CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Infoscience - École polytechnique fédérale de Lausanne



# A Local Nonlinear Model for the Approximation and Identification of a Class of Systems

Heinz Koeppl

***Abstract*—Based on Volterra series the work presents a novel local nonlinear model of a certain class of linear-analytic systems. The special form of the expressions for the Laplace-domain Volterra kernels of such systems is exploited to obtain an approximation structure that results in an appealingly simple feed-forward block structure. It comprises a composition of the linearization and the multivariate nonlinear function of the original system. Although based on Volterra series the model does not involve a truncation in the power series expansion nor in the memory depths. Compared to the exponential increase in parameters of classical memory truncated Volterra models, the structure offers an economic parametrization. The model is shown to be linear identifiable in one step if a priori information about the linearized dynamics is provided. We present simulation results for a simple nonlinear circuit showing the validity of the model.**

## I. Introduction

The proposed model is based on a Volterra series description of the forced dynamics of a nonlinear state space system around a locally stable equilibrium point. Applying an approximation step we are able to distill the corresponding Volterra series into a feedforward block arrangement involving only two types of well characterized blocks, namely the linearization of the system and its original multivariate state space function (see Fig. 4). The approximation is not based on a truncation of the series expansion nor on the truncation of the memory depth as normally done for Volterra models. Due its feed-forward structure the model inherits the stability properties of the linearization. The corresponding Volterra kernels of the model are exactly equivalent up to the second order (up to third order for anti-symmetric state space functions) to the kernels of the original nonlinear system. It is the author's opinion that the model's simple composition in terms of well characterized sub-block and its novelty, renders the model interesting on its own. Nevertheless, he position it to other approaches in the following.

Volterra series provide an elegant description of the local dynamics of a forced nonlinear system [1], [2]. Although theoretically sound, Volterra series have several shortcomings in their practical implementation, i.e., their exponential increase of parameter with order of the series permits only low such orders and short memory lengths. There are several approximations to Volterra series that circumvent this combinatorial explosion at the cost of generality, such as the Wiener model, the Hammerstein model, the linear-nonlinear-linear (LNL) model (see [3] for an overview). An advanteguous feature of Volterra models, that is not shared among all its approximations is the linear identifiability. The proposed model is able to retain this feature in a situation of practical interest. Equivalent to Volterra series, Wiener series [1] have been shown to exhibit a block cascade structure similar to that of the proposed model [4]. An important distinction though is that for the Wiener series the dimensionality of the involved prefilter bank is undetermined and can be arbitrarily large (as the result solely relies on the application of the Stone-Weierstrass theorem [4]) while for the presented model it is determined by the dimension of the original nonlinear system.

The remaining part of the work proceeds as follows. In Section II we introduce the class of nonlinear systems that we aim to approximate and give its exact Laplace-domain Volterra series representation. The proposed approximation is given in Section III, where we first exemplify the derivation for a two-dimensional system. An identification algorithm for the model is proposed in Section IV. Simulation results are given in Section V, while Section VI draws the conclusions and provides an outlook.

## II. The System and its Volterra Series

We consider the class of linear-analytic systems [5] of the particular form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{b}u \\ y &= \mathbf{c}^T\mathbf{x} \quad \text{with } \mathbf{x}(0) = \mathbf{0}, \end{aligned} \tag{1}$$

with $\mathbf{f} : \mathbb{R}^N \to \mathbb{R}^N$ assumed to be an analytic function. Although here we concentrate on the single-input-single-output (SISO) case, the main results of the paper can be generalized to the multi-input-multi-output (MIMO) case. Furthermore, we assume a local asymptotically stable equilibrium point at $\mathbf{x}_0 = \mathbf{0}$. With the introduction of an appropriate change of coordinates, systems with $\mathbf{x}_0$ not at the origin can be reduced to system (1) with $\mathbf{x}_0 = \mathbf{0}$. If the input $u$ perturbs the system (1) such that the trajectories always remain within the basin of attraction of $\mathbf{x}_0$ the system locally has fading memory [4] and a local nonlinear input-output (i/o) model in terms of a Volterra series expansion exists. To develop the expansion we make use of our above assumptions on $\mathbf{f}$ and write

$$\mathbf{f}(\mathbf{x}) = \sum_{n=1}^{\infty} \mathbf{A}_n \mathbf{x}^{\otimes n}, \tag{2}$$

where the applied Kronecker notation means the $(n-1)$-fold Kronecker product of $\mathbf{x}$ with itself and $\mathbf{A}_n$ denote the rectangular coefficient matrices of the expansion. In the following we develop our results for the space of square integrable functions $\mathrm{L}_2$, but other formulations are naturally possible. We define the Volterra series operator that maps input to state as the multi-linear operator $\mathsf{V} : u \mapsto \mathbf{x}$ with $u \in \mathrm{L}_2(\mathbb{R})$ and $\mathbf{x} \in \mathrm{L}_2^N(\mathbb{R})$, $\mathsf{V} = \sum_{n=1}^{\infty} \mathsf{V}_n$ with $\mathsf{V}_n(\alpha u) = \alpha^n \mathsf{V}_n(u)$ for every $\alpha \in \mathbb{R}$. Utilizing this definition the first line of (1) reads

$$\frac{\mathrm{d}}{\mathrm{d}t} \circ \mathsf{V}u = \mathbf{f} \circ \mathsf{V}u + \mathbf{b}u, \tag{3}$$

H. Koeppl is with the Laboratory of Nonlinear Systems, School of Computer and Communication Sciences, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland.

where "$\circ$" denotes operator composition that for clarity is also applied to functions throughout the work. We use the convention that the operators proceed their arguments. To determine the operators $\mathsf{V}_n$ we equate like powers of the input. Switching to a representation in terms of kernels of $\mathsf{V}_n$ yields the following set of equations. The Laplace-domain Volterra kernel of order $n$ associated with $\mathsf{V}_n$ are denoted as $\mathbf{V}(s_1,\ldots,s_n)$. We introduce the convention that the order of homogeneity of the kernel can be read off from the number of its arguments. For the first order we obtain from (3) $s\mathbf{V}(s)U(s)=\mathbf{A}_1\mathbf{V}(s)U(s)+\mathbf{b}U(s)$ and thus

$$\mathbf{V}(s)=(s\mathbf{I}-\mathbf{A}_1)^{-1}\mathbf{b}, \tag{4}$$

where $U(s)$ denotes the Laplace transform of $u(t)$. Equating terms with $U(s_1)U(s_2)$ gives the second-order Volterra kernel

$$\mathbf{V}(s_1,s_2)=\left[(s_1+s_2)\mathbf{I}-\mathbf{A}_1\right]^{-1}\mathbf{A}_2\mathbf{V}(s_1)\otimes\mathbf{V}(s_2), \tag{5}$$

where we applied the Kronecker formulation for multivariate Volterra operators developed in [6]. Throughout this document we use the hierarchy of operations $\otimes\rightarrow\times\rightarrow+$. For the third-order kernel we obtain

$$\begin{aligned}\mathbf{V}(s_1,s_2,s_3)&=\left[(s_1+s_2+s_3)\mathbf{I}-\mathbf{A}_1\right]^{-1}\\&\times\left\{\mathbf{A}_2\mathbf{V}(s_1,s_2)\otimes\mathbf{V}(s_3)+\mathbf{V}(s_3)\otimes\mathbf{V}(s_1,s_2)\right.\\&\left.+\,\mathbf{A}_3\mathbf{V}(s_1)\otimes\mathbf{V}(s_2)\otimes\mathbf{V}(s_3)\right\}.\end{aligned} \tag{6}$$

The general form of the Volterra kernels for system (1) for $n\geq 2$ can be obtained by applying the general composition rule for multivariate Volterra operators [6]. The result is given by the following expression, where exceptionally for ease of reading the order of the Volterra kernels is also indicated by their subscripts,

$$\begin{aligned}\mathbf{V}_n(s_1,\ldots,s_n)&=\left[(s_1+\cdots+s_n)\mathbf{I}-\mathbf{A}_1\right]^{-1}\\&\times\sum_{\ell=2}^{n}\underbrace{\sum_{m_1=1}^{n}\cdots\sum_{m_\ell=1}^{n}}_{u(\ell)=n}\mathbf{A}_\ell\bigotimes_{k=1}^{\ell}\mathbf{V}_{m_k}(s_{l(k)},\ldots,s_{u(k)}),\end{aligned} \tag{7}$$

with the upper and lower index bound functions $l(k)=l(m_1,\ldots,m_{k-1})\equiv 1+\sum_{j=1}^{k-1}m_j$ and $u(k)=u(m_1,\ldots,m_k)\equiv\sum_{j=1}^{k}m_j$, respectively, with $l(1)=1$. For later reference let us rewrite (7) by splitting the sum as

$$\begin{aligned}&\mathbf{V}_n(s_1,\ldots,s_n)=\\&\quad\left[(s_1+\cdots+s_n)\mathbf{I}-\mathbf{A}_1\right]^{-1}\mathbf{A}_n\bigotimes_{k=1}^{n}\mathbf{V}_1(s_k)\\&\quad+\left[(s_1+\cdots+s_n)\mathbf{I}-\mathbf{A}_1\right]^{-1}\\&\quad\times\sum_{\ell=2}^{n-1}\underbrace{\sum_{m_1=1}^{n}\cdots\sum_{m_\ell=1}^{n}}_{u(\ell)=n}\mathbf{A}_\ell\bigotimes_{k=1}^{\ell}\mathbf{V}_{m_k}(s_{l(k)},\ldots,s_{u(k)}).\end{aligned} \tag{8}$$

The scalar kernels for the i/o map $u\mapsto y$ with $y\in\mathrm{L}_2(\mathbb{R})$ are obtained as $H(s_1,\ldots,s_n)=\mathbf{c}^T\mathbf{V}(s_1,\ldots,s_n)$.

## III. The Local nonlinear Model

Subsequently we give an intuitive derivation of the proposed model by studying the structure of the second-order kernel for a two-dimensional system (1) in diagonal representation. Afterwards we derive the model from the general composition rule for multivariate Volterra operators for arbitrary dimensions in the original non-diagonal representation.

As the i/o behavior of the system (1) is unchanged by a nonsingular state transformation we apply the linear transform $\tilde{\mathbf{x}}=\mathbf{P}\mathbf{x}$. Assuming that $\mathbf{A}_1$ is diagonalizable we choose $\mathbf{P}$ that renders $\tilde{\mathbf{A}}_1=\mathbf{P}\mathbf{A}_1\mathbf{P}^{-1}$ diagonal, i.e., $\tilde{\mathbf{A}}_1=\mathrm{diag}(\tilde{A}_{1,11},\ldots,\tilde{A}_{1,NN})$, where the notation $\tilde{A}_{n,lm}$ means the $lm$-th element of the matrix $\tilde{\mathbf{A}}_n$. The function $\mathbf{f}(\mathbf{x})$ in the transform domain reads

$$\tilde{\mathbf{f}}(\tilde{\mathbf{x}})=\mathbf{P}\mathbf{f}(\mathbf{P}^{-1}\tilde{\mathbf{x}}). \tag{9}$$

The corresponding higher-order coefficient matrices $\tilde{\mathbf{A}}_n$ of the series expansion (2) of $\mathbf{f}(\mathbf{x})$ transform as

$$\tilde{\mathbf{A}}_n=\mathbf{P}\mathbf{A}_n\mathbf{P}^{-1}\otimes\cdots\otimes\mathbf{P}^{-1}, \tag{10}$$

where a $(n-1)$-fold Kronecker product is applied. The diagonalized representation is useful to see more easily the possible decomposition of higher-order Volterra operators in terms of interconnections of linear operators. Consider now the case of the second-order Volterra kernel $\mathbf{V}(s_1,s_2)$ in (5) of (1). Applying the diagonal representation and taking the output function with $\tilde{\mathbf{c}}=(\mathbf{P}^{-1})^T\mathbf{c}$ into account, i.e., $H(s_1,s_2)=\tilde{\mathbf{c}}^T\tilde{\mathbf{V}}(s_1,s_2)$, one can map the second-order Volterra operator to a cascade and multiplicative structure of linear systems. For example the explicit terms for the second-order kernel $H(s_1,s_2)$ in correspondence to (5) and (4) for a two-dimensional system (1) are

$$\begin{aligned}H(s_1,s_2)&=\begin{pmatrix}\tilde{c}_1\\\tilde{c}_2\end{pmatrix}^T\begin{pmatrix}\tilde{T}_{11}(s_1+s_2)&0\\0&\tilde{T}_{22}(s_1+s_2)\end{pmatrix}\\&\times\begin{pmatrix}\tilde{A}_{2,11}&\tilde{A}_{2,12}&\tilde{A}_{2,13}&\tilde{A}_{2,14}\\\tilde{A}_{2,21}&\tilde{A}_{2,22}&\tilde{A}_{2,23}&\tilde{A}_{2,24}\end{pmatrix}\begin{pmatrix}\tilde{T}_{11}(s_1)&0\\0&\tilde{T}_{22}(s_1)\end{pmatrix}\\&\otimes\begin{pmatrix}\tilde{T}_{11}(s_2)&0\\0&\tilde{T}_{22}(s_2)\end{pmatrix}\begin{pmatrix}\tilde{b}_1\\\tilde{b}_2\end{pmatrix}\otimes\begin{pmatrix}\tilde{b}_1\\\tilde{b}_2\end{pmatrix},\end{aligned} \tag{11}$$

with the definition of the diagonal transfer matrix $\tilde{\mathbf{T}}(s)\equiv(s\mathbf{I}-\tilde{\mathbf{A}}_1)^{-1}$ and with the following property for Kronecker products [7]

$$(\tilde{\mathbf{T}}(s_1)\tilde{\mathbf{b}})\otimes(\tilde{\mathbf{T}}(s_2)\tilde{\mathbf{b}})=\tilde{\mathbf{T}}(s_1)\otimes\tilde{\mathbf{T}}(s_2)\tilde{\mathbf{b}}\otimes\tilde{\mathbf{b}}.$$

The second-order i/o Volterra kernel then reads

$$\begin{aligned}H(s_1,s_2)&=\sum_{k_1=1}^{N}\sum_{k_2=1}^{N}\sum_{k_3=1}^{N}\\&\frac{\tilde{c}_{k_1}\tilde{A}_{2,k1,k_3+N(k_2-1)}\tilde{b}_{k_2}\tilde{b}_{k_3}}{(s_1+s_2-\tilde{A}_{1,k_1,k_1})(s_1-\tilde{A}_{1,k_2,k_2})(s_2-\tilde{A}_{1,k_3,k_3})}.\end{aligned} \tag{12}$$

Each summand of the second-order transfer function $H(s_1,s_2)$ in (12) is a particular instance of a cascade connection of linear filters shown in Fig. 1. Every term in (12) can be realized by a multiplicative connection of two first order low pass filters cascaded by another first-order low pass filter. The representation (12) of the Volterra kernel can be seen as the analog to the partial-fraction expansion for the

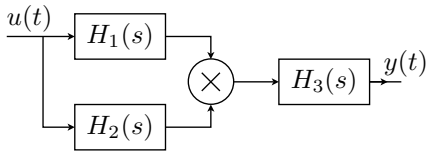


Fig. 1. General cascade structure of linear filters and quadratic nonlinearity.

transfer function of linear systems. With this observation and a rearrangement of the linear blocks we give in Fig. 2 the exact representation of the second-order transfer function $H(s_1, s_2)$ of the operator $\mathbf{c}^T \mathsf{V}_2$ for system (1) with $N = 2$.

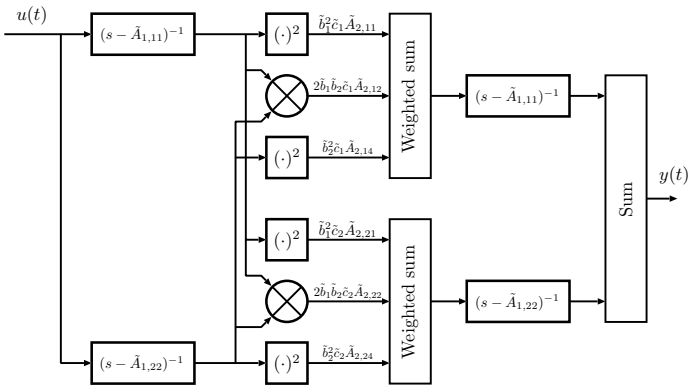


Fig. 2. Block diagram equivalent to the i/o behavior of a second-order Volterra operator of the system (1) for state space dimension $N = 2$.

Some remarks are in order. First, the system in Fig. 2 can be thought of consisting of three parts, namely the central block that realizes a general quadratic vector function of $N = 2$ arguments and the two adjacent parts comprising the diagonal linear transfer function $\tilde{\mathbf{T}}(s)$. Considering the general expression for the higher-order kernels (8), reveals that its first summand is analogous to the above discussed expressions in (12) for higher orders. For example the third-order kernel would involve a cubic vector function of $N = 2$ arguments and the same adjacent transfer functions $\tilde{\mathbf{T}}(s)$. Thus, these particular terms present in all higher-order kernels can be collectively accounted for in the same structure Fig. 2, by just replacing the purely quadratic function by a two-dimensional polynomial function. Consequently, if we separate the contributions from $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{b}}$ this polynomial function can be expressed as $\tilde{\mathbf{g}}(\tilde{\boldsymbol{\vartheta}}) \equiv \tilde{\mathbf{f}}(\tilde{\boldsymbol{\vartheta}}) - \tilde{\mathbf{A}}_1 \tilde{\boldsymbol{\vartheta}}$ with $\tilde{\boldsymbol{\vartheta}} \equiv (\tilde{\vartheta}_1, \ldots, \tilde{\vartheta}_N)^T$ the output of the first linearization block $\tilde{\mathbf{T}}(s)$ (cf. Fig. 3). To this end the approximate i/o realization for a system (1) with $N = 2$ becomes the system shown in Fig. 3, where also the first-order Volterra operator (4) is taken into account. All of

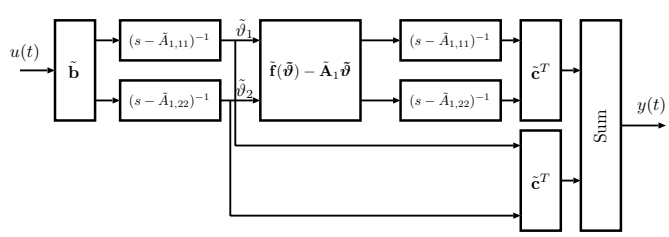


Fig. 3. Block diagram of the approximate linear-nonlinear-linear realization of the Volterra operator of the system (1), with $N = 2$ in diagonal representation.

the above considerations can also be performed in the original domain with $\mathbf{A}_1$ having general form. The same line of argumentation leads to the final approximative structure given in the original coordinates, shown for arbitrary dimension $N$ in Fig. 4. Subsequently we prove this claim.

Thus, the proposition is that the cascade operator with the composition

$$\mathsf{W} \equiv \mathsf{V}_1 + \mathsf{T} \circ \mathbf{g} \circ \mathsf{V}_1 \tag{13}$$

exhibits Volterra kernels that are identical to the first summand in (8) for $n \geq 2$ and (4) for $n = 1$ of the kernel expressions for (1). Here $\mathsf{T}$ denotes the linear MIMO operator $\mathsf{T} : \mathbf{b}u \mapsto \boldsymbol{\vartheta}$ with the kernel $\mathbf{T}(s) \equiv (s\mathbf{I} - \mathbf{A}_1)^{-1}$. Using the results of [6] we first determine the kernels $\mathbf{K}(s_1, \ldots, s_n)$ of the composition $\mathsf{K} \equiv \mathsf{T} \circ \mathbf{g}$. One obtains

$$\mathbf{K}(s_1, \ldots, s_n) = \mathbf{T}(s_1 + \cdots + s_n)\mathbf{A}_n,$$

for $n \geq 2$ and $\mathbf{K}(s) = 0$ for all $s$ by construction of $\mathbf{g}$. Utilizing this intermediate result we can compute the kernels of $\mathsf{K} \circ \mathsf{V}_1$ that in turn with (13) coincide with the kernels $\mathbf{W}(s_1, \ldots, s_n)$ of the cascade $\mathsf{W}$ for orders $n \geq 2$. Thus, they read

$$\begin{aligned}\mathbf{W}(s_1, \ldots, s_n) &= \mathbf{K}(s_1, \ldots, s_n) \bigotimes_{k=1}^{n} \mathbf{V}(s_k) \\ &= \mathbf{T}(s_1 + \cdots + s_n)\mathbf{A}_n \bigotimes_{k=1}^{n} \mathbf{V}(s_k),\end{aligned}$$

for $n \geq 2$ and $\mathbf{W}(s) = \mathbf{V}(s)$ for the linear part $n = 1$. We observe that this is identical to the first summand in (8) for $n \geq 2$ and (4) for $n = 1$. We conclude the prove. The kernels for the i/o map $u \mapsto y$ of the cascade structure are obtained as $\mathbf{c}^T \mathbf{W}(s_1, \ldots, s_n)$ for all $n$.

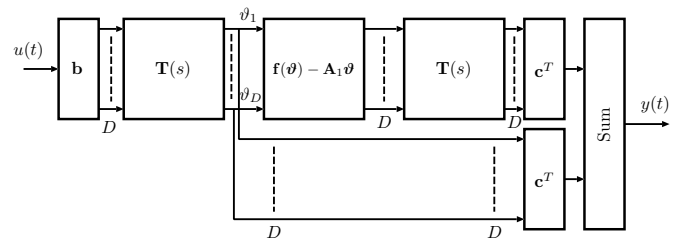


Fig. 4. Block diagram of the approximate linear-nonlinear-linear cascade realization $\mathsf{W}$ of the Volterra operator $\mathsf{V}$ for the system (1); no state transformation matrix $\mathbf{P}$ is applied; arbitrary dimension $N$.

This local model has the following intriguing property. It can be considered an augmentation of a local linear model of a nonlinear system. In fact, the lower branch of the structure in Fig. 4 is identical to the linear approximation of (1). The second parallel branch, containing the original nonlinearity of (1), gives the nonlinear correction to the linear model.

As indicated the exact equivalence between the second-order operator $\mathbf{c}^T \mathsf{V}_2$ and the cascade structure in Fig. 2 does not hold for higher orders. For orders $n \geq 3$ and $n \geq 4$ for general and anti-symmetric nonlinearities $\mathbf{f}(\mathbf{x})$ respectively, additional terms are present in the Volterra kernel representation (8) of $\mathsf{V}_n$ that can not be cast into a cascade structure (13). Confer for instance to the first summand in (6).

To derive this model we assumed complete information about the underlying system (1), while another interesting application of the model is in situations, where the model parameters have to be estimated from noisy input-output (i/o)

measurement of a partially unknown system. This application is discussed subsequently.

## IV. The Identification of the model

In many situation, only i/o measurements of a nonlinear system is available and the problem is to estimate an accurate local i/o model of it. Throughout this section we assume model matching conditions, i.e., the structure of the system to be identified exhibits the structure of the model. Studies investigating the effect of model mismatch for this model need to be done but are not subject of this paper. Due its particular structure we can distinguish two variants of the identification problem for this model. The general case comprises the situation where both the linearization $\mathbf{T}(s)$ as well as the nonlinear function $\mathbf{g}(\boldsymbol{\vartheta})$ is unknown. In the second situation we do have a priori information about the linearization and only have to identify the function $\mathbf{g}(\boldsymbol{\vartheta})$. Although, this condition seems to be ad-hoc at first, it actually corresponds to a situation encountered in practice. For instance in the generation of macromodels for large-scale weakly-nonlinear circuits, such as transceiver front-ends [8] the linear part of the system can accurately be extracted from a small-signal analysis. In the absence of a priori knowledge about the linear characteristics a sequential procedure can be applied, where first the linear part is estimated [9], [10], [11] (possibly with a smaller amplitude than nominal) and second the function $\mathbf{g}(\boldsymbol{\vartheta})$ is estimated.

In the following we detail on a particular implementation for the estimation of the function $\mathbf{g}(\boldsymbol{\vartheta})$ in the presence of information about the linearization $\mathbf{T}(s)$. We model every component of the vector function $\mathbf{g}(\boldsymbol{\vartheta})$ as a basis function network (BFN) with (cf. Fig. 5)

$$g_k(\boldsymbol{\vartheta}) = \boldsymbol{\phi}_k^T(\boldsymbol{\vartheta})\mathbf{w}_k,$$

with $\boldsymbol{\phi}_k \equiv (\,\phi_{k,1}(\boldsymbol{\vartheta}), \ldots, \phi_{k,M}(\boldsymbol{\vartheta})\,)^T$ and $\mathbf{w}_k$ denoting the column vector of parameters for component $k$. From Fig. 5 one realizes that this is not the classical regression setup as the target vector for the regression is not directly available but merely $y(t)$ is available. Therefore we have to filter all regressors with the known linear operator $\mathsf{T}$. The model output can thus be written as

$$y(t) = \mathrm{Tr}\left\{(\mathsf{T}\boldsymbol{\Theta})^T(t)\mathbf{W}\right\} + \boldsymbol{\vartheta}^T(t)\mathbf{w}_0, \tag{14}$$

with the matrix of regressors

$$\boldsymbol{\Theta}(t) \equiv (\,\boldsymbol{\phi}_1(\boldsymbol{\vartheta}(t)), \ldots, \boldsymbol{\phi}_N(\boldsymbol{\vartheta}(t))\,) \tag{15}$$

and the $M \times N$ matrix of parameters

$$\mathbf{W} \equiv (\,\mathbf{w}_1, \ldots, \mathbf{w}_N\,). \tag{16}$$

The symbol $\mathrm{Tr}(\cdot)$ denotes the trace operation on a matrix. The additional term $\boldsymbol{\vartheta}^T(t)\mathbf{w}_0$ in (14) corresponds to the linear part, i.e., the lower branch in Fig. 5. In the estimation the readout vector $\mathbf{c}$ for the linear as well as for the nonlinear branch is absorbed into the parameter $\mathbf{w}_k$. Another way to write (14), more suitable for estimation, is

$$y(t) = \boldsymbol{\psi}^T(t)\mathbf{w}, \tag{17}$$

with the $N(M+1) \times 1$ vector of filtered basis functions

$$\boldsymbol{\psi}(t) \equiv (\,\boldsymbol{\vartheta}^T(t), \mathrm{vec}^T((\mathsf{T}\boldsymbol{\Theta})(t))\,)^T \tag{18}$$

and the $N(M+1) \times 1$ vector of all BFN parameters

$$\mathbf{w} \equiv (\,\mathbf{w}_0^T, \mathrm{vec}^T(\mathbf{W})\,)^T, \tag{19}$$

where $\mathrm{vec}(\cdot)$ denotes the columwise concatenation of a matrix into a vector. The linear estimation problem then reads

$$\mathbf{y} = \boldsymbol{\Psi}\mathbf{w}, \tag{20}$$

with the vector of output measurements $\mathbf{y} \equiv (\,y(\Delta t), \ldots, y(N_s\Delta t)\,)^T$ and the matrix of filtered basis vectors $\boldsymbol{\Psi} \equiv (\,\boldsymbol{\psi}(\Delta t), \ldots, \boldsymbol{\psi}(N_s\Delta t)\,)^T$ at $N_s$ different time points. This linear, in general overdetermined problem can now be solved by any standard linear regression algorithm. For the example in Section V we deploy sparse Bayesian regression [12].

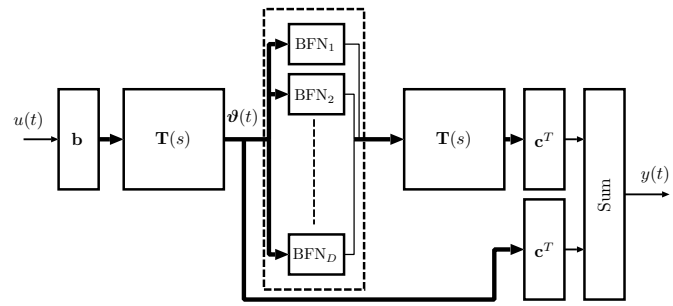


Fig. 5. Block diagram of the nonlinear cascade model with the nonlinear function realized by a set of basis function networks.

## V. Simulation Results

To show the practical validity of the proposed local nonlinear model we apply it to the simple nonlinear circuit depicted in Fig. 6. Naturally, this does not account for a full assessment of the approximation capabilities of the model but merely serves as a first prove of principle. The parameters of the circuit are chosen such that its linear characteristic matches a second-order Chebychev type I low pass filter. The system is made nonlinear by assuming that the inductor exhibits a magnetic flux characteristics $\varphi(j) = \frac{L_0}{\eta}\tanh(\eta j)$ [13, p.36]. The resulting differential equation has the form (1) with

$$\mathbf{f}(\mathbf{x}) \equiv \begin{pmatrix} -\dfrac{1}{R_{in}C}x_1 - \frac{1}{C}x_2 \\ \dfrac{\cosh^2(\eta x_2)}{L_0}(x_1 - R_l x_2) \end{pmatrix}, \tag{21}$$

and $\mathbf{b} = (\,\frac{1}{R_{in}}, 0\,)^T$, $\mathbf{c} = (\,0, R_l\,)^T$, where the components $x_1$ and $x_2$ of $\mathbf{x}$ denote the capacitor voltage $v_C(t)$ and inductor current $j_L(t)$, respectively. First, assuming complete knowledge about the system a linear approximation as well as the novel approximation is computed. Their responses are compared to the output of the nonlinear circuit by applying a broadband discrete-multi-tone (DMT) signal with $512$ carriers and phases drawn from a uniform distribution $\mathcal{U}[0, 2\pi)$. This broad-band signal was chosen in order not to bias the results due to a particular choice of input signal and to guarantee a persistent excitation [11] of the system for the identification experiment discussed below. A normalized mean square error is computed for different peak voltages of $v_{in}(t) = u(t)$, ranging from $5-50$V. The error is normalized to single out the obvious scaling of the error with the input signal amplitude. The results are given in Fig. 7. The linear as well as the nonlinear model show the natural tendency of local models to give less accurate prediction for larger amplitudes. For small

amplitudes the proposed model achieves a large improvement of up to 20dB with respect to the linear model.

In the second experiment only knowledge of the linear characteristics of the circuit is assumed. To approximate the two-dimensional nonlinear function $\mathbf{g}(\boldsymbol{\vartheta})$ the algorithm outlined in Section IV is applied. A radial basis function network is used for each of the two dimensions (although from our a priori knowledge of (21) we know that the first component of the nonlinear function $\mathbf{g}(\boldsymbol{\vartheta}) = \mathbf{f}(\boldsymbol{\vartheta}) - \mathbf{A}_1\boldsymbol{\vartheta}$ is zero). For the estimation of the network parameters with (20), the linear regression algorithm [12] is deployed. Each basis function network initially consists of spherical Gaussian kernels

$$\phi_{k,i}(\boldsymbol{\vartheta}) = \exp(\frac{1}{2\sigma^2}||\boldsymbol{\vartheta} - \mathbf{c}_i||^2) \quad \text{for} \quad k = 1, 2$$

with centers $\mathbf{c}_i$ placed on a two-dimensional equidistant grid of size $10 \times 10$ over a predetermined amplitude range of $\boldsymbol{\vartheta}(t)$, i.e., the accessible output signal of the linearization $\mathbf{T}(s)$ (cf. Fig. 5). The width of all kernels is chosen to be $\sigma = 0.7$. The model is trained with a signal of 20V peak, while the validation is done in terms of its normalized mean square generalization error $\varepsilon_g$ with different phase distributions but same peak-to-average-ratio of the validation DMT input signal over the range from $5 - 50$V. Due to the deployment of a sparse regression algorithm only a small subset of $18$ from the initial $200$ parameters remain non-zero and constitute the final model parameters. In Fig. 7 the performance of the identified model is illustrated. Note that the discussed least squares fit of the function $\mathbf{g}(\boldsymbol{\vartheta})$ can have superior performance than the analytically obtained cascade model (also shown in Fig. 7). The reason for this is that the least squares fit can compensate for the model mismatch between the generative system (1) and the model (13).

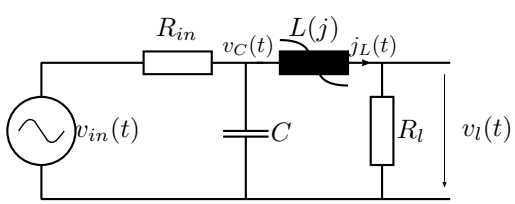


Fig. 6. Simple exemplary circuit including the inductor $L(j)$ as nonlinear component used to test the proposed approach.

## VI. Conclusion

We derived a novel local nonlinear model for a certain class of linear-analytic systems. Although based on the Volterra series representation of the i/o behavior of the original system, it alleviates some of the shortcomings of classical Volterra series modeling. Furthermore, we proposed an identification algorithm for the new model and showed first simulation results for a simple nonlinear circuit. The model gives an improvement to a linear model of up to 20dB in terms of the mean square error. This paper serves as an introduction of the novel model or approximation structure and is complete in this respect. Clearly as for any new model, many things remain to be done and the author proposes the following two important research avenues. First, to be applicable to general nonlinear circuits it has to be investigated whether the formulation can be extended to circuit equations being more general than (1), i.e., of differential-algebraic type. Paralleling the approximation error bounds for Taylor series and Volterra series [14], a further important challenge is to derive corresponding error bounds for this structure.

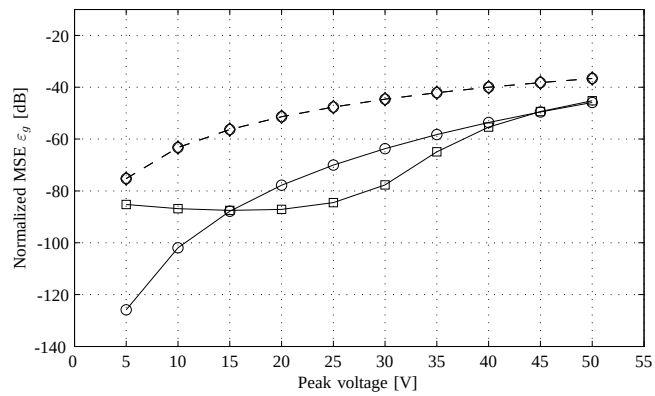


Fig. 7. Normalized mean-square error between the output of the reference circuit and the output of three different models: linear approximation (dashed), novel local nonlinear model (solid); with complete knowledge (circle) and estimation of the nonlinear function using radial basis functions (square); amplitude of the training signal is 20V.

## Acknowledgment

The author would like to thank Prof. Martin Hasler for helpful discussions. The work is supported by the Swiss National Science Foundation, grant no. 200020-117975/1

## References

[1] W. J. Rugh, *Nonlinear System Theory*. London: Johns Hopkins University Press, 1981.
[2] P. Wambacq and W. Sansen, *Distortion Analysis of Analog Integrated Circuits*. Dordrecht, The Netherlands: Kluwer Academic Publisher, 1998.
[3] O. Nelles, *Nonlinear System Identification*. Springer-Verlag, 2001.
[4] S. Boyd and L. O. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Transactions on Circuits and Systems*, vol. 32, no. 11, pp. 1150–1171, November 1985.
[5] A. Isidori, *Nonlinear control systems*, 3rd ed. London: Springer Verlag, 1995.
[6] H. Koeppl, "The composition rule for multivariate Volterra operators and its application to circuit analysis," in *IEEE International Symposium on Circuits and Systems*, New Orleans, LA, USA, May 27–30 2007, pp. 441–444.
[7] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp. 772–781, 1978.
[8] H. Koeppl and P. Singerl, "An effcient scheme for nonlinear modeling and predistortion in mixed signal systems," *IEEE Transactions on Circuits and Systems-Part II*, vol. 53, no. 12, pp. 1368–1372, 2006.
[9] P. Van Overschee and B. De Moor, *Subspace identification for linear systems*. Dordrecht: Kluwer Academic Publisher, 1996.
[10] L. Ljung, *System Identification: theory for the user*. Upper Saddle River, NJ: Prentice-Hall, 1999.
[11] J. Schoukens and R. Pintelon, *System Identification: A Frequency Domain Approach*. IEEE Press, 2001.
[12] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, June 2001.
[13] L. O. Chua, *Introduction to Nonlinear Network Theory*. McGraw-Hill, 1969.
[14] S. Boyd, L. O. Chua, and C. A. Desoer, "Analytical foundations of Volterra series," *IMA Journal of Mathematical Control and Information*, vol. 1, no. 3, pp. 243–282, 1984.