

# Optimisation of neural state variables estimators of two-mass drive system using the Bayesian regularization method

M. KAMINSKI\* and T. ORLOWSKA-KOWALSKA

Institute of Electrical Machines, Drives and Measurements, Wrocław Institute of Technology,  
19 Smoluchowskiego St., 50-372 Wrocław, Poland

**Abstract.** The paper deals with the application of neural networks for state variables estimation of the electrical drive system with an elastic joint. The torsional vibration suppression of such drive system is achieved by the application of a special control structure with a state-space controller and additional feedbacks from mechanical state variables. Signals of the torsional torque and the load-machine speed, estimated by neural networks are used in the control structure. In the learning procedure of the neural networks a modified objective function with the regularization technique is introduced. For choosing the regularization parameters, the Bayesian interpretation of neural networks is used. It gives a possibility to calculate automatically these parameters in the learning process. In this work results obtained with the classical Levenberg-Marquardt algorithm and the expanded one by a regularization function are compared. High accuracy of the reconstructed signals is obtained without the necessity of the electrical drive system parameters identification. Simulation results show good precision of both presented neural estimators for a wide range of changes of the load speed and torque. Simulation results are verified by the laboratory experiments.

**Key words:** electrical drive, two-mass system, state estimation, neural networks, training methods, Bayesian regularization.

## 1. Introduction

In the recent years an increase of the neural networks (NN) applications in electrical drives has been observed. NN models are often implemented in control of the speed or position of the drive systems [1], state variable estimation of the electrical motors [2], control of the power electronics systems [3] or in the fault detection and diagnosis of industrial processes [4]. Most papers appearing in the technical literature concern drive systems with a stiff connection of the motor and load machine. However, in the real drives, connecting elements of the mechanical part of the drive system are characterized by a limited stiffness, e.g. rolling-mill drives, conveyer belt or cage host drives. In many cases the appearing elasticity of connecting shaft in the drive can considerably deteriorate precision of the speed or position control, and in special cases can even lead to the loss of stability. These complications of the system work are caused by oscillations of the state variables resulting from the elasticity of the mechanical part of a drive. One of the methods often applied for solving this problem is a modification of the control structure by introducing additional feedbacks from chosen state variables [5, 6]. Therefore, the problem of getting the information about required feedback signals appears, as in the most of real systems the measurements of all mechanical state variables necessary for the control structure are impossible (problem with sensors' installation) or associated with the increasing costs (additional measuring arrangements).

The application of NN for the state variables estimation of electrical drives with the elastic coupling enables obtaining

the required signals without necessity of parameter identification of the drive required in algorithmic estimators [7].

The issue of the state variables estimation in the two-mass drive system is more complicated than in the drive with perfectly stiff shaft. In the system with an elastic shaft we can deal with the shaft torque influenced also the electromagnetic torque of the drive. This phenomenon is very important as the electromagnetic torque (or current) is usually incorporated into the input vector of the NN estimator. Therefore applying techniques taken from the theory of neural networks for improving their generalization properties is advantageous [8].

The data generalization realized by the NN appoints the degree of the possibility and accuracy for solving the assumed task by trained network in the case of appearing the input vector elements not taken into account in the training process.

One of the most effective methods used for improving this ability of NN is a regularization. It consists in the modification of the objective function used in the training algorithm, which is minimized in the each iteration. In the extended form of such cost function, elements dependent on values of the inter-neural-connection coefficients are added to the standard function. A problem of the selection of the regularization parameters in the modified objective function is appearing.

In this work the regularization method based on the Bayesian interpretation of NN is applied. This algorithm gives analytical formulas for the automatic computation of the optimal regularization parameters [9–14]. Simulation and experimental results presented in this paper, lead to the conclusion,

---

\*e-mail: marcin.kaminski@pwr.wroc.pl

that the Bayesian regularization method can improve significantly the quality of state variables estimation of the drive system with an elastic joint.

The paper is divided into six sections. After a short introduction the mathematical model of the two-mass drive system is presented. Then the speed control structure with additional feedbacks from the shaft torque, the motor and the load speeds, is described. Next the issues related to the Bayesian regularization and neural networks are presented. The influence of the regularization in the classical Levenberg-Marquardt algorithm on the quality of the state variable estimation is presented. The designed NN estimators are applied in the control structure and tested under simulation and experimental research.

## 2. Mathematical model of the two-mass drive system and control structure

The electrical drive with elastic joint can be described by different mathematical models, depending on the exactness of the elastic shaft modeling. Usually, such a drive is analyzed as a system composed of two masses connected by an elastic shaft, where the first mass represents the moment of inertia of the drive and the second mass refers to the moment of inertia of the load side. In the case of a small value of the shaft inertia in comparison with the motor and load inertia moments, the mechanical coupling is treated as inertia free. Also the internal damping of the shaft is often neglected. With these assumptions, the mechanical part of the considered two-mass drive can be described by the following mathematical model (in per unit system [p.u.]), where nonlinear phenomena, like backlash or friction torques are neglected:

$$T_1 \frac{d\omega_1(t)}{dt} = m_e(t) - m_c(t), \quad (1)$$

$$T_2 \frac{d\omega_2(t)}{dt} = m_c(t) - m_L(t), \quad (2)$$

$$T_c \frac{dm_c(t)}{dt} = \omega_1(t) - \omega_2(t), \quad (3)$$

where  $\omega_1, \omega_2$  – the motor and load speeds,  $m_e, m_L$  – the shaft and load torques,  $T_1, T_2$  – the mechanical time constants of the motor and load machine,  $T_c$  – the stiffness time constant.

The classical cascade control structure consists of two major control loops: the inner control loop contains the current controller, the power converter and the motor. After optimization, the current control loop can be replaced by the first-order inertial component with small time constant. During the design process of the speed loop the dynamics of the torque loop is very often neglected [6]. In most of applications, classical PI or PID speed controllers are applied [5, 6]. In this paper the state controller with an integral action is applied for speed control. Structure of the drive with two-mass system and the state-space controller is presented in Fig. 1.

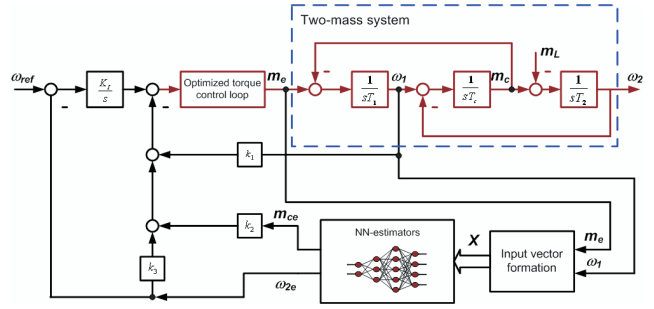


Fig. 1. The block diagram of the control structure

Gain values of this control structure can be adjusted according to the pole placement method. So the speed controller parameters and gains of the feedbacks from such state variables as: the motor and load side speeds as well as the torsional torque are calculated using following formulas:

$$k_1 = 4T_1\xi_r\omega_r, \quad (4)$$

$$k_2 = T_1T_c \left( 2\omega_r^2 + 4\xi_r^2\omega_r^2 - \frac{1}{T_cT_2} - \frac{1}{T_cT_1} \right), \quad (5)$$

$$k_3 = \omega_r^2 k_1 T_2 T_c - k_1, \quad (6)$$

$$K_I = T_1 T_2 T_c \omega_r^4, \quad (7)$$

where  $\omega_r, \xi_r$  – are the assumed values of the resonant frequency and the damping coefficient of the closed-loop structure.

In the control structure, presented above, different state feedbacks are introduced, so the information about the shaft torque  $m_c$ , motor speed  $\omega_1$  and load speed  $\omega_2$  is required. The measurement of the motor speed  $\omega_1$  is simple and trouble free, but the measurement of the shaft torque and the load speed can be difficult or expensive. In this case special estimation structures based on neural network (NN) can be used.

In the technical literature many methods for neural networks structure optimization can be found. Most of them require the initial choice of NN structure and then selected neural connections are eliminated. One of the simplest ways to choose a connection between nodes for elimination is an analysis of absolute values of NN weights. The other method consists in checking the influence of each connection on the generalization error, as in [15, 16]. Other solution for this problem is adding the regularization element to the cost function [17]. It consists in the modification of the objective function used in the training algorithm, which is minimized in any iteration. In the extended form of such cost function elements dependent on the values of the inter-neural-connection coefficients are added to the standard function. A problem of the selection of regularization parameters in the modified objective function is appearing. In this work regularization method based on the Bayesian interpretation of NN is applied. This algorithm gives analytical formulas for the automatic computation of the optimal regularization parameters [9–14], so it is much more convenient in the practical applications and thus is used in this paper for the neural estimator optimization of state variables of the two-mass drive system.

### 3. Bayesian regularization in neural networks

The neural networks training process can be defined as a minimization of the objective function. In the considered case analyzed cost function is described by the following equation:

$$F = \alpha E_W + \beta E_D, \quad (8)$$

where element  $E_D$  is a sum of squares of NN calculation errors for each input sample, and  $E_W$  is a regularization term presented below:

$$E_D = \sum_{j=1}^M (d_j - y_j)^2, \quad (9)$$

$$E_W = \sum_{i=1}^W w_i^2, \quad (10)$$

where  $d_j$  – desired output values;  $y_j$  – actual output values of the neuron;  $M$  – dimension of the vector  $\mathbf{d}$ ,  $w_i$  – weights;  $W$  – the total number of weight and biases in the network.

In relation to the objective function (8), a problem of the parameters  $\alpha$  and  $\beta$  selection is appearing. The regularization parameters describe the influence of the additional term in the cost function. The second coefficient decides on matching of the training data and the first one enforces the smoothness of NN output signals [17]. If  $\beta$  is relatively significant in comparison with  $\alpha$ , the training error is smaller and effect is like in the classical algorithm. In the other case training causes smaller weights and leads to a smoother output signal of the network [11]. Therefore, to achieve a good quality of state estimation, the optimal values for these factors are important. In many cases those parameters can be chosen using cross-validation techniques, but this procedure is very time-consuming. In this paper, the comparison between the classical Levenberg-Marquardt algorithm and after the regularization procedure with parameters chosen by Bayesian optimization method is presented.

In the Bayesian interpretation of NNs, optimization of inter-neural weights corresponds to the increasing of the probability:

$$P(\mathbf{w}|\mathbf{D}, \alpha, \beta, \mathbf{A}) = \frac{P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{A})P(\mathbf{w}|\alpha, \mathbf{A})}{P(\mathbf{D}|\alpha, \beta, \mathbf{A})}, \quad (11)$$

where  $\mathbf{w}$  – weight coefficient vector,  $\mathbf{D}$  – training data,  $\mathbf{A}$  – structure of the neural network,  $P(\mathbf{w}|\alpha, \mathbf{A})$  – describes the information on the weights' values before introducing the training data,  $P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{A})$  – probability of obtaining the established response of the NN for suitable inputs, depending on parameters of the network,  $P(\mathbf{D}|\alpha, \beta, \mathbf{A})$  – normalization element.

Under assumption that noises in the input data (measurements), used in the process of NN training is a Gaussian, as well as probability of the weight distribution is a Gaussian, suitable elements in the equation (11) are described by the following formulas:

$$P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{A}) = \frac{1}{Z_D(\beta)} \exp(-\beta E_D) \quad (12)$$

and

$$P(\mathbf{w}|\alpha, \mathbf{A}) = \frac{1}{Z_W(\alpha)} \exp(-\alpha E_W), \quad (13)$$

where

$$Z_D(\beta) = \left(\frac{\pi}{\beta}\right)^{\frac{M}{2}} \quad (14)$$

and

$$Z_W(\alpha) = \left(\frac{\pi}{\alpha}\right)^{\frac{W}{2}} \quad (15)$$

thus we get

$$P(\mathbf{w}|\mathbf{D}, \alpha, \beta, \mathbf{A}) = \frac{\frac{1}{Z_D(\beta)} \frac{1}{Z_W(\alpha)} \exp(-(\alpha E_W + \beta E_D))}{P(\mathbf{D}|\alpha, \beta, \mathbf{A})}. \quad (16)$$

For the optimization of  $\alpha$  and  $\beta$  parameters in the objective function, the following equation is taken into account:

$$P(\alpha, \beta|\mathbf{D}, \mathbf{A}) = \frac{P(\mathbf{D}|\alpha, \beta, \mathbf{A})P(\alpha, \beta|\mathbf{A})}{P(\mathbf{D}|\mathbf{A})}. \quad (17)$$

Under the assumption that a distribution of regularization coefficients  $\alpha$  and  $\beta$  is uniform, maximal values of the probability  $P(\alpha, \beta|\mathbf{D}, \mathbf{A})$  are obtained for the biggest values of the element  $P(\mathbf{D}|\alpha, \beta, \mathbf{A})$ . Probability  $P(\mathbf{D}|\mathbf{A})$  is independent from those parameters.

After suitable transformations, equations describing  $\alpha$  and  $\beta$  parameters for minimization of the objective functions are obtained [9–11]:

$$\alpha = \frac{\gamma}{2E_W(\mathbf{w}_{MP})} \quad (18)$$

and

$$\beta = \frac{M - \gamma}{2E_D(\mathbf{w}_{MP})}, \quad (19)$$

where

$$\gamma = W - 2\alpha \text{trace}(\mathbf{H})^{-1} \quad (20)$$

and  $\mathbf{w}_{MP}$  – minimum point of the objective function,  $\mathbf{H}$  – hessian matrix of the cost function.

The parameter  $\gamma$  means effective number of parameters of the NN, however  $W$  is a number of all parameters of the network.

### 4. Simulation results

The NN estimators are tested in the control structure with the state space controller and additional feedbacks from the shaft torque, motor and the load speeds of the two-mass drive. The main parameters of the drive system are the following:  $T_1 = T_2 = 203$  ms and  $T_c = 2.6$  ms. The assumed values of the resonant frequency and the damping coefficient of the drive are respectively:  $\omega_r = 45$  s<sup>-1</sup> and  $\xi_r = 0.7$ . A sampling time taken in a simulation for the NN training and testing is set to 0.1 ms.

The structure of the NN estimators is the same for both presented estimators and is described as follows:

$$\{NN\} = \{6-7-8-1\}, \quad (21)$$

which means: 6 inputs, 7 neurons in the first hidden layer, 8 neurons in the second hidden layer, 1 output neuron. The

decision about a number of the hidden layers was taken basis on the Kolmogorov theorem [16] and practical experience with the application of neural network in the analyzed problem

[15]. For the hidden layers the nonlinear tangensoidal activation functions are applied. The linear activation functions are selected as the output function of the neural estimators.

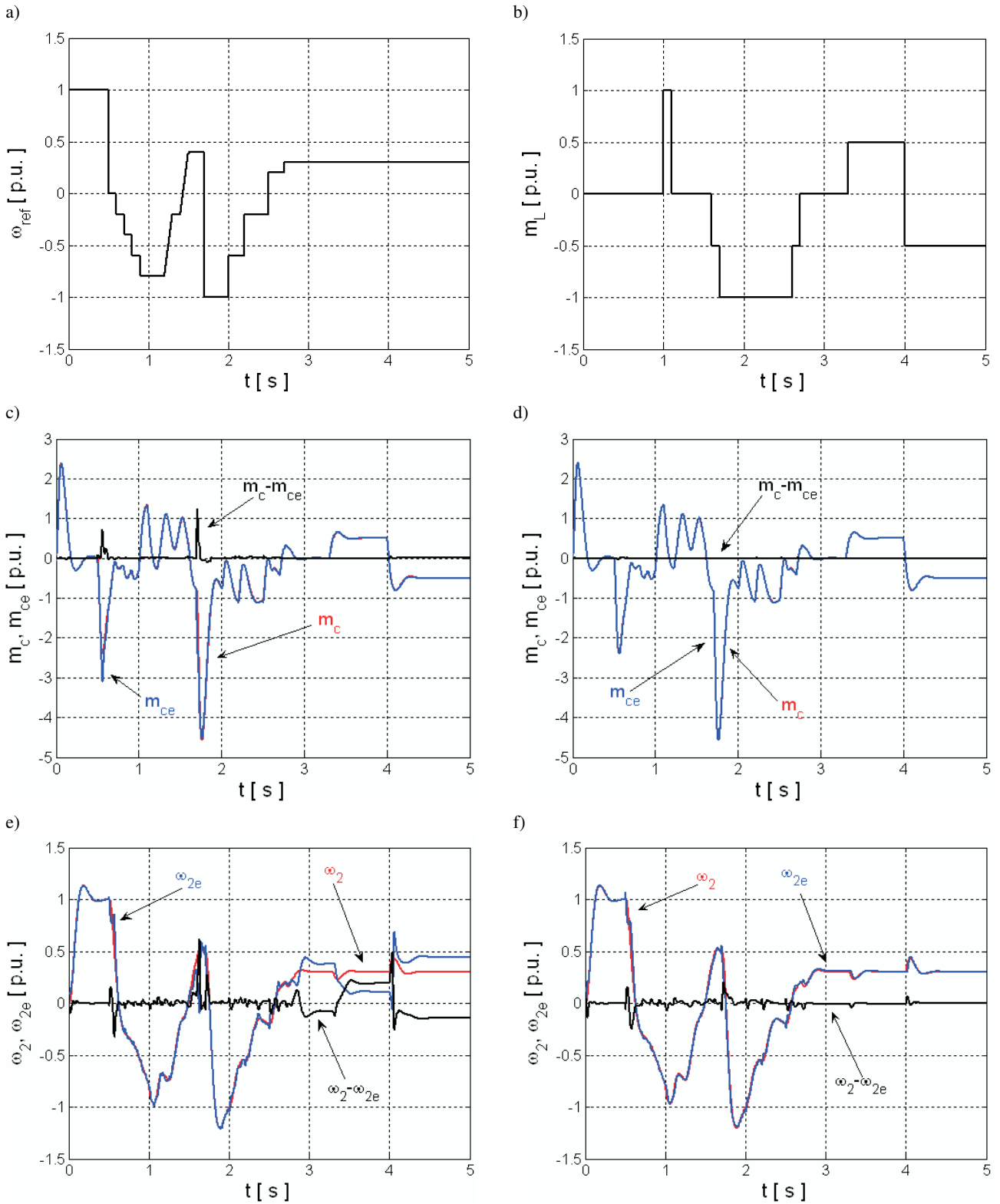


Fig. 2. Simulated transients of the reference speed (a), load torque (b), real and estimated torsional torque  $m_c$  (c,d) and load speed  $\omega_2$  (e,f) and their estimation errors for NN trained with classical Levenberg-Marquardt algorithm (c,e) and after adding Bayesian regularization (d,f)

Input signals of the NN are the torque  $m_e$  and the speed of the motor  $\omega_1$ . However, the outputs signals of NN are estimated state variables: torsional torque  $m_c$  and speed of the load machine  $\omega_2$ . In an addition, for better precision of state variable estimation, the input vector is expanded by historical elements including samples of the electromagnetic torque and the motor speed delayed by one and two periods. Thus the input vector is defined in the following way:

$$\mathbf{X} = [\omega_1(k), \omega_1(k-1), \omega_1(k-2), m_e(k), m_e(k-1), m_e(k-2)]. \quad (22)$$

For the disturbance reduction of the high dynamic inputs signals and measurement noise, the first-order filters are applied.

Results of states variables estimation obtained using the classical Levenberg-Marquard training algorithm as well as the algorithm with the Bayesian regularization, are presented in Fig. 2.

In order to obtain a quality measure of the estimated signals, the estimation errors of NNs are calculated in the following way:

$$Err = \frac{\sum_{i=1}^N |x_i - \hat{x}_i|}{N} \cdot 100, \quad (23)$$

where  $x_i$  – real value,  $\hat{x}_i$  – estimated value,  $N$  – number of samples.

The estimation errors (average error per sample) for neural networks trained with the classical Levenberg-Marquardt algorithm are about 7.93 for the load speed, and about 1.93 for the shaft torque. Such big estimation errors, especially in the case of the load speed estimate (also in a steady-state), cause the oscillation in the closed-loop drive system, when estimated state variables are included in the state feedbacks and can lead to an unstable operation of the drive. After application of the Bayesian regularization, both neural estimators present much higher quality of the estimation and estimation errors are: 1.64 for the load speed and 0.15 for the shaft torque. Steady-state errors are not observed now in the load speed estimate. Analyzed state variables are estimated with a high precision, which is very important in the situation, when those estimators are included in the feedback loops of the drive system.

## 5. Experimental results

The laboratory set-up is composed of two DC machine motors (0.5 kW each). The motor is connected to a load machine by an elastic shaft (a steel shaft of 5 mm diameter and 600 mm length). The stiffness of the connection depends on the shaft diameter. The speeds of the driven motor and the load machine are measured by incremental encoders (36000 pulses

per rotation). On the laboratory set-up the LEM sensors for current measurements are implemented. Measured data and control signals are connected with digital and analog I/O of the dSPACE 1102 card. The motor is driven by a power converter. The load machine in the drive system is controlled using DSP also (Fig. 3).

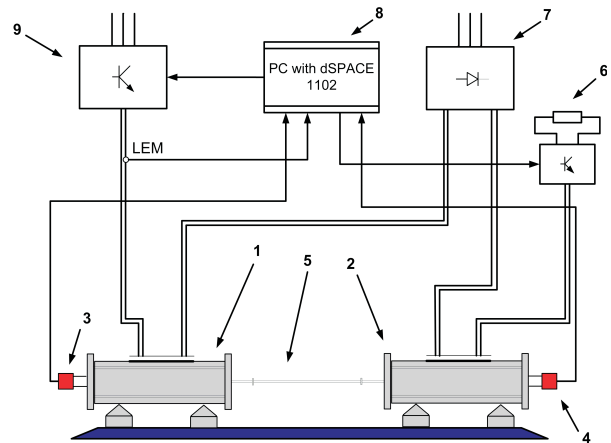


Fig. 3. Schematic diagram of the experimental set-up, where: 1 – motor machine, 2 – load machine, 3, 4 – encoders, 5 – shaft, 6 – resistor, 7 – rectifier, 8 – control structure, 9 – power converter

In Fig. 4 transients of the closed-loop control structure, with NN state estimators in feedback loops, is presented. The assumed value of the resonant frequency of the closed-loop system is  $\omega_r = 45 \text{ s}^{-1}$  and the damping coefficient is  $\xi_r = 0.7$ . There is no shaft torque sensor in the laboratory set-up. Therefore, in order to check the estimated shaft torque shape, the Kalman filter is applied [7]. The input filters of NN estimators are applied with the following time constant: for the motor speed  $T_{\omega_1} = 0.01 \text{ s}$ , for the electromagnetic torque (current)  $T_{m_e} = 0.05 \text{ s}$ . In experiments only state estimators obtained with the application of the Bayesian regularization during the training process of NN, are tested.

Presented results are obtained for estimators implemented in the closed-control-loop. At the time about  $t = 1.4 \text{ s}$  the passive load torque  $T_L = 1.4 \text{ [p.u.]}$  is applied. Despite the load torque is bigger than used in the training process, the neural estimator of the torsional torque is able to estimate it properly, due to the generalization ability of the NN. The estimation error of the torsional torque obtained in the control structure is about 4.23, however imprecision of the load speed estimation is very low and close to 0.86. The NN estimators provide the very good estimation accuracy and the operation of the closed-loop system with estimators in the feedback paths is stable, with good ability of torsional vibration dumping.

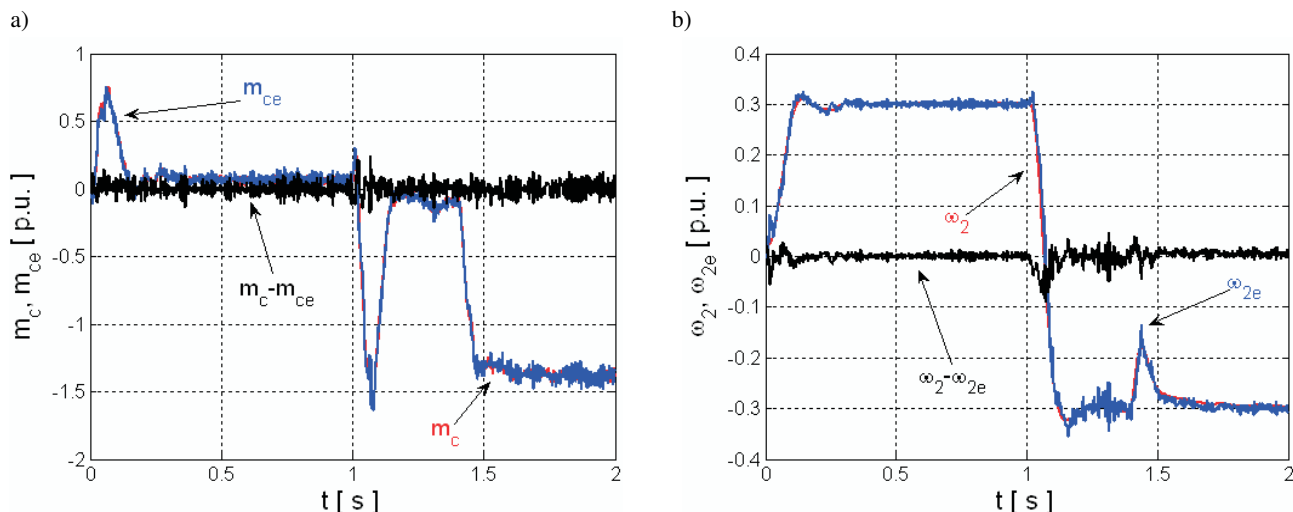


Fig. 4. Experimental transients of the real and estimated torsional torque  $m_c$  (a) and load speed  $\omega_2$  (b), and their estimation errors for NN estimators trained according to the Levenberg-Marquardt algorithm with Bayesian regularization

## 6. Conclusions

The use of artificial neural networks for state variable estimation, in the implementation of drive systems with elastic couplings, allows the reconstruction of chosen signals with very high accuracy. Their main advantage is the design simplicity, as the neural estimators do not need a mathematical model and system parameters, only the training data is required.

The described regularization procedure implemented in the NN training enables improvement of the ability to the data generalization. Application of the Bayesian regularization eliminates the problem of the selection of regularization parameters. Introduction of the automatic regularization into the Levenberg-Marquardt learning algorithm leads to better results of the load speed and shaft torque reconstruction than those obtained using a classical algorithm. Correct work of the designed estimators in the closed-loop system is confirmed experimentally in the laboratory set-up.

**Acknowledgements.** This research work was supported by the Ministry of Science and Higher Education, Poland, under Grant N510 358336 (2009–2010).

## REFERENCES

- [1] M. Cristea and A. Dinu, "A new neural networks approach to induction motor speed control", *IEEE Power Electronics Specialists Conf. 2*, 784–787 (2001).
- [2] M.F.S.V. D'Angel and P.P. Costa, "State estimation for induction machines using a neural network back-propagation technique", *IEEE Int. Conf. on Systems, Man, and Cybernetics 4*, 2613–2618 (2000).
- [3] L.M. Grzesiak and J. Sobolewski, "Energy flow control system based on neural compensator in the feedback path for autonomous energy source", *Bull. Pol. Ac.: Tech.* 54, (3), 335–340 (2006).
- [4] J. Korbicz, "Robust fault detection using analytical and soft computing methods", *Bull. Pol. Ac.: Tech.* 54, (1), 75–88 (2006).
- [5] G. Zhang and J. Furusho, "Speed control of two-inertia system by PI/PID control", *IEEE Trans. on Industrial Electronic* 47 (3), 603–609 (2000).
- [6] T. Orłowska-Kowalska and K. Szabat, "Vibration suppression in two-mass drive system using PI speed controller and additional feedbacks – comparative study", *IEEE Trans. Ind. Electronics* 54 (2), 1193–1206 (2007).
- [7] K. Szabat, T. Orłowska-Kowalska, and K. Dyrz, "Extended Kalman filters in the control structure of two-mass drive system", *Bull. Pol. Ac.: Tech.* 54 (3), 315–325 (2006).
- [8] D.R. Hush and B.G. Horne, "Progress in supervised neural networks", *IEEE Signal Processing Magazine* 10 (1), 8–39 (1993).
- [9] D.J.C. MacKay, "Bayesian Interpolation", *Neural Computation* 4 (3), 415–447 (1992).
- [10] D.J.C. MacKay, "A practical Bayesian framework for back-propagation networks", *Neural Computation* 4 (3), 448–472 (1992).
- [11] F. Dan Foresee and M.T. Hagan, "Gauss-Newton approximation to Bayesian learning", *IEEE Int. Conf. on Neural Networks* 3, 1930–1935 (1993).
- [12] D. Mirikitani and N. Nikolaev, "Recursive Bayesian Levenberg-Marquardt training of recurrent neural networks", *Int. Joint Conf. on Neural Networks* 1, 282–287 (2007).
- [13] R. Gencay and M. Qi, "Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging", *IEEE Trans. Ind. Electronics* 12 (4), 726–734 (2001).
- [14] K.K. Aggarwal, Y. Singh, P. Chandra, and M. Puri, "Bayesian regularization in a neural network model to estimate lines of code using function points", *J. Computer Sciences* 1 (4), 505–509 (2005).
- [15] T. Orłowska-Kowalska, M. Kaminski, and K. Szabat, "Mechanical state variable estimation of the drive system with elastic coupling using optimised MLP neural networks", *Bull. Pol. Ac.: Tech.* 56 (3), 239–246 (2008).
- [16] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, 1996.
- [17] F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures", *Neural Computation* 7 (2), 219–269 (1995).