FILTER-CHAIN MODELS FOR IDENTIFICATION OF NONLINEAR DYNAMICAL SYSTEMS

K. Voigtländer and H.-H. Wilfert
Fraunhofer-Institute for Information and Data Processing (IITB)
Zeunerstr. 38, D-01069 Dresden
e-mail: voigtlae@adp-dresden.de

Abstract. The modeling of nonlinear dynamical systems is considered in this paper. It is assumed that only sampled input/output data of the system under investigation are available. The most popular way for creating a black-box-model is a common nonlinear difference-equation-approach. Some basic features of such an approach are related to the corresponding properties of a filter-chain model consisting of a linear dynamical system followed by a nonlinear readout map. The filter-chain model has some very good structural characteristics but needs to be optimized with respect to its approximation efficiency. So the construction of a suitable filter system - which enables an efficient modeling - and the construction of an adjusted nonlinear readout map from a given data set is considered.
To illustrate the relation between a proper filter selection and an efficient modeling some theoretical reflections concerning an optimal filter design for an approximation of a given nonlinear system are presented afterwards. The both discussed methods are based on a VOLTERRA-Kernel representation and a state space description of the plant and yields to an adjusted filter-chain model and a bilinear filter model respectively.

1. The nonlinear difference-equation-approach and the filter-chain model

Within this section both models are established and compared with respect to some basic features arising from its structural properties.
Concerning the assumed time-continuous character of the process it is fair to ask for an appropriate time-continuous model. To deal with the acquired sampled input-output data a corresponding time-discrete version of the continuous model is necessary. In general it is impossible to obtain an analytic discrete version of a time-continuous differential model and so a quasi continuous simulation (with a numerical integration) is necessary. To avoid such a time consuming approximation it is common practice to cancel the request for a time-continuous model and to establish a time-discrete difference-equation model structure a priori (Fig. 1). Unfortunately, once the parameterization is done the model cannot be re-mapped into a continuous form.

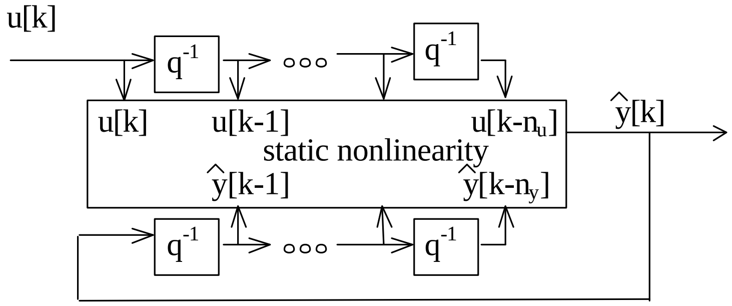

Fig. 1 Nonlinear difference-equation model

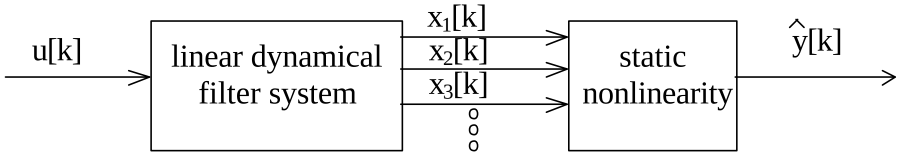

Fig. 2 General form of the nonlinear filter-chain-model

On the contrary the time continuous filter-chain model - its general form consists of a linear dynamical system followed by a nonlinear readout map - can be exactly mapped into a time discrete analytic equivalent by

assuming a hold element for the interpolation of the input signal between two samples. So the continuous model can be parametrized by using the fitting properties of the time-discrete equivalent (Fig. 2) .
Considering stability properties, aspects of parameter estimation and occasions for a structure selection within the static nonlinearity one has to admit that the filter-chain model offers excellent features (Table 1).
Especially the parameterization of the difference equation model with respect to an equation error is critical because - in contrast to linear theory - there is only a mysterious link between equation and output error. Even the output error can become infinite (caused by an instable model) while the equation error is sufficient small [5].

Table 1 Comparison of difference-equation model and filter-chain model

feature	difference-equation model	filter-chain model
relation between time continuous and time discrete model	no analytic discrete equivalent existing; numerical integration with respect to a selected interpolation element necessary	analytic calculation with respect to a selected interpolation element possible
guarantee of global asymptotic stability	an a posteriori proof generally impossible; an a priori guarantee yields to unfeasible approximation restrictions	stability of the linear dynamic part sufficient and easy to guarantee
parameterization of the nonlinearity	feasible with respect to an insufficient equation error; very difficult with respect to the output error	feasible with respect to the output error
structure selection	only possible with respect to an insufficient equation error	possible with respect to the output error
considered process class	nonlinear state space description with a unique and always defined solution	fading memory systems [1]: a unique stationary solution has to be reached asymptotically for any bounded input

While a lot of structural advances of the filter-chain model have been recognized, where is the drawback of this approach? The answer can be found very easily by checking some aspects of approximation efficiency. In the case of the difference-equation model the number of delay elements determining the input dimension of the difference model is strongly related to the order of the process considered. It has to be increased only if the output-map has not a unique inverse map. Unfortunately in the case of the filter-chain model the number of filters and so the input dimension of the readout-map must generally tend to infinity for an exact representation. But with an appropriate chosen filter system the input dimension can remain small while the accuracy of the approximation is sufficient. So the construction of a suited filter-system is a milestone within the estimation of a filter-chain model.

2. Identification of filter-chain models

The identification process involves the suitable determination of the eigenvalues of the filter and the construction of an adjusted readout map.

The construction of a suited filter-system from measured input/output data
The most popular example of a simple not adjusted filter system (a simple tapped delay line) is involved in VOLTERRA’s famous approach. Consequently a large number of delay-filters is required for a reasonably well approximation and yields to a huge number of parameter to estimate within the polynomial readout map. One can ameliorate this problem by replacing the tapped delay filters with other filters. If these filters are well adjusted a reasonable reduction of the required number of states is possible. WIENERS’s model with LAGUERRE filters or the use of KAUTZ filters [9] are examples of such well known filter systems. Recent papers discussing the problem of a suited filter parameterization for an efficient approximation of linear system [6] recommend adjustments based on the impulse response of the plant. For nonlinear systems only rules of thumb concerning the transition time of the nonlinear plant are available [4]. Obviously such recommendations are only helpful when the plant has dominant linear parts and weak nonlinearities. To deal with the crucial problem of a suitable filter parametrization in a more general way a geometrical approach is suggested now.

One possible approach for a filter-design based on measured process-data can be derived according to geometrical reflections. The filter-operators of the model perform an embedding [7] of the input series into the state-space of the model. The followed readout map provides a static mapping of this state-space to the model output. To make the model output equal to the plant output one has to establish such an embedding which allows a static mapping from state-space of the model to the plant output (apart from some measurement disturbances). Hence a plot of the plant output versus the state-space of the model is useful to judge the embedding quality because it has to prove a static dependence of the data. Fig. 3 shows embeddings with a first-order low-pass-filter. While a low and a large time constant result in a very poor embedding a proper chosen time constant yields to an appropriate arrangement of the data. For a further separation an embedding with two filters is necessary. Similar the parameters of these filters have to be tuned in such a way that the embedded data form a surface as flat as possible.

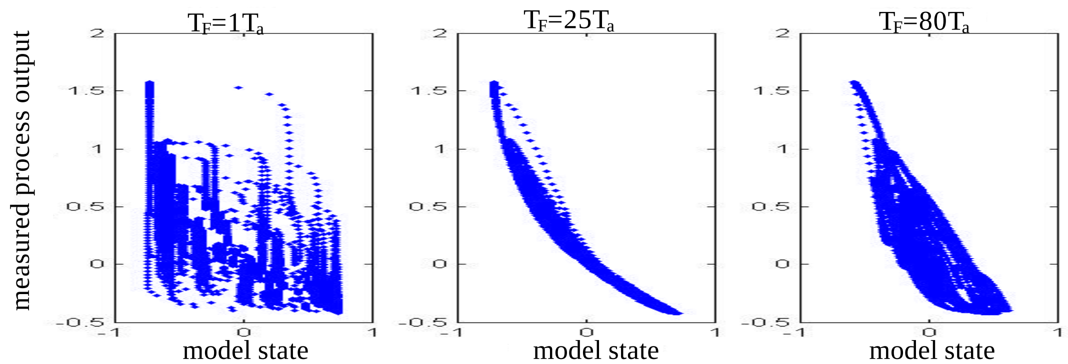

Fig. 3 Measured process output over filter-embedded input series

Summary, a model state-space must be created which yields to a unique assignment of the embedded input data to the measured process output. The optimal model filters can be found by an optimization of the filter parameters with respect to the „roughness" of embedded data. Two proposals for the measurement of this „roughness" are shown in Fig. 4. While the box-counting method provides a cumulative estimation of local variances a parametric measurement calculates the residuum of the embedded data with respect to a smooth parametric approach.

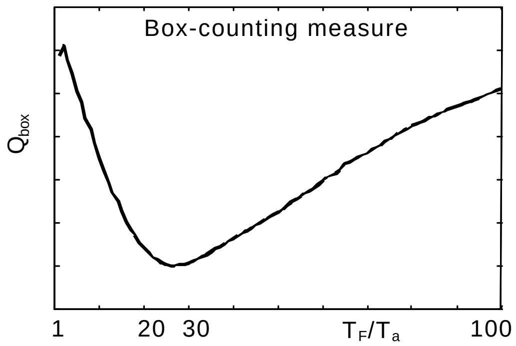

Fig. 4a
Cumulative local (100 boxes) variance

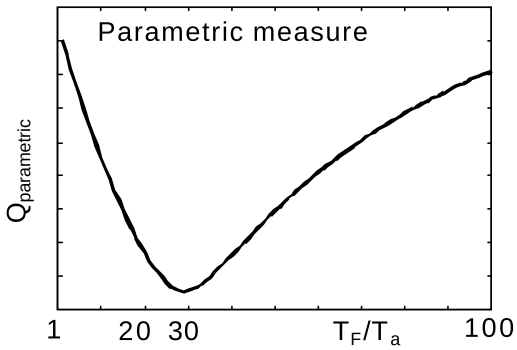

Fig. 4b
Residuum of a 3rd order polynomial approach

In praxis the demonstrated optimization might be limited to about two eigenvalues. A further suited embedding can be achieved by a repeated use of the optimized eigenvalues in the sense of a filter-chain.

The construction of an adjusted readout map

After a suited filter system has been chosen or optimized an adjusted model output (readout map) must be established. A simple but universal approach for an output map is given by

$$\hat{y}[k] = \sum_i \hat{p}_i \varphi_i\left(x[k]\right)$$

were $\{\varphi_i\}$ is a specific approximation basis. Personal preferences are crucial for the actual choice of a basis (polynoms, radial basis functions, sigmoidal functions, wavelets, splines, walsh functions, ...) and further discussion is superfluous without any a priori information of the function to be approximated.

Although the choose of a suited filter system reduces the number of required states enormously (e.g. in comparison with the classical VOLTERRA model) any nonlinear expansion will still yield to a large number of p-parameters for an accurate approximation. Unfortunately the estimated number of parameters must be limited for a good regression because of output disturbances. So an optimal collection of regressors finally to represent the model has to be estimated using structure selection methods [3]. For the stepwise choice of most important regressors from the basis-pool a forward regression algorithm [2] has been proved to be a suitable one. The algorithm is useful to establish a ranked orthogonal regressor sequence. Finally the optimal number of orthogonal regressor which has to be incorporated in the model must be determined. The recommended way of cross-validation is a special generalization of PRESS-algorithm [2] where parts of the data are sequential excluded from the parametrization process. The resulting parameter-sets will be applied on the removed data part respectively. The partitioning of the data has to be done with respect to the excitation of the process and includes disturbances. By monitoring this cross-validation error while the forward regression is in progress an optimal model complexity can be found. So it is possible to establish an adjusted readout map with respect to special signal considerations (measurement time, excitation, disturbances).

3. Derivation of suitable filter systems for the approximation of nonlinear systems

The previous given approaches for a filter design are based on measured input/output and are therefore suited for an identification process. But for a good understanding of the relation between a proper filter parametrization and an efficient modeling some theoretical reflections concerning an optimal filter design for an effective approximation of a given nonlinear system are very useful.
Within this chapter two different approaches for the construction of an optimal filter system are presented. The first method is based on a VOLTERRA kernel representation of the plant and results in a proper adjustment of a first order filter for the approximation of the plant with a filter-chain model. The second calculation yields to a bilinear filter system which eigenvalues are related to the eigenvalues of the linearized plant at the equilibrium point.

Kernel approach
Starting with the LAPLACE representation of the VOLTERRA kernels of the plant

$$K^1(s_1) \;\; ; \;\; K^2(s_1,s_2) \;\; ; \;\; K^3(s_1,s_2,s_3) \;\; ; \;\; ...$$

one can replace the complex variables s with a filter operator F of a considered first order filter F(s). An expansion of the resulting kernels as a TAYLOR series yields to

$$K^1\big(F(s_1)\big)=\sum_{i=0}^{\infty} c_i F^i(s_1) \quad ; \quad K^2\big(F(s_1),F(s_2)\big)=\sum_{i=0}^{\infty}\sum_{j=i}^{\infty} c_{ij} F^i(s_1) F^j(s_2) \quad ; \quad \ldots$$

which have a simple filter-chain representation (Fig. 5) [8]. Further the coefficients c_i,c_{ij}, ... represent the TAYLOR expansion of the required readout-map to be established.

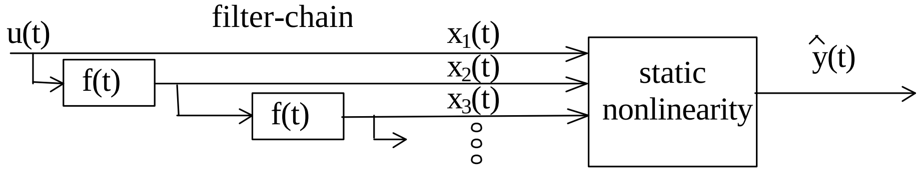

Fig. 5 Special form of the nonlinear filter-chain-model

Mainly the c-coefficients depend on the parameters of the applied filter F resp. its impulse response f(t). Hence the goal is a filter parametrization which yields to a fast decrease of the c's towards zeros. This ensures a feasible approximation quality even with only a few filters involved in the final model. Even though an analytic estimation of an adjusted filter-set requires knowledge about the VOLTERRA kernels of the system to be modeled this method yields to a good understanding of the relation between filter characteristics and approximation efficiency.

CARLEMAN approach
The second approach is based on the state-space description of the analytic linear system

$$T\,\underline{\dot{x}} = \underline{a}(\underline{x}) + \underline{b}(\underline{x})u(t)$$
$$y(t) = \underline{c}^T\,\underline{x}$$

The introduction of a new state vector $\underline{x}^{\otimes}$ which contains products of the original state up to a specific order yields to a bilinear description

$$T\,\underline{\dot{x}}^{\otimes} = A\underline{x}^{\otimes} + N\underline{x}^{\otimes}u(t) + Bu(t)$$
$$y(t) = (\underline{c}^T \quad 0 \quad \ldots)\,\underline{x}^{\otimes}$$

of the above system. From this representation it is now possible to derive a kernel representation [8] which can be realized as a junction of filter systems S_i with a linear output map (Fig. 6). The eigenvalues Λ_1 of the systems S_1 must be equal to the eigenvalues of the linearized plant at the equilibrium point. The system S_2 consists of filters equal to those of system S_1 and contains filters whose eigenvalues Λ_2 can be calculated by summing the eigenvalues of S_1 in pairs. The additional eigenvalues Λ_3 within the system S_3 result in a further pair-wise summation of the eigenvalues Λ_1 and Λ_2.

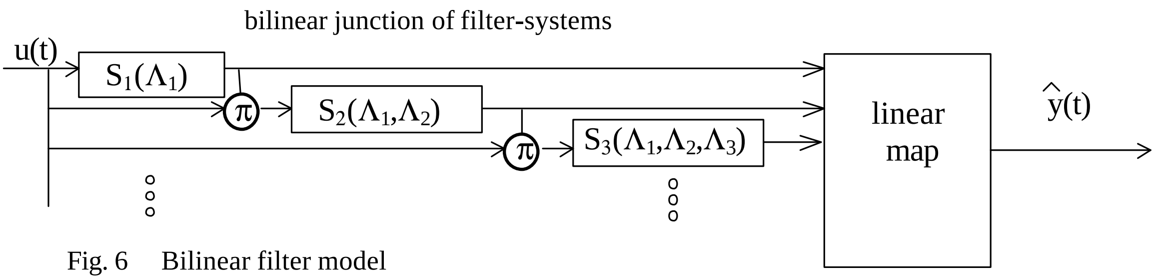

Fig. 6 Bilinear filter model

While the merely linear output map is an advantage of this model the number of state signals inflationary increases from one of filter system to the next. This is because every input signal of a filter-system has to pass each filter-element. Compared with the simple calculation of a time discrete equivalent of the filter-chain model (Fig. 5) it is still possible but time consuming to get a time discrete equivalent of the bilinear model for the use with sampled data since an exponential matrix must be calculated within every time step. But at least with this approach an exact representation of the first N kernels is possible with N filter-systems parameterized in relation to the eigenvalues of the linearized plant.

Summary

The superior structural properties of the filter-chain model have been worked out. The analysis of the approximation task yields to suggestions for a suitable filter parametrization. Considering the importance of a proper filter parametrization for a high approximation efficiency two different approaches are given for the identification of a filter-chain model from sampled input-output data.

References

[1] Boyd, S.; Chua, L.O.: Fading memory and the problem of approximating nonlinear operators with volterra series. IEEE Trans. on CAS, vol. 32 (1985), no. 11, pp. 1150-1161.

[2] Draper, N.R.; Smith, H.: Applied regression analysis, John Wiley, New York 1981.

[3] Haber, R.; Unbehauen, H.: Structure identification of nonlinear dynamic systems - a survey on input/output approaches. Automatica, vol. 26, no. 4 (1990), pp. 651-677.

[4] Kurth, J.: Identifikation nichtlinearer Systeme mit komprimierten Volterra-Reihen. Dissertation RWTH-Aachen, Fortschritt-Berichte VDI, VDI-Verlag, Düsseldorf 1995. Reihe 8, Nr. 459.

[5] Nelles, O.: On the identification with neural networks as series-parallel and parallel models. International Conference on Artificial Neural Networks (ICANN), Paris, October 1995.

[6] Oliviera a Silva, T.: On the determination of optimal pole position of Laguerre filters. IEEE Trans. Signal Process., vol. 43, no. 9, pp. 2079-2087.

[7] Packerd, N.H.; Crutchfield, J.P.; Farmer, J.D.; Shaw, R.S.: Geometry from a time series. Physical review letters, vol. 45 (1980), no. 9, pp. 712-716.

[8] Rugh, W.J.: Nonlinear system theory. The Johns Hopkins University Press, Baltimore and London 1981.
[9] Wahlberg, B.; Mäkilä, P.M.: On approximation of stable linear dynamical systems using Laguerre and Kautz functions. Automatica, vol. 32 (1996), no. 5, pp. 693-708.