the *uniformly*-spaced Discrete Fourier Transform (DFT) points given *nonuniformly* sampled values of the function.

The relation between the K_1K_2-point 2-D inverse DFT of uniform samples of $\tilde{F}$ and the nonuniform samples values $\tilde{F}(\omega_{1,k_1}, \omega_{2,k_2})$, $k_1 = 1, 2, \ldots, K_1$, $k_2 = 1, 2, \ldots, K_2$ is given below

$$\begin{pmatrix} \tilde{F}(\omega_{1,1}, \omega_{2,1}) \\ \tilde{F}(\omega_{1,2}, \omega_{2,1}) \\ \vdots \\ \tilde{F}(\omega_{1,K_1}, \omega_{2,K_2}) \end{pmatrix} = \frac{1}{K_1 K_2} \Phi \begin{pmatrix} \tilde{f}(-\frac{K_1}{2}, -\frac{K_2}{2}) \\ \tilde{f}(-\frac{K_1}{2}+1, -\frac{K_2}{2}) \\ \vdots \\ \tilde{f}(\frac{K_1}{2}-1, \frac{K_2}{2}-1) \end{pmatrix} \tag{8}$$

where Φ is the transformation matrix given in (Kim and Bose, 1990).

This approach is especially attractive in the N-D case, as it gives a closed-form interpolation formula. It allows us first to find the matrix which relates the uniformly-spaced frequency samples to the original nonuniformly-spaced sample values of the function itself. The required uniformly-spaced interpolation is then obtained by performing the inverse DFT. However, we must notice that the matrix may not always be invertible in the N-D case, even when all the sample points are distinct. Kim and Bose (1990) give some necessary conditions for invertibility in the 2-D case together with a computational algorithm. The conditions restrict to some extent the possible irregular spread patterns of samples, as might be expected.

Having transformed the nonuniformly-spaced samples in one domain into uniformly spaced samples in the other domain, we can reconstruct the obtained function by the methods based on the Shannon Theorem and its generalisations. For the problem of uniformly spaced interpolation we may use a neural network as discussed in (Żbikowski and Dzieliński, 1995).

4.2. New Method

The problem we are solving is stated as follows. We are given a finite number $K_1 \cdot K_2 \cdot \ldots \cdot K_N$ of nonuniformly spread samples $\tilde{f}(\lambda_k)$, where $\lambda_k = (\lambda_{1,k_1}, \ldots, \lambda_{N,k_N})$, of the non-linear function $\tilde{f} = \tilde{f}(x)$, where $x = (x_1, \ldots, x_N) = (y_t, y_{t-1}, \ldots, y_{t-n+1}, u_t, u_{t-1}, \ldots, u_{t-m+1})$, i.e., $N = m + n$. We want to find the function $\tilde{f}$. From the statement of our problem we know that $\tilde{f}$ is of bounded support (space-limited). We assume that *all* the sampled values $\tilde{f}(\lambda_k)$ are given, which means that we deal with an *off-line* problem.

The main idea of our method is based on the transformation of the nonuniform sampling problem in the space domain to a *uniform* problem in the Fourier Transform domain, similarly to the Kim and Bose method. However, taking advantage of the specific features of the problem of reconstruction of f in (1), we propose a simpler approach. Since $\tilde{f}$ is of bounded support, in order to reconstruct it properly we need its Continuous Fourier Transform (CFT) $\tilde{F}$. Therefore, our solution consists of two basic steps. First, we find an approximation of the CFT on the basis of given nonuniformly sampled values of $\tilde{f}$, i.e., $\tilde{f}(\lambda_k)$. Then we find the Fourier inverse of the approximation of $\tilde{F}$ to get an approximation of $\tilde{f}$. Both steps involve approximations and we will point out the sources of them in the sequel.

Let us note that the CFT for the space-limited function $\tilde{f}$ is given by the following pair of relations

$$\tilde{F}(\omega) = \int_{\mathbb{R}^N} \tilde{f}(x) \mathrm{e}^{-j\omega \cdot x} \, \mathrm{d}x = \int_A^B \tilde{f}(x) \mathrm{e}^{-j\omega \cdot x} \, \mathrm{d}x \tag{9}$$

$$\tilde{f}(x) = \left(\frac{1}{2\pi}\right)^N \int_{\mathbb{R}^N} \tilde{F}(\omega) \mathrm{e}^{j\omega \cdot x} \, \mathrm{d}\omega \tag{10}$$

where $\int_A^B$ denotes the multiple definite integral $\int_a^b \cdots \int_a^b \int_c^d \cdots \int_c^d$. Integration limits a, b correspond to the y components and c, d to the u components of vector x, respectively.

In the first step of our algorithm we approximate (9) by a finite Riemann sum of the form:

$$\tilde{F}_K(\omega) = \sum_{k=0}^{K-1} \tilde{f}(\lambda_k) \mathrm{e}^{-j\omega\cdot\lambda_k} \Delta_k \tag{11}$$

Here $\sum_{k=0}^{K-1}$ stands for the multiple summation $\sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \cdots \sum_{k_N=0}^{K_N-1}$ with $k = (k_1, \ldots, k_N)$ and $K = (K_1, \ldots, K_N) \in \mathbb{N}^N$; the K_i^S are related to the observation horizon time of system (1). Also, $\lambda_k = (\lambda_{1,k_1}, \ldots, \lambda_{N,k_N})$ is the vector of the k-th sample from $D \subset \mathbb{R}^N$ (see (1) and Section 3.1) and $\Delta_k = \prod_{i=1}^{N} \Delta_{i,k_i} = \Delta_{1,k_1} \cdot \Delta_{2,k_2} \cdot \ldots \cdot \Delta_{N,k_N}$, where $\Delta_{i,k_i} = \lambda_{i,k_i+1} - \lambda_{i,k_i}$. Also $\lambda_{i,k_i+1} > \lambda_{i,k_i}$ for all $k_i = 0, 1, \ldots, K_i - 1$ and $i = 1, 2, \ldots, n, n+1 \ldots, N$, where

$$\lambda_{i,K_i} = \begin{cases} b & \text{for } i = 1, \ldots, n \\ d & \text{for } i = n+1, \ldots, N \end{cases}$$

Furthermore $\omega = (\omega_1, \ldots, \omega_N) \in \mathbb{R}^N$ and finally, $\omega\cdot\lambda_k$ is the Euclidean inner product $\sum_{i=1}^{N} \omega_i \lambda_{i,k_i}$. Let us note that (11) bears a close resemblance to the DCFT of f. In the case of a space-limited $\tilde{f}$ such a DCFT is naturally truncated, i.e. it equals the Truncated DCFT.

The core issue of the proposed algorithm is how to use (11) to get the best approximation of $\tilde{f}$. There seems to be several natural moves possible. However, some of them do not yield a satisfactory answer to our problem. We cannot invert (11) directly, because in this way we get the sampled values $\tilde{f}(\lambda_k)$ and not the function of continuous arguments we are looking for. Another problem is how to define such an inverse transform in this case. Yet another possibility is to convolve (11) with some other function of ω, say $G = G(\omega)$ and then perform an inverse transform. In this case we also face the same problem of inverse transform definition. Moreover, since $\tilde{f}$ is not band-limited, this inverse would not result in the reconstruction of $\tilde{f}$. This problem would remain even if we knew how to construct G for a band-limited $\tilde{f}$.

Therefore, we propose to sample $\tilde{F}_K$ of (11) and use these sampled values as a basis for further approximation of $\tilde{f}$. The crucial observation is that we may choose the sample locations arbitrarily. The simplest choice is to sample $\tilde{F}_K$ *uniformly*. The sampling intervals on all N axes are provided by the N-D Sampling Theorem (see Section 3.2). Given the specific shape of the support D of $\tilde{f}$ we are able to calculate the appropriate sampling intervals for $\tilde{F}$ from (6), where the w_j's represent the space extent of $\tilde{f}$ and the v_i's give the required sampling intervals of $\tilde{F}$. Choosing the hyperrectangular sampling geometry for $\tilde{F}$, we obtain *unique* values for sampling intervals assuming the same geometry for the locations of repetitive parts of D.

Let us discuss this issue in detail for a simple 2-D example.

Example 1. In this case, $m = n = 1$, $N = m + n = 2$ and the NARMA model has the form

$$y_{t+1} = f(y_t, u_t)$$

where $y \in [a, b]$ and $u \in [c, d]$. Set $x = (x_1, x_2) = (y_t, u_t)$.

For the rectangular sampling geometry we obtain from (6)

$$v_1 \cdot w_1 = 2\pi \qquad \text{and} \qquad v_2 \cdot w_2 = 2\pi$$

where $v_1 = [\omega_{S_1}, 0]^T$, $v_2 = [0, \omega_{S_2}]^T$, $w_1 = [\max\{|a|, |b|\}, 0]^T$ and $w_2 = [0, \max\{|c|, |d|\}]^T$. In this way, we get the desired sampling intervals ω_{S_1} and ω_{S_2} in the form

$$\omega_{S_1} = \frac{2\pi}{\max\{|a|, |b|\}} \qquad \text{and} \qquad \omega_{S_2} = \frac{2\pi}{\max\{|c|, |d|\}}$$

This is not the case for any sampling geometry. If either of the geometries is not hyperrectangular, then our result would not be unique. ■

The result for the 2-D case may be generalised to $N = m + n$ dimensions

$$\omega_{S_i} = \begin{cases} \dfrac{2\pi}{\max\{|a|, |b|\}}, & \text{for } i = 1, \ldots, n \\ \dfrac{2\pi}{\max\{|c|, |d|\}}, & \text{for } i = n+1, \ldots, N \end{cases} \tag{12}$$

Sampling the DCFT gives us the following representation of $\tilde{F}_K$:

$$\tilde{F}_K(\omega) = \sum_{l \in \mathbb{Z}^N} \tilde{F}_K(\omega_S^l) \operatorname{sinc}(\omega - \omega_S^l) \tag{13}$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_N)$, $\omega_S^l = (l_1\omega_{S_1}, l_2\omega_{S_2}, \ldots, l_N\omega_{S_N})$, $l = (l_1, \ldots, l_N)$ and $\operatorname{sinc}(\omega - \omega_S^l) = \prod_{i=1}^N \operatorname{sinc}(\omega_i - l_i\omega_{S_i}) = \prod_{i=1}^N \dfrac{\sin(\omega_i - l_i\omega_{S_i})}{\omega_i - l_i\omega_{S_i}}$.

Let us note that the sampled version of (11) is equivalent to the following generalised Discrete Fourier Transform:

$$\tilde{F}_K(\omega_S^l) = \sum_{k=0}^{K-1} \tilde{f}(\lambda_k) \mathrm{e}^{-j\omega_S^l \cdot \lambda_k} \Delta_k \tag{14}$$

The inverse Fourier Transform of (13) would give us the sought $\tilde{f}$ in the form of an infinite linear combination of rectangular windowing functions. The theoretical background for such a reconstruction is provided by Theorem 1. However, from the practical point of view we are not able to use the infinite number of samples. Moreover, this approximation is useless for control applications, as rectangular windowing functions result in an awkward representation. Our problem is to reconstruct $\tilde{F}_K$ in its entirety out of a given *finite* number of values of $\tilde{F}_K(\omega_S^l)$ and obtain a useful

representation of the dynamic system, i.e., $\tilde{f}$. Therefore, we have to introduce further approximation steps. Firstly, we will approximate sinc functions in (13) by some other functions of similar (but smooth) Fourier inverses. Gaussian functions seem to be suitable in this context. Secondly, we use only a finite number of terms in the approximation of (13). The adequacy of acting this way is ensured by the so-called "$2X\Omega$ Theorem" (see (Dzieliński and Żbikowski, 1996; Slepian, 1976) for details). This result allows us to use a finite linear combination of non-linear functions in order to approximate the given function with a desired accuracy

$$\int_{-X/2}^{X/2} \left[f(x) - \sum_{l=0}^{K-1} a_l \phi_l(x) \right]^2 \mathrm{d}x < \varepsilon \tag{15}$$

where the number of terms in (15) K depends on the accuracy ε. The functions used in (15) have to be Fourier transformable, i.e., belong to $L^1(\mathbb{R}^{m+n}) \cap L^2(\mathbb{R}^{m+n})$. Taking this into account we end up with the following approximate representation of $\tilde{F}_K$ of (11):

$$\tilde{F}_K^a(\omega) = \sum_{l=-K/2}^{K/2-1} \tilde{F}_K(\omega_S^l) \mathrm{e}^{-\frac{\|\omega-\omega_S^l-c_l\|^2}{\sigma_l^2}} \tag{16}$$

valid for K even, where $\frac{\|\omega-\omega_S^l-c_l\|^2}{\sigma_l^2} = \frac{\|\omega_1-l_1\omega_{S_1}-c_{l_1}\|^2}{\sigma_{l_1}^2} + \ldots + \frac{\|\omega_N-l_N\omega_{S_N}-c_{l_N}\|^2}{\sigma_{l_N}^2}$. For K odd the summation indices in (16) need a straightforward reformulation. Taking the inverse Fourier Transform of (16), we obtain the following approximate representation of $\tilde{f}$

$$\begin{aligned}
\tilde{f}^a(x) &= \left(\frac{1}{2\pi}\right)^N \int_{\mathbb{R}^N} \tilde{F}_K^a(\omega) \mathrm{e}^{j\omega\cdot x} \,\mathrm{d}\omega \\
&= \left(\frac{1}{2\pi}\right)^N \int_{\mathbb{R}^N} \sum_{l=-K/2}^{K/2-1} \tilde{F}_K(\omega_S^l) \mathrm{e}^{-\frac{\|\omega-\omega_S^l-c_l\|^2}{\sigma_l^2}} \mathrm{e}^{j\omega\cdot x} \,\mathrm{d}\omega \\
&= \left(\frac{1}{2\pi}\right)^N \sum_{l=-K/2}^{K/2-1} \tilde{F}_K(\omega_S^l) \int_{\mathbb{R}^N} \mathrm{e}^{-\left(\frac{\|\omega-\omega_S^l-c_l\|^2}{\sigma_l^2} + j\omega\cdot x\right)} \,\mathrm{d}\omega \\
&= \left(\frac{1}{2\sqrt{\pi}}\right)^N \sum_{l=-K/2}^{K/2-1} \tilde{F}_K(\omega_S^l) \sigma \mathrm{e}^{j(\omega_S^l + c_l)} \mathrm{e}^{-\frac{x^2\sigma_l^2}{4}}
\end{aligned} \tag{17}$$

where $\sigma_l = (\sigma_{l_1}, \sigma_{l_2}, \ldots, \sigma_{l_N})$, $\sigma = \sigma_{l_1} \cdot \sigma_{l_2} \cdot \ldots \cdot \sigma_{l_N}$, $x^2\sigma_l^2 = x_1^2\sigma_{l_1}^2 + \ldots + x_N^2\sigma_{l_N}^2$ and $\omega_S^l + c_l = l_1\omega_{S_1} + c_{l_1} + \cdots + l_N\omega_{S_N} + c_{l_N}$.

Combining this with the existence results from Neural Approximation Theory (see (Żbikowski and Dzieliński, 1995) and references therein), we may note that (17)

gives a neural network architecture with Gaussian RBFs as network nodes. However, the function reconstructed from (17) may take complex values. This means that we have to implement an additional step of 'phase discrimination' in our procedure. When the phase value of the complex number (reconstructed function value) lies in the interval $(-\pi/2, \pi/2)$, we assume the phase to be equal to 0 (i.e., the value of the complex number is equal to its modulus); when it lies in the interval $(\pi/2, 3/2\pi)$, we assume the phase to be equal to π (the value of the complex number is equal to its modulus with negative sign).

4.2.1. Algorithm of the Method

The summary of the algorithm is as follows:

1. From *nonuniform* samples $\lambda_k, \tilde{f}(\lambda_k)$, $0 \leq k_i \leq K_i - 1$, $K_i \in \mathbb{N}$, $i = 1, \ldots, N$ compute the approximate Fourier Transform representation: formula (11).
2. Sample $\tilde{F}_K(\omega)$ *uniformly* with the Nyquist frequency: formula (12).
3. Reconstruct an approximation of $\tilde{F}_K(\omega)$ out of a given finite number of values of $\tilde{F}_K(\omega_S^l)$ using (16) (to replace sinc functions used in (13) use RBFs or other functions of similar spectral properties).
4. Invert (16) according to (17) and discriminate the phase.

4.2.2. Advantages of the Method

The main advantage of this algorithm is its relative simplicity. By an appropriate choice of basis functions in the reconstruction formula (16) we may obtain a neural network approximation of the NARMA model (1). This serves as a basis to further control applications of this method (Section 5).

Let us now recap the advantages of the proposed algorithm:

1. *Tailored to the real-world data from input-output discrete-time models.* The data entering the algorithm are exactly the same data we measure as inputs and outputs of the plant;
2. *Computationally simple.* In comparison with many function approximation and reconstruction algorithms it requires only simple transformations and computations. As opposed to other approaches, this algorithm neither involves matrix inversion (like the Kim and Bose method), nor iterations (the Sandberg method);
3. *Wide applicability—mild assumptions.* The only assumption on the given function is that both its direct and inverse Fourier transform exist;
4. *Flexible neural implementation.* In order to approximate the interpolation filter of (13), we may use several neural architectures. The use of Gaussian RBFs is justified by their 'nice' properties. However, other basis functions may be considered;

5. *Usefulness for control purposes.* Appropriate choice of basis functions in (13) ensures the control applicability of the reconstructed function and especially its smoothness and accuracy.

4.2.3. Comparison with the Sanner and Slotine Approach

We now focus on the basic differences between our and the Sanner and Slotine approach (Sanner and Slotine, 1992) which was a primary inspiration for our research.

- Our approach is based on the assumption of the availability of the sampled values of the function f spread nonuniformly, while Sanner and Slotine adopted the simplified view of uniform samples locations;
- We consider a discrete-time input-output models, while they dealt with a continuous-time setting;
- In our approach the function to be reconstructed is assumed to be space-limited, which results in very mild assumptions on f. Sanner and Slotine required f to be band-limited, which means very high smoothness;
- In our method the reconstruction is done in the frequency domain, while in theirs in the spatial domain.

4.2.4. Extensions and Further Research

The method presented seems to be a promising tool in the area of non-linear dynamic systems modelling. The version discussed in this paper is based on the assumption of the availability of all samples of the function f. This means it is *off-line* in character. This assumption has a direct influence on the definition of the Riemann sum in (11). To define the intervals Δ_{i,k_i} we need all the λ_{i,k_i}'s. Let us note that the order in which these values appear in the model does not correspond to the order in which they are summed up. This means we re-number the data in comparison to their natural indexing. While this is not a problem in the *off-line* approach, it may cause some difficulties when trying to construct an *on-line* version of the method.

However, it is possible to reformulate this method in a recursive manner, i.e., adapt it to account for incoming data *on-line*. The core issue is to decide how to define the summation pattern similar to the one given by (11). The most natural way is to subdivide the region in which $\tilde{f}$ is non-zero into an increasing number of intervals along with incoming data. Thus, we start with 2^N subregions obtained from the initial values $\lambda_0 = (y_{t_0}, y_{t_0-1}, \ldots, y_{t_0-n+1}, u_{t_0}, u_{t_0-1}, \ldots, u_{t_0-m+1})$, then after the arrival of the first measured value of the function $\tilde{f}$ the number of subdivisions increases to 3^N etc. At each step we are able to evaluate the approximate value of (11) in the form

$$\tilde{F}_i(\omega) = \sum_{k=0}^{i} \tilde{f}(\lambda_k) \mathrm{e}^{-j\omega\cdot\lambda_k} \Delta_k, \qquad i = 0, 1, \ldots, K-1 \tag{18}$$

where Δ_{i,k_i}'s in Δ_k are computed on-line. In the first instance these are only two subintervals along each axis obtained from λ_0. They are of the form $\{[a, y_{t_0}], [y_{t_0}, b]\}$, $\{[a, y_{t_0-1}], [y_{t_0-1}, b]\}$, ..., $\{[a, y_{t_0-n+1}], [y_{t_0-n+1}, b]\}$, $\{[c, u_{t_0}], [u_{t_0}, d]\}$, $\{[c, u_{t_0-1}], [u_{t_0-1}, d]\}$, ... $\{[c, u_{t_0-m+1}], [u_{t_0-m+1}, d]\}$. Next, in each of these pairs of subintervals one of them (it depends on whether the value of the incoming sample lies in the first or second subinterval of the pair) is further subdivided giving the triple of subintervals along each axis and so on. This proceeds so long as new measurements arrive. Finally, after all $K_1 \cdot K_2 \cdot \ldots \cdot K_N$ samples have arrived, we obtain the same $\tilde{F}_K(\omega)$ as with (11). The on-line method based on the above approach is the subject of research in progress.

Some other interesting details of the method are also being investigated. The problem of determining a function of compact support from the values of its Fourier transform on a finite segment is linked to the problems of *analytic continuation* which is an ill-posed problem in the sense of Hadamard. The solution to such a problem is not unique. However, it is possible to find an approximate solution with a small error by the universal methods of regularisation (see Tikhonov and Arsenin, 1977).

5. Relevance for Neurocontrol

The neural model of a dynamic system obtained by the nonuniform N-D sampling approach was devised with control purposes in mind. It is, however, approximate and it is not straightforward to quantify precisely its accuracy.

In this section we assess the existing feedback linear based control from the point of view of applicability of our neural model. Next, we present an overview of a novel approach to nonlinear control design—BIBO redesign—which is aimed at designing a stable control system based on scarce knowledge about the plant model. This is especially relevant to neural models which are not known accurately, as is inevitable if we only have raw input-output data and no *a priori* knowledge of the NARMA model, as in our demanding problem setting.

5.1. Feedback Linearisation Based Control

Usual approaches to control of non-linear systems need precise knowledge of system non-linearities. This is especially the case when feedback linearisation is used (Isidori, 1989). The essence of this idea is twofold. First, we transform a state-space model of the plant into new coordinates and then we cancel the non-linearities (fully or partially) by non-linear feedback. This procedure is obviously impossible without the exact knowledge of the plant non-linearities. One way is to represent the plant non-linearities by neural network's models discussed in previous sections. The method presented here is simply a neural implementation of the Chen-Khalil approach (Chen and Khalil, 1992), using our neural models in (20) below.

Let us consider an adaptive control problem for a SISO system given in the form:

$$y_{t+1} = f(y_t, \ldots, y_{t-n+1}, u_{t-1}, \ldots, u_{t-m}) + g(y_t, \ldots, y_{t-n+1}, u_{t-1}, \ldots, u_{t-m})u_t \quad (19)$$

Note that we have u_{t-m} in (19), absent in (1). This is immaterial to the generality of considerations and is done to generate a convenient form of the formulae below.

By an appropriate choice (Isidori, 1989) of state variables for (19) we may obtain the state-space model

$$x_{t+1} = \bar{f}(x_t) + \bar{g}(x_t)u_t \tag{20}$$

Now, applying the linearising feedback control law $u_t = 1/\bar{g}(x_t)\left[-\bar{f}(x_t)+r_t\right]$ we may obtain a linearised version of the plant. This may be achieved under the following assumptions on the system under consideration: f and g are smooth (differentiable a sufficient number of times); f vanishes at the origin and g is bounded away from zero over a compact subset of $\mathbb{R}^{n+m}$; the system is minimum phase, i.e., its zero dynamics has a globally exponentially stable equilibrium point at the origin; the non-linear functions f and g can be exactly represented by the neural networks $\hat{f}$ and $\hat{g}$ on a compact subset of $\mathbb{R}^{n+m}$ in the way discussed in previous sections.

Provided that our neural models represent plant non-linearities precisely, the control law (20) can exactly cancel the model non-linear dynamics giving this way the desired output signal r_t. This approach requires the exact knowledge of the plant dynamics, which is particularly difficult in the nonlinear, discrete-time context. Therefore it is not well-suited to neural networks based models which are approximations of the real plant. The modelling error that always exists may destroy the stability properties of the system. The disadvantages can be avoided when applying the BIBO redesign techniques.

5.2. BIBO Redesign

Trying to apply the presented method of nonlinear systems modelling with neural networks to control, we have to face one critical issue. The neural networks modelling of the NARMA systems gives only an approximate model. The accuracy of the right-hand side of such a model may be described by the following lower and upper bounds

$$f_1(y_k,\ldots,y_{k-n+1}) \leq f(y_k,\ldots,y_{k-n+1};u_k,\ldots u_{k-m+1}) \leq f_2(y_k,\ldots,y_{k-n+1}) \tag{21}$$

The question is how to design a controller for this inaccurate model so that the closed-loop system would remain stable.

In this section we sketch the BIBO stable design of neural control systems. This methodology is described in detail in (Żbikowski and Dzieliński, 1996b). Applying the results from the qualitative theory of difference equations (Agarwal, 1989) we introduced a novel design methodology of BIBO stable control systems based on the neural approximation of the NARMA model of nonlinear plant (Żbikowski and Dzieliński, 1996b). This has been achieved by difference inequalities methods (Agarwal, 1989).

The first issue is the adequacy of an *approximate* (due to modelling errors) NARMA model of (1) for control purposes. In other words, we would like to know if the neural NARMA representation of the real system (1) would behave in the same way as the system itself, when influenced by the control signal.

From the mathematical point of view this question is formulated as follows. Consider the true NARMA model and a function constituting an upper bound on the norm of the modelling error, i.e., the difference between the approximate NARMA model and real system (1). This function should bound the norm uniformly in u (for all admissible control signals u). Is the norm of the difference between the solutions of the approximate NARMA model and the real system (1) also bounded uniformly with respect to the control signal u? A positive answer to this question is very important from the control design viewpoint. It means that applying a specific control signal to the approximate model would cause similar (in the sense of the norm) behaviour of the real system. This applies to the stability analysis as well, because if we are able to prove stability of the control system with the model, we will prove in this way stability of the control system with the real plant. We can solve this problem using the differential inequalities given in (Pachpatte, 1970); we have described the technical details of the approach in (Żbikowski and Dzieliński, 1996b).

Let us note very important consequences of this result for the purposes of modelling and control of dynamic systems. If we are able to find a plant model being close enough to the real plant uniformly on the set of admissible control signals $\mathcal{U}$, then the discrepancy between the solutions of the model and the plant is bounded. This is exactly the case in the neural modelling of the NARMA systems described in Section 4.2, where we obtain a neural approximation of the real plant. This means, in particular, that a control signal stabilising such a model would also stabilise the real plant. However, for some classes of nonlinearities the set of admissible controls $\mathcal{U}$ may be restricted and we may be able to establish stability for a subset of $\mathcal{U}$ only. On the other hand, if we were able to prove the boundedness of the output for any u, we would obtain global results. This is not easy in general.

Using this approach it is possible (but non-trivial) to prove BIBO stability for the NARMA model (1) with its right-hand side not exactly known. This result reflects the case of neural network identification of the model, where the obtained right-hand side is known approximately.

The proposed methodology allows proving BIBO stability of the closed-loop system provided its approximate NARMA model and its error bounds are known (nonlinear) functions. The approximate NARMA model is obtained by the N-D sampling neural network approach. Our design aim is to find such control signals (design such a controller) for which we are able to prove boundedness of solutions of difference equations corresponding to the bounds of the NARMA model.

The main result of this methodology may be summarised by the following theorem.

Theorem 2 (BIBO redesign). *Consider an approximate NARMA model f^* of nonlinear control system (1). If there exist two functions $f_1(y(t),\ldots,y(t-n+1))$ and $f_2(y(t),\ldots,y(t-n+1))$ forming uniformly in u (for all $u \in \mathcal{U}$) a lower and upper bound of $f^*(y(t),\ldots,y(t-n+1);u(t),\ldots,u(t-m+1))$, and such that the solutions of difference equations with f_1 and f_2 as their right-hand sides are bounded for bounded initial conditions, then the NARMA model and (1) are BIBO stable.*

The proof of this theorem may be found in (Żbikowski and Dzieliński, 1996b).

6. Conclusions

A synergy of Fourier Analysis and nonuniform multi-dimensional sampling resulting in a novel approach to the NARMA model identification was considered for neurocontrol with feedforward networks. The engineering formulation of the problem, allowing only discrete-time input-output data and little *a priori* knowledge about the discrete-time model, has made the setting particularly demanding.

The paper concentrated on theoretical and practical control aspects of the multi-dimensional nonuniform sampling approach, which was shown to be a natural setting for modelling of NARMA systems in the context of neurocontrol. A proper tool for this methodology involves multi-dimensional Fourier Analysis. The arising fundamental problems of: function extension, approximation errors, nonuniformity of samples were addressed. The space-limited solution was shown to be the most promising. This led to a harmonically-limited N-D Fourier transform, which enabled its reconstruction in the multi-dimensional frequency domain via feedforward neural networks.

A novel method for the non-linear function approximate interpolation from a finite set of its irregular samples was proposed. It was analysed in some detail and its advantages: computational simplicity, wide applicability, flexible neural implementation and usefulness for control purposes were pointed out. A comparison with other reconstruction methods for retrieving non-linear N-D functions was given and the relevance to neurocontrol was also discussed.

Function interpolation with the proposed method allows theoretically the use of linearising control for affine NARMA systems. This strategy requires smoothness, which is not necessary for the neural reconstruction where only absolute and square integrability are needed. However, the exact knowledge of plant nonlinearities is crucial. Otherwise the modelling error influence on the system's behaviour must be taken into consideration. This concern is central to the neural BIBO redesign technique proposed.

Acknowledgements

Work supported by ESPRIT III Basic Research Project No. 8039 *Neural Adaptive Control Technology*.

References

Åström K.J. and Wittenmark B. (1990): *Computer-Controlled Systems, 2nd Edition.* — Englewood Cliffs: Prentice-Hall.

Agarwal R.P (1989): *Difference Equations and Inequalities. Theory, Methods and Applications.* — New York, Basel, Hong Kong: Marcel Dekker, Inc.

Ash J.M. (1976): *Multiple trigonometric series*, In: Studies in Harmonic Analysis (Ash J.M., Ed.). — MAA Studies in Mathematics, Mathematical Association of America, Washington, pp.76–96.

Chen F.C. and Khalil H.K. (1992): *Adaptive control of nonlinear systems using neural networks.* — Int. J. Control, v.55, No.6, pp.1299-1317.

Chen S. and Billings S.A. (1989): *Representation of non-linear systems: the NARMAX model.* — Int. J. Control, v.49, No.3, pp.1013–1032.

Dzieliński A. and Żbikowski R. (1995a): *Feedforward neural networks: n-D systems theoretic aspects.* — Proc. 3rd European Control Conference, Rome, Italy, v.2, pp.1595–1600.

Dzieliński A. and Żbikowski R. (1995b): *Multidimensional sampling aspects of neurocontrol with feedforward networks.* — Proc. 4th Int. Conf. *Artificial Neural Networks*, ICANN'95, Cambridge, England, pp.240–244.

Dzieliński A. and Żbikowski R. (1996): *Irregular sampling approach to neurocontrol with feedforward neural networks.* — Annals of Mathematics and Artificial Intelligence, (to appear).

Dzieliński A., Żbikowski R. and Kaczorek T. (1996): *Persistency of excitiation of identification with N-D sampling neural networks.* — Proc. IEEE Int. Symp. *Industrial Electronics*, Warsaw, Poland, v.1, pp.121–126.

Feichtinger H.G. and Gröchenig K. (1993): *Theory and practice of irregular sampling*, In: Wavelets: Mathematics and Applications (Benedetto J. and Frazier M. Eds.). — CRC Press, pp.305–363.

Fiddy M.A. (1987): *The role of analyticity in image recovery*, In: Image Recovery: Theory and Application (Stark H., Ed.). — Orlando: Academic Press, pp.499–529.

Havin V. and Jöricke B. (1994): *The Uncertainty Principle in Harmonic Analysis.* — Berlin: Springer-Verlag.

Isidori A. (1989): *Nonlinear Control Systems, 2nd Edition.* — Berlin: Springer-Verlag.

Kalkkuhl J.C. and Hunt K.J. (1996): *Discrete-time neural model structures for continuous-time non-linear systems: Fundamental properties and control aspects*, In: Neural Adaptive Control Technology (Żbikowski R. and Hunt K.J., Eds.). — World Scientific series in Robotics and Intelligent Systems, v.15, Singapore, London: World Scientific, pp.3–40.

Kim S.P. and Bose N.K. (1990): *Reconstruction of 2-D bandlimited discrete signals from nonuniform samples.* — IEE Proc., Part F, v.137, No.3, pp.197–204.

Leontaritis I.J. and Billings S.A. (1985): *Input-output parametric models for non-linear systems. Part I: Deterministic non-linear systems.* — Int. J. Control, v.41, No.2, pp.303–328

Marvasti F. (1993): *Nonuniform sampling*, In: Advanced Topics in Shannon Sampling and Interpolation Theory (Marks II R.J., Ed.). — New York: Springer-Verlag, NY, pp.121–156.

Pachpatte B.G. (1970): *On some n-th order finite difference inequalities.* — Proc. National Academy of Sciences, India, Section A, v.40, No.II, pp.235–240.

Petersen D.P. and Middleton D. (1962): *Sampling and reconstruction of wave-number-limited functions in N-dimensional Euclidean spaces.* — Information and Control, v.5, No.4, pp.279–323.

Pinsky M.A., Stanton N.K. and Trapa P.E. (1993): *Fourier series of radial functions in several variables.* — J. Functional Analysis, v.116, No.1, pp.111–132.

Ronkin L.I. (1974): *Introduction to the Theory of Entire Functions of Several Variables.* — Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, v.44.

Sandberg I.W. (1994): *The reconstruction of band-limited signals from nonuniformly spaced samples.* — IEEE Trans. Circuits and Systems–I: Fundamental Theory and Applications, v.41, No.1, pp.64–66.

Sanner R.M. and Slotine J.-J.E. (1992): *Gaussian networks for direct adaptive control.* — IEEE Trans. Neural Networks, v.3, No.6, pp.837–863.

Shannon C.E. (1949): *Communication in the presence of noise.* — Proc. IRE, v.37, No.1, pp.10–21.

Slepian D. (1976): *On bandwidth.* — Proc. IEEE, v.64, No.3, pp.292–300.

Stein E.M. and Weiss G. (1971): *Introduction to Fourier Analysis on Euclidean Spaces.* — Princeton, NJ: Princeton University Press.

Tikhonov A.N. and Arsenin V.Y. (1977): *Solution of Ill-posed Problems.* — New York: John Wiley and Sons.

Żbikowski R. and Dzieliński A. (1995): *Neural approximation: A control perspective*, In: Advances in Neural Networks for Control Systems (Hunt K.J., Irwin G.R. and Warwick K., Eds.). — Advances in Industrial Control, Berlin, New York: Springer-Verlag, pp.1–26.

Żbikowski R. and Dzieliński A. (1996a): *Nonuniform sampling approach to control systems modelling with feedforward networks*, In: Neural Adaptive Control Technology, (Żbikowski R. and Hunt K.J., Eds.). — World Scientific series in Robotics and Intelligent Systems, v.15, Singapore, London: World Scientific, pp.71–112.

Żbikowski R. and Dzieliński A. (1996b): *Difference inequalities approach to BIBO stability of approximate NORMA models.* — (submitted).