IDENTIFICATION OF POLYNOMIC SYSTEMS
- A SHORT REVIEW


Research Report No.106

Department of Control Engineering
University of Sheffield
Mappin Street, Sheffield S1 3JD

January 1980
IDENTIFICATION OF POLYNOMIAL SYSTEMS

S. A. Billings

Early in the twentieth century Prechet showed that a large class of nonlinear time invariant systems can be represented by the functional series

\[ y(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \tau_2, \ldots, \tau_n) \frac{\partial^n u(t-\tau_1) \cdots \partial^n u(t-\tau_n)}{\partial \tau_1 \partial \tau_2 \cdots \partial \tau_n} \, d\tau_1 \cdots d\tau_n \]  

(1)

Volterra studied these functionals extensively and the series has become widely known as the Volterra series. To facilitate the identification of nonlinear systems Wiener used the Volterra series as a basis and applied a Gram-Schmidt orthogonalisation procedure to construct a new functional series

\[ y(t) = \sum_{n=0}^{\infty} C_n(k_n u(t)) \]  

(2)

where the functionals \( C_n \) are orthogonal to one another for a Gaussian white noise stimulus. The first few terms of the Wiener functional series are

\[ G_1[k_1, u(t)] = \int_{-\infty}^{\infty} k_1(\tau) u(t-\tau) \, d\tau \]

\[ G_2[k_2, u(t)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_2(\tau_1, \tau_2) u(t-\tau_1) u(t-\tau_2) \, d\tau_1 d\tau_2 \]

\[ -A \int_{-\infty}^{\infty} k_2(\tau_1, \tau_1) \, d\tau_1 \]  

(3)

The coefficients of the n'th degree G-functional are the coefficients of the n'th degree Hermite polynomial. Although the Wiener series is equivalent to the Volterra series, the former spans the function space more efficiently due to the orthogonality of the functionals.

Identification of a nonlinear system involves the measurement of the kernels \( k_n \) in eqn (2) and Wiener applied Cameron and Martin's idea of representing each functional term by a Fourier-Hermite series and expanded the kernels in a series of Laguerre functions \( L_m(\tau) \) as

\[ k_n(\tau_1, \tau_2, \ldots, \tau_n) = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} C_{m_1 \ldots m_n} \xi_{m_1}(\tau_1) \cdots \xi_{m_n}(\tau_n) \]  

(4)

S. A. Billings is with the Department of Control Engineering, University of Sheffield.
Laguerre functions were selected because they can be represented by a series of phase shifted electrical networks called lattice networks. Measurement of the coefficients $C_{m_1...m_n}$ is achieved by correlating the output of the unknown system $y_b(t)$ with the output $y_a(t)$ of a known system for a white Gaussian input to yield

$$C_{m_1...m_n} = \frac{1}{n! A^n} y_a(t) y_b(t)$$  \hspace{1cm} (5)

The known system is constructed such that its output for a white Gaussian input is $G_n[b_n, u(t)]$ where the kernel in the leading term is synthesised as a product of Laguerre functions.

The synthesis of a nonlinear system using Wiener's method can be visualised as a linear system with multiple outputs representing the expansion of the past of the input in terms of Laguerre polynomials, in cascade with a nonlinear no-memory system representing the kernel functions followed by a network of amplifiers and adders representing the coefficients of the Laguerre expansion.

Whilst Wiener's formulation is very elegant theoretically it is impractical and difficult to apply because of the excessive number of coefficients required. Identification of even a simple system containing a second order nonlinearity would require the evaluation of approximately $10^{10}$ coefficients.

The functional representation of nonlinear systems and Wiener's ideas have been studied by several authors notably, Barrett, Bose, Brilliant, Flake, George, Harris, Singleton and Zadeh. References to these authors contributions are contained in a recent bibliography compiled by Barrett.

An alternative method of measuring the Wiener kernels $k_n$ of a nonlinear system was developed by Lee and Schetzen using correlation techniques and a white Gaussian input process. The procedure consists of computing multidimensional correlation functions between the white Gaussian input and the system output to yield

$$k_n(\tau_1, ..., \tau_n) = \frac{1}{n! A^n} \{y(t) - \sum_{m=0}^{n-1} G_m[k_m, u(t)]u(t, \tau_m)\}u(t, \tau_1) ... u(t, \tau_n)$$  \hspace{1cm} (6)

The second term on the rhs of eqn (6) only has a value on the diagonal and is included to remove impulse functions which would otherwise appear when $\tau_1 = \tau_2 = ... = \tau_n$. Although the method removes many of the difficulties associated with the Wiener formulation the amount of computation required can still be excessive. For a single input system the $n$th order kernel must be estimated at $((n+m-1)!)/(n!(m-1)!)$ points where $m = \mu/\Delta t$, $\mu$ is the system memory and $\Delta t$ the sampling interval. Computing time therefore increases almost exponentially with the order of the kernel to be evaluated.

The estimation error associated with eqn (6) will be more severe at the diagonal points because of the presence of low order
integral terms and it has been suggested that more accurate kernel estimates at these points can be obtained by interpolation between the nondiagonal points rather than direct estimation. The method has been quite widely applied notably to biological, structural and hydrological systems. Most of the practical applications have however been restricted to quadratic systems and it has recently been suggested\cite{12} that the excessive errors associated with the diagonal kernel estimates introduce fundamental difficulties which preclude the identification of third or higher order kernels.

Solution of the identification problem based on the Volterra series involves measurement of the Volterra kernels. The common approach has been to approximate the kernels by an expansion of orthogonal functions. Methods of solution include gradient type algorithms\cite{14-18}, pattern recognition techniques\cite{19}, and other related methods\cite{20-23}.

The identification of polynomial systems using higher order spectra has been investigated by Brillinger\cite{24} who derived an asymptotically unbiased estimate of the n'th degree frequency domain transfer function. French and Butz\cite{25} developed a frequency domain method of measuring the Wiener kernels based on the FFF algorithm. The method is analogous to the Lee and Schetzen algorithm\cite{10} in the time domain but results in a considerable reduction in the computational requirement. Recently Barker and Davy\cite{26} have shown that estimates of the first two Volterra kernels can be computed using Fourier transform techniques with a pseudorandom ternary input.

Volterra series have been widely applied in the analysis of communication systems and several authors (Bedrosian and Rice, Brilliant, Barrett, Bussgang, Narayanan, Zames)\cite{9} have contributed to this area.

In the search for methods of simplifying the measurement techniques and reducing the number of data points and computations required to identify either the Volterra or Wiener kernels numerous authors have investigated quasi-white and pseudorandom inputs as an alternative to white noise. Kooper and Gyftopoulos\cite{27} first described the practical measurement of a second order Volterra kernel by cross-correlation using a ternary m-sequence. Although the identification time using these inputs was reduced by a factor of 70, compared with a Gaussian white input, anomalies appeared in the fourth order autocorrelation functions which Kooper and Gyftopoulos could not explain. Similar anomalies were also observed by Simpson\cite{28} in the fourth order autocorrelation function of pseudorandom inverse repeat sequences. Ream\cite{29} investigated these anomalies, and Barker and Pradisthayon\cite{30} showed that the non-zero values in the higher than second order autocorrelation functions of m-sequences are due to deterministic characteristics of these sequences. Barker, Obidegiou and Pradisthayon\cite{31} noted that some pseudorandom signals are more suitable than others for the measurement of the second order Volterra kernel and proposed a criterion for input selection. The selection of suitable input sequences was later studied by Kadri and Lamb\cite{32}, and Barker and Davy\cite{33} have recently derived an improved identification procedure based on ternary inputs. Kernels of order higher than second can be estimated if a compound pseudorandom input\cite{34}

\[ u(t) = \sum_{i=1}^{k} x_i(t) \]
is used. Although the anomalies associated with the higher order autocorrelation functions are avoided multilevel inputs must be employed to isolate the correlation functions associated with each kernel. However, if the system kernels are factorable all the kernels can be identified sequentially from a single level compound input when the \( x_i \)'s are independent, zero mean pseudorandom sequences and \( \phi_{x_i x_i}(\tau) = \delta(\tau), i = 1, \ldots, k \).

Recently Paymarelis has introduced CPNS (constant switching pace symmetric random signals) as an alternative to band limited Gaussian white noise and has applied these inputs to the identification of various biological systems. Although the autocorrelation functions of these signals exhibit fluctuations over the whole argument space, Paymarelis has shown that the accuracy of the second order kernel estimates using these inputs is superior to pseudorandom inputs.

Various authors have considered the identification of a restricted class of systems, notably cascade systems composed of a linear system followed by a nonlinear element in cascade with another linear system. Gardiner and Economakos have suggested methods of identifying the linear kernel associated with this system by injecting multilevel inputs and isolating the kernel outputs. Identification of the higher order kernels of these systems was considered by Webb using multilevel single frequency tests and by Sander and Williamson using tensor techniques. Many authors have investigated the identification of other systems within this class; Brown and Simpson and Power considered a feedforward system, Goldberg and Durling developed an algorithm for systems composed of a linear subsystem sandwiched between two nonlinear systems, Lawrence analysed feedback systems, Godfrey and Briggs studied processes with direction dependent responses, and many others have considered the identification of the Hammerstein, Uryson and other related model structures. Recently the separable class of random processes, which were introduced by Nuttall and studied by Balasubramanian and Atherton, West and Douc were used as a basis to formulate a unified identification procedure for most of the system structures mentioned above. Haber and Keviczky give a comprehensive summary of many nonlinear model structures.

If the algorithms for restricted classes of nonlinear systems outlined above are to be implemented it is often necessary to determine the structural form, or the type of model representation which best approximates to the process prior to the identification. Although in general this is a very complex problem, the structure of cascade systems can be determined using simple tests based on correlation analysis. Saridis and Efstatdar have investigated the characterization of the structure of unknown nonlinear systems using pattern recognition techniques based on cross-covariance functions.

A measure of the degree of nonlinearity was introduced by Rajbman using dispersion functions. Auto and cross-dispersion functions have been defined, based on the conditional mean, to detect the nonlinear relationship between signals which cannot be measured using linear correlation methods. Unfortunately dispersion functions are difficult to compute and similar information can be obtained by evaluating \( \phi_{x x}(\tau) \) which is a much simpler measure of nonlinear effects.
Algorithms for the identification of both stochastic and
deterministic bilinear systems have been proposed\textsuperscript{67}. Balakrishnan
and Bruni, Di Pillo and Koch proposed algorithms using maximum
likelihood techniques. Baheti and Mohler applied correlation
analysis in conjunction with least squares, and, Beghelli and
Guidorzi\textsuperscript{68} used a simple least squares estimator based on an input/
output expansion. Recently, Baheti, Mohler and Spang\textsuperscript{69} considered
the identification of the first two kernels in a Volterra series
expansion of bilinear systems using correlation analysis and derived
a similar algorithm based on a related integral equation. Unfortu-
nately most of these algorithms appear to be either very difficult
to implement or apply only to a simplified class of bilinear systems.

Parameter estimation methods for nonlinear systems can be
classified according to the model structure and are based on either
linear or nonlinear-in-the-parameter models. A review of the
techniques associated with the latter approach is given by Seinfeld\textsuperscript{70}.
Recent developments in this field include an algorithm which com-
penstates for uncertain model structure and external disturbances by
introducing time varying parameters\textsuperscript{71}. Estimation using linear-
in-the-parameter models has been based largely on the Hammerstein
model\textsuperscript{14-49} and discrete Volterra series\textsuperscript{14-23} as discussed above.
Other authors have considered polynomic expansions of the system
states\textsuperscript{72,73}. The Group Method of Data Handling (GMDH) has been
developed by Ivakenko\textsuperscript{74} using the principles of heuristic self-
organisation to solve complex problems with large dimensionality
and short data sequences. Data splitting algorithms, cluster anal-
ysis and the application of results from catastrophe theory to
nonlinear system identification have been discussed by Mehra\textsuperscript{75,76}.

Considerable progress has been made in the identification of
nonlinear systems over the last two decades. Properties of the
functional expansions, design of algorithms, selection and
properties of inputs and identification of specific nonlinear
model structures have been studied. There is however a distinct
lack of applications of nonlinear system identification techniques
and this undoubtedly reflects the fact that further research is
required to develop improved identification and structure detection
algorithms to reduce the number of data points and simplify the
measurement techniques.

References

3. Wiener, N.: "Nonlinear problems in random theory", Wiley,
   1958.
4. Cameron, P., Martin, W.: "The orthogonal development of non-
   linear functionals in a series of Fourier-Fermite functionals",
5. HARRIS, G.H., Lapidus, L.: "The identification of nonlinear


44. LAWRENCE, P.J.: "An iterative identification procedure for systems with a nonlinear gain in the feedback or feedforward path", 7th IFAC Congress, Finland, 1978.


