Cascade identification of nonlinear systems

S L Podvalny and E M Vasiljev
Department of Automated and Computer Systems, Voronezh State Technical University, 14, Moscow av., Voronezh, 394026, Russia
E-mail: vgtu-ait@yandex.ru

Abstract. This work solves the problem of identification of nonlinear systems the standard implementation of which are the telecommunication systems. Nonlinearity of the reception and transmitting nodes of communication channels causes the intermodulation distortion of transmitted signals. Compensation of this distortion is based on the identification of nonlinear characteristics of the abovementioned nodes and their dynamic parameters. An analysis the traditional approach to the solution of a problem of nonlinear identification was made to obtain a solution based on the classical concept of “a black box”. It was shown that using this approach to produce the concentrated model of a nonlinear system with adaptation by its output signal faces inevitably the basic restrictions caused by incommutativity of nonlinear and linear nodes of a system. To overcome this restriction the we offer the stage-by-stage composition and identification of the nonlinear model parameters that are based on the principles of a multialternative structure and functioning of composite systems i.e.: multilevelness, modularity and separation of functions. The results of a pilot study of numerous variants of this model are shown here, viz., the concentrated model and multistage model with adaptation by output signal and a multistage model with adaptation by a vector of state. High efficiency of the model with stage-by-stage identification of nonlinear system was confirmed.

1. Condition of a problem
Presently, the most urgent problem which exists in telecommunication complexes is how to provide a low level of distortions of the transmitted signals. Among the sources of such distortions are the nonlinear amplitude and phase characteristics of the transmitting and reception nodes. Nonlinearity of these characteristics leads to the emergence of intermodulation distortion (figure 1), it prevents, in particular, the use of high order quadrature modulation (OFDM modulation) which is widely used in modern telecommunication data transmission systems [1].

Compensation of a nonlinear distortion is achieved, as a rule, by the introduction of additional components to a communication channel, e.g.: the pre-distorters in transmitters and nonlinear equalizers in the receivers. These devices are constructed using the models of the transmission and reception nodes whose nonlinear characteristics should be close to the corresponding characteristics of real channels of telecommunication system [2-7].

Reproduction of such characteristics in the model is reached by the solution of the problem of identification of the initial nonlinear system parameters and the corresponding digital processing of the distorted signals with purpose of their correction [8-10].
In figure 2 the structure of nonlinear system is shown comprising, in general case, a number of $n$ cascades.

Each cascade can be presented as consistently connected with both the non-inertial link, said link having the nonlinear characteristic $f_i(x)$, which reflects the dependence of a link transfer coefficient on the signal amplitude $x(\omega)$, and with the dynamic link $w_i(j\omega) = w_i(\omega) \cdot \exp(j \cdot \varphi(\omega))$ with phase frequency characteristic $\varphi(\omega)$, $i = 1, ..., n$, which also has the nonlinear form resulting in the inconstancy of temporary shifts of the signal components in the operating band of frequencies $\omega$.

The signal $X(j\omega) = x(\omega) \cdot \exp(j \cdot \varphi(\omega))$ which is read at the input of such multicascade system undergoes the repeated nonlinear amplitude and phase distortions imposed by the parameters $f_i(x)$ and $\varphi_i(\omega)$.

The identification of unknown characteristics of initial system which is performed with the purpose of compensation of the nonlinear distortions arising in system assumes the obtaining of the estimates $\hat{f}_i(x)$ and $\hat{\varphi}_i(\omega)$.

The most widespread approach to the solution of an objective of identification is the representation of an initial nonlinear system as the model with concentrated parameters $\hat{F}(x)$, $\hat{\Phi}(\omega)$ and $\hat{W}(j\omega)$, as shown in figure 3, [11-13].
Using this representation of the system, the problem of identification is how to find some estimates $\hat{F}(x)$, $\hat{\Phi}(j\omega)$ and $\hat{W}(j\omega)$ the cumulative influence of which on the signal $X(j\omega)$ passage is equivalent to the distributed influence of the initial real system on this signal [14-16]. The criterion of quality of identification is the difference between the output signals of the real system and its model:

$$e(j\omega) = Y(j\omega) - \hat{Y}(j\omega).$$

(1)

Despite the considerable success achieved in the identification of the concentrated model parameters, the abovementioned approach to the solution of the problem has the basic shortcoming which is the matter of principle, i.e. in the concentrated parameters model the cascade structure of real system is not accounted or even allowed.

2. Solution

In the theoretical plan the nonlinear operations $f_i(x)$ and $w_i(j\omega)$, that are carried out by each cascade have no properties of a commutativity.

In fact, the attempt to obtain a concentrated representation of the specified operations (figure 2) does not provide an equivalence of the model and real data transmission channel:

$$w_i(j\omega) \cdot f_i[x_i(j\omega)] \neq f_i[w_i(j\omega) \cdot x_i(j\omega)], \quad i = 1, n.$$  \hspace{1cm} (2)

This circumstance, primarily, leads to an unjustified complication of characteristics $\hat{F}(x)$ and $\hat{\Phi}(j\omega)$ of the concentrated model whereas the initial nonlinear dependences $f_i(x)$ and $\phi_i(\omega)$ can be simple enough [17,18]; secondly, the same circumstance is an obstacle for the creation of high-quality data transmission systems with low distortion level.

Therefore, the refusal to apply the concentrated model of a nonlinear system and the transition to its step by step, cascade by cascade identification gives a chance to minimize a mistake in comparison with the standard structure shown in figure 2.

Step by step identification of cascades of a nonlinear system assumes a possibility to use full vector $X$ of its states which can be obtained as a result of direct measurement of coordinates or by means of the corresponding observer which calculating their estimates.

The structure of a system with step by step cascades identification is shown in figure 4.
Figure 4. Structure of step by step cascades identification of a nonlinear system.

The step by step structure of cascades reflects the use of principles of modularity and multilevelness comprising the concept of multialternativity and of complex systems functioning in the methodological plan [19-21].

The modularity of the identification process is implemented by its decomposition to local algorithms which provide an adaptation process activated mainly by the signal of an error of some cascade numbered as $i$.

This technique allows to reduce significantly, by $n$ times, the number of the configured settings of parameters in each local adaptation algorithm and, respectively, to increase identification quality.

Besides, the cascade structure of identification allows to perform an independent definition of the estimates $\hat{f}_i(x_i), \hat{\phi}_i(\omega)$ in the separate processors, i.e. to organize parallel calculations. An important advantage of the cascades step by step identification is the possibility to form several hierarchical levels of the process of adaptation due to the input of the signal of an error to the parameters setting contour of the previous cascade numbered as $i-1$, see figure 3.

The hierarchical structure of the process of general identification of a nonlinear system provides the coherence of local processes of adaptation and their submission to the ultimate goal of identification which is the identity of output signals $\hat{Y}(j\omega)$ and $Y(j\omega)$.

3. Experimental testing of the offered solution

3.1. Model of the preset nonlinear system

As the preset nonlinear system we consider the two-stage structure containing the links connected in series $f_1(x_1), w_1(j\omega), f_2(x_2), w_2(j\omega)$, (see figure 2), with the characteristics show below, figure 5:

\[ f_1(i) = 2.5[1 - \exp[-0.7x_1(i)\text{sign}x_1(i)]] \cdot \text{sign}x_1(i); \]
\[ y_1(i+1) = \{f_1(i)[1 - \exp(-\tau/T_1)] + y_1(i) \cdot \exp(-\tau/T_1)\} \cdot k_1; \]
\[ f_2(i) = 1.25[1 - \exp[-0.8y_1(i)\text{sign}y_1(i)]] \cdot \text{sign}y_1(i); \]
\[ y_2(i+1) = f_2(i)[1 - \exp(-\tau/T_2)] + y_2(i) \cdot \exp(-\tau/T_2), \]
where $\tau$ – is sampling the processes in the system by time; $k_1$ – the coefficient of transfer of the first frequency-dependent link, $k_1 = 1$, for the second link we take the same coefficient as 1; $T_1, T_2$ – time constants of frequency-dependent links $w_1(j\omega), w_2(j\omega)$. $T_1 = 0.001$ s, $T_2 = 0.005$ s; $x_1$ – a signal at the system input, $x_i = X(j\omega)$; $y_1, y_2$ - signals at the stage 1 and stage 2 outputs, respectively, $y_2 = Y(j\omega)$.

Figure 5. Characteristics of the considered system: a) nonlinear characteristics of stages; b) signals at the system’s input and output.

3.2. The concentrated system model with adaptation by output

We will choose the concentrated (one-staged) model as a basic option to apply the adaptive model having the structure shown in figure 3.

We approximate the nonlinear characteristic of this system with the polynomial model of the fourth order [22]:

$$
\hat{F}(x_i) = \sum_{i=1}^{4} h_i x_i^{i-1} = h_1 x_1 + h_2 x_1 \cdot x_1 + h_3 x_1 \cdot x_1^2 + h_4 x_1 \cdot x_1^3.
$$

We will reproduce inertial properties of a two stage concentrated model using the link of second order applying consecutively two differential circuits of the first order:

$$
q(i + 1) = \{\hat{F}[x_1(i)]\cdot[1 - \exp(-\tau/\hat{T}_1)] + q(i) \cdot \exp(-\tau/\hat{T}_1)\} \cdot \hat{k}_1;
$$

$$
y(i + 1) = q(i) \cdot [1 - \exp(-\tau/\hat{T}_2)] + y(i) \cdot \exp(-\tau/\hat{T}_2),
$$

where $q$ – is the result of calculation of the first differential circuit; $\hat{F}, \hat{T}_1, \hat{T}_2, \hat{k}_1$ – estimates of the corresponding parameters of the preset nonlinear system. As a result the constructed nonlinear system model contains seven variable settings: $h_1, h_2, h_3, h_4, \hat{T}_1, \hat{T}_2, \hat{k}_1$.

Nonlinearity of characteristics causes the multy-extreme nature of a function which is the purpose of this adaptation process, namely, the square deviation between output signals of the preset system and those of its model:

$$
\varepsilon(h_1, h_2, h_3, h_4, \hat{T}_1, \hat{T}_2, \hat{k}_1) = \sum_{i=1}^{m} [y(i) - \hat{y}(i)]^2 \to \min,
$$

5
where \( m \) – is a number of the time intervals used for calculation of a criterion in the reviewed example. For the solution of a task (10) with restrictions:

\[
\hat{T}_1 > 0; \quad \hat{T}_2 > 0; \quad \hat{k}_1 > 0; \\
h_i \geq 0; \quad i = \overline{1,4},
\]

the adaptation algorithm is used which is based on the method of the interfaced gradients with an updating using the number of steps \( r \) equal to the number of variables \( (r = 7) \), i.e.:

\[
P(i) = P(0) - \lambda_i \nabla \varepsilon(0); \\
S(i) = -\nabla \varepsilon(i) - \eta(i) \nabla \varepsilon(i - 1); \\
P(i + 1) = P(i) - \lambda_i S(i); \\
i = \overline{1,r},
\]

where \( P = [h_1 \ h_2 \ h_3 \ h_4 \ \hat{T}_1 \ \hat{T}_2 \ \hat{k}_1]^T \) – is the vector of the configured system settings; \( \nabla \varepsilon(i) \) – gradient of function at the current step \( i \) of adaptation; \( \lambda_i \) – the multiplier which determines step length from the initial point \( i \) to \( i + 1 \) point of the intermediate extremum in the direction \( \nabla \varepsilon(i) \) or \( S(i) \); \( S(i) \) – the vector interfaced to the \( P(i - 1) \) vector:

\[
P(i - 1)^T \cdot G \cdot S(i) = 0,
\]

\( G \) – Hesse's matrix; \( \eta \) – the directed multiplier which is defined by the expression:

\[
\eta(i) = \frac{\nabla^2 \varepsilon(i)[\nabla \varepsilon(i) - \nabla \varepsilon(i - 1)]}{\nabla \varepsilon^T(i - 1) \nabla \varepsilon(i - 1)}.
\]

In figure 6 process of adaptation of parameters \( \hat{T}_1, \hat{T}_2 \) is shown which runs until their final values \( \hat{T}_1 = 0.00094, \hat{T}_2 = 0.00501 \) are obtained in 660 steps of search (please note that the values of time constants of the dynamic links of the preset nonlinear system are \( T_1 = 0.001 \) s, \( T_2 = 0.005 \) s).

The multi-extreme nature of the target function (10) is overcome by way of updating the search process and, in general, the curve of function \( \varepsilon(N) \) changes has monotonous character, figure 6b.

Final value of the target function \( \varepsilon(660) = 0.198 \).

![Figure 6](image.png)

**Figure 6.** An illustration of the concentrated nonlinear model adaptation process: a) time constants \( \hat{T}_1, \hat{T}_2 \) control; b) convergence of target function \( \varepsilon(N) \) to minimum.
An average square deviation of a model and real output signals:

\[ S = \left( \sum_{i=1}^{m} \{ y(i) - \hat{y}(i) \}^2 / (m-1) \right)^{1/2} = \left( 0.198 / (250 - 1) \right)^{1/2} = 0.028 \]  
(15)

corresponds to the error level (-31) dB. The obtained of nonlinear distortion level is insufficiently low according to the contemporary requirements to telecommunication channels.

Let's consider a variant of the two stage model calculation from which, according to paragraph 2, better results may be expected.

### 3.3. Two-stage model of a system with adaptation by output

The structure of two-stage model of a system with adaptation by output is shown in figure 7.

![Figure 7. Structure of two-stage model of a nonlinear system with adaptation by output.](image)

Using the adaptation by output, similar to the one-stage model, assumes the use of a general assessment of adaptation quality as an error at the model output (see figure 7), however it gives the chance to increase an adequacy of this model due to the reproduction of multistage structure of a real system.

Moreover, it gives an opportunity to lower significantly an order of the polynomials describing the nonlinear characteristics of each stage.

We approximate the nonlinear characteristics of stages of this system with the second order polynomial models:

\[ \hat{f}_1(x_1) = \sum_{i=1}^{2} h_{i1} x_1 \cdot [x_1]^{-1} = h_{11} x_1 + h_{12} x_1 \cdot [x_1]; \]  
(16)

\[ \hat{f}_2(x_2) = \sum_{i=1}^{2} h_{i2} x_2 \cdot [x_2]^{-1} = h_{21} x_2 + h_{22} x_2 \cdot [x_2]. \]  
(17)

Here we reproduce the inertial properties of each stage in this model using the differential circuits of the first order:

\[ z(i+1) = (\hat{f}_1 [x_1(i)] \cdot [1 - \exp(-\tau / \hat{T}_1)] + z(i) \cdot \exp(-\tau / \hat{T}_1)) \cdot \hat{k}_1; \]  
(18)

\[ y(i+1) = \hat{f}_2 [z(i)] \cdot [1 - \exp(-\tau / \hat{T}_2)] + y(i) \cdot \exp(-\tau / \hat{T}_2), \]  
(19)

where \( z \) – is a signal at the output of the first stage; \( \hat{f}_1, \hat{f}_2, \hat{T}_1, \hat{T}_2, \hat{k}_1 \) – estimates of the corresponding parameters of the preset nonlinear system.

Adaptation of this model was made utilizing an algorithm (12) by the criterion:
\[ \varepsilon(h_{11}, h_{12}, h_{21}, h_{22}, \hat{\hat{T}}_1, \hat{\hat{T}}_2, \hat{k}_1) = \sum_{i=1}^{m} [y(i) - \hat{y}(i)]^2 \rightarrow \text{min}, \]  

which supports seven configured settings, as before.

As a result of adaptation the final value of the target function \( \varepsilon(h_{11}, h_{12}, h_{21}, h_{22}, \hat{T}_1, \hat{T}_2, \hat{k}_1) = 0.124 \) corresponding to the error level (-33) dB is received.

This new advantage of a two-stage structure shown in figure 7 in comparison with one-stage shown in figure 3 is not practically significant because in the described variant of a two-stage structure the idea of functions separation is incomplete: the process of adaptation of each stage is carried out not separately and using the common criterion (see figure 7).

Let’s proceed to the variant of a model which implements completely the identification method based on the idea of multialternativity.

### 3.4. Two-stage model of a system with adaptation by vector of state

The diagram of the two-stage model of system with adaptation by vector of state is similar to the diagram shown in figure 4. In this structure the adaptation of each stage is made separately, according to its own state (an output signal). At the same time the model’s first stage occupies higher hierarchical level in relation to the second stage which is the output stage.

Separated, hierarchically organized process of adaptation gives the chance to approximate each stage with polynomials of high orders. In the variant of a model considered below the third order of approximation is used for the first stage:

\[
\hat{f}_1(x_1) = \sum_{i=1}^{3} h_{11} x_1 [x_1]^{-1} = h_{11} x_1 + h_{12} x_1 \cdot [x_1] + h_{13} x_1 \cdot [x_1]^2
\]

and the second order of approximation is for the second stage, i.e. the total number of the configured variable settings for the specified stage is five and three, respectively, (see models (17)-(19).

Functions of the adaptation goals can be drawn as:

\[
\varepsilon_1(h_{11}, h_{12}, h_{13}, \hat{T}_1, \hat{k}_1) = \sum_{i=1}^{m} [z(i) - \hat{z}(i)]^2 \rightarrow \text{min} ;
\]

\[
\varepsilon = \varepsilon_2(h_{21}, h_{22}, \hat{T}_2) = \sum_{i=1}^{m} [y(i) - \hat{y}(i)]^2 \rightarrow \text{min} .
\]

As a result of experimental test of a two-stage model with adaptation by vector of state the final value of a criterion \( \varepsilon = 0.025 \), corresponding to the error level (-40) dB was obtained. Comparative characteristics of the variants of identification models considered above are shown in the table 1.

| No. | Model variant | Model order | Absolute value of the sum of error squares | Average value of a square deviation, dB |
|-----|--------------|-------------|-------------------------------------------|---------------------------------------|
| 1   | One-stage    | 7           | 0.198                                     | -31                                   |
|     | concentrated model |       |                                           |                                       |
| 2   | Two-stage model with adaptation by output | 7           | 0.124                                     | -33                                   |
|     | Two-stage model with adaptation by vector of state | 5 (stage 1) 3 (stage 2) | 0.025                                     | -40                                   |
This result confirms the efficiency of the stage identification based on the principles of multialternativity.

4. Conclusion
The problem of identification of nonlinear system considered in this work was solved traditionally on the basis of a "black box" concept and usually assumed the concentrated analysis of the identified system and formation of the criterion of adaptation by its output. Such a solution of this task is unable to provide the level of signals distortion assumed by modern requirements in the process of signal transmission via the nonlinear communication channels.

The approach to the nonlinear identification we offer here relies on the principles of multialternative structure and functioning of complex systems, such as: multilevelness, modularity and separation of functions of the internal subsystems. According to these principles the multistage model structure of a nonlinear system is offered. An adaptation of stages is provided utilizing the local criteria of quality calculated independently for each stage. At the same time there is a hierarchical interrelation of adaptation processes in the stages, i.e. the combination of the model’s control processes independence and their structural subordination.

Stage-by-stage performance of functions of model adaptation gives way to a decomposition of the general problem of high dimension to several simple tasks and to break thereby the barrier of a "dimensional curse" arising in case of identification of complex systems.

Additional benefit of a stage-by-stage identification is the possibility of the parallel on-line process of simulation using independent computers.

References
[1] Tarighat A and Sayed A H 2005 MIMO OFDM receivers for systems with IQ imbalances IEEE Transactions on Signal Processing 53 pp 3583-96
[2] Ifeachor E C and Jervis B W 2001 Digital Signal Processing : A Practical Approach ( Prentice Hall Inc.) p 960
[3] Diniz P S R 2013 Adaptive Filtering Algorithms and Practical Implementation (New York: Springer Science, Business Media) p 230
[4] Wang W-X, Lai Y-C and Grebogi C 2016 Data based identification and prediction of nonlinear and complex dynamical systems Phys. Rep. 644 pp 1–76
[5] Apolinário J A 2009 QRD-RLS adaptive filtering (New York: Springer) p 356
[6] Karimi-Ghartemani M and Iravani M R 2002 A nonlinear adaptive filter for online signal analysis in power systems: applications IEEE Transactions on Power Delivery 17(2) pp 617–622
[7] Lee K A, Gan W S and Kuo S M 2009 Subband Adaptive Filtering: Theory and Implementation (UK, Chichester: John Wiley & Sons, Ltd.) p 324
[8] Roweis S T and Lawrence K S 2000 Nonlinear dimensionality reduction by locally linear embedding Science 290 pp 2323–26
[9] Roll J, Nazin A and Ljung L 2005 Non-linear system identification via direct weight optimization Automatica 41(3) pp 475–490
[10] Lan Y and Mezic I 2013 Linearization in the large of nonlinear systems and Koopman operator spectrum. Physica D 242 pp 42–53
[11] Sijbert J, Zhang Q, Ljung L, Benveniste A, Delyon B, Glionenec P Y, Hjamarsson H and Juditskys A 1995 A nonlinear black-box modeling in system identification: a unified overview Automatica 31 (12) pp 1691-1724
[12] Brunton S L, Proctor L P and Kutz J N 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems Proc. of the National Academy of Sciences 113(15) pp 3932–37
[13] Sayed A H 2008 Adaptive filters (NJ, Hoboken: John Wiley & Sons, Inc.) p 785
[14] Billings S A 2013 Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains (New York: John Wiley & Sons, Inc.) p 574

[15] Kumar R, Srivastava S, Gupta J R P and Mohindru A 2018 Self-recurrent wavelet neural network–based identification and adaptive predictive control of nonlinear dynamical systems Int. Journal of Adaptive Control and Signal Processing 32 (3) pp 1326-58

[16] Mandic D P and Goh V S L 2009 Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely linear and Neural Models (UK, Chichester: John Wiley & Sons, Inc.) p 324

[17] Giri F and Bai E W 2010 Block-Oriented Nonlinear System Identification (Heidelberg: Springer) p 431

[18] Brouri A 2017 Two stages nonlinear systems identification International Journal of Control Systems and Roboti, 2 pp 12-15

[19] Podvalny S L and Vasiljev E M 2015 A multi-alternative approach to control in open systems: origins, current state, and future prospects Automation and Remote Control 76 (8) pp 1471-99

[20] Podvalny S L, Vasiljev E M and Barabanov V F 2014 Models of multi-alternative control and decision-making in complex systems Automation and Remote Control 75 (10) pp 1886-90

[21] Podvalny S L and Vasiljev E M 2018 Multi-alternative control of large systems 13th Int. Scientific-Technical Conference on Electromechanics and Robotics “Zavalishin’s Readings”, St. Petersburg, MATEC Web of Conferences 161

[22] Paduart J, Lauwers L, Swevers J, Smolders K, Schoukens J and Pintelon R 2010 Identification of nonlinear systems using polynomial nonlinear state space models. Automatica 46 pp 647–656