2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Dynamic Neural Network with Guaranteed Sensitivity to External Influences

Danil Chernozubov
Center "Supersonic"
Lomonosov Moscow State University
Moscow, Russia
danil.chernozubov66@gmail.com

Arthur Mukhamedov
Center "Supersonic"
Lomonosov Moscow State University
Moscow, Russia
a.mukhamedov@vrmsu.ru

Grigory Bugriy
Center "Supersonic"
Lomonosov Moscow State University
Moscow, Russia
gregbugr@vrmsu.ru

Viktor Chertopolokhov
Center "Supersonic"
Lomonosov Moscow State University
Moscow, Russia
psvr.msu@gmail.com

Isaac Chairez
Inst. of Adv. Mat. for the Sust. Manuf.
Tecnologico de Monterrey
Guadalajara, Mexico
isaac.chairez@tec.mx

***Abstract*—Mathematical models are used to represent a vast array of complex processes in engineering, physics, biology, social science, and economics. Model parameters with significant impacts on identification outcomes are ascertained through parametric sensitivity analysis. Authors previously considered an approximation models for systems with uncertain dynamics using a dynamic neural network. The results obtained from studying the problem of predicting the response variable, indicated that this model has a structural flaw. This flaw manifests as an insensitivity of the weight coefficients to external influences, leading to inaccurate predictions. This insensitivity is marked by the minimal contribution of weight coefficient components in the identification process. This paper discusses modifying learning laws to enhance the sensitivity of the weight coefficients to external signals. Through Lyapunov stability analysis, stable algorithms for weight component evolution that minimize identification error were derived.**

***Index Terms*—Neural Networks, Differential Neural Networks, Non-parametric Identifiers, Guaranteed Sensitivity.**

I. Introduction

Over the past decades, control theory has undergone significant evolution through the incorporation of system identification methodologies. One such identification method is the application of differential neural networks (DNNs). DNNs have shown considerable potential in approximating trajectories inherent in dynamical systems with uncertain mathematical models [1]. Despite their ability to blend existing data, typical differential neural networks show limited capacity for extrapolation beyond their training range. A persistent issue is their observed insensitivity to input after identification.

Parameter sensitivity is essential when building a mathematical model, even based on DNN. Modern mathematical models approximate complex and intricate engineering, physical, biological, social, and economic processes. Therefore, by highlighting the parameters that make the most significant contribution to the approximation process, or conversely, by highlighting those parameters that practically do not affect the identification quality, it is possible to reduce the system's dimensionality, which will positively affect the computational power of the model.

Analyzing sensitivity within the realm of control theory provides valuable perspectives on how the input impact the overall approximation performance of the system [2], [3]. Crucial elements of this analysis are sensitivity derivatives, also known as sensitivity functions. The computations of these derivatives in state-space models have been explored and presented in several studies [4], [5].

The book [6] states the necessity of carrying out identification experiments under output feedback or in closed-loop situations. It clarifies the logic behind prerequisites, such as system instability or continuous supervision necessities for safety, productivity, or economic reasons. It also provides a profound understanding of complex problems related to the closed-loop operation of the nonlinear system approximated by DNN.

Various system parameter sensitivity analysis approaches, from statistical to differential, are considered in [7]. Various optimizations for calculating parameter sensitivity to accelerate models that require high computational power are discussed in [8]. In recent research, sensitivity analysis has been widely broadened to encompass optimal control issues using the approximated model based on DNN [9]–[11].

The primary focus of this study is the relative sensitivity of DNN components. This research introduces a new architecture for a non-parametric adaptive approximation model based on DNNs, suitable for systems with uncertain dynamics. The approximation model combines the dynamics of the DNN and the projector of the system's weight coefficients, using

The paper was prepared with the partial financial support of the Tecnologico de Monterrey, Institute of Advanced Materials for Sustainable Manufacturing under the grant Challenge-Based Research Funding Program 2022 number I006-IAMSM004-C4-T2-T and with the partial financial support of the Ministry of Science and Higher Education of the Russian Federation within the framework of the Center "Supersonic" program (agreement 075-15-2022-331 of April 26, 2022).

an extended form of weights [12]. Implementation of the non-differentiable projection operator ensures guaranteed sensitivity of the model's weight coefficients to external influences. The learning laws for tuning the weights of the continuous projection DNN were developed using a controlled function, similar to the Lyapunov function [13]. Using the Attractive Ellipsoid Method (AEM) [14] allows us to analyze the convergence quality of the developed approximation model.

The influence of uncertainty on the performance of dynamic systems has attracted much attention over the past few decades. Robust control, which means control that remains stable amidst external disturbances, offers a practical and systematic method to design control systems affected by uncertainties. However, it still necessitates an approximate representation of the system.

The aforementioned AEM [14], [15] initially developed for nonlinear systems and subsequently adapted for linear ones, employs asymptotically attracting (invariant) ellipsoids. The main goal of this method is to define robust control and the corresponding attracting ellipsoid such that the system's trajectories converge within a small neighborhood of zero. Applying AEM to linear systems with constant bounded disturbances is considered in [16].

The main contributions of this study include:

- New learning laws keeping weight inside or outside predefined ellipsoids are derived, and a theorem on the practical stability of the identification error is formulated and proven considering the secured sensitiveness of the identifier response to the external input.
- The implementation of these learning laws demonstrated the feasibility of maintaining desired sensitivity in the vestibulo-ocular reflex data when using controlled motions as inputs.

The text of the work is organized as follows. Section II describes the results obtained during the preceding research. In section III-A, the concept of relative sensitivity and the ideological approach to its implementation are introduced. In section III-B, the main results of the present work are given. Section IV describes the method of obtaining data on which numerical testing of the built identifier was performed. Conclusions and final remarks are given in sections V and VI.

II. Differential Neural Network with Secure Sensitivity to External Influences

The system with uncertain dynamics was considered, represented by the following differential equation:

$$\frac{d}{dt}x = f(x(t), u(t)) + \eta(t). \quad (1)$$

The state of system (1) $x = x(t) \in \mathbb{R}^n$ is a function of time. The vector-function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ describes uncertain dynamics, is Lipschitz with respect to the first argument with a positive constant $L_f > 0$, and continuously and linearly depends on the vector-function of external influences $u = u(t) \in \mathbb{R}^m$. The vector-function of external influences $u(t)$ is integrable over the entire time interval of the system's consideration and $\|u\|^2 \leq u^+,\ u^+ > 0$. The permissible class of disturbances, represented by the vector-function $\eta = \eta(t) \in \mathbb{R}^n$, belongs to the following set Σ

$$\Sigma = \left\{\eta \mid \|\eta\|^2 \leq \eta_0 > 0\right\}. \quad (2)$$

Such a class of disturbances is permissible, considering the origin of signals influencing the dynamics of the visual-orientation reflex.

The identifier corresponding to this system is:

$$\frac{d}{dt}\hat{x}(t) = A\hat{x} + W_1(t)\sigma_1(\hat{x}(t)) + W_2(t)\sigma_2(\hat{x}(t))u, \qquad \hat{x}(0) = \hat{x}_0 \in \mathbb{R}^n. \quad (3)$$

Here: $\hat{x}(t) \in \mathbb{R}^n$ – identifier state; $A \in \mathbb{R}^{n\times n}$ – Hurwitz matrix, responsible for the linear part of the identifier's dynamics; $W_1(t) \in \mathbb{R}^{n\times p}$ and $W_2(t) \in \mathbb{R}^{n\times p}$ – weight matrices; $\sigma_1 : \mathbb{R}^n \to \mathbb{R}^p$, $\sigma_2 : \mathbb{R}^n \to \mathbb{R}^{p\times m}$ – activation function matrices.

Numerical testing was carried out on data obtained as a result of experiments on a platform with four degrees of freedom, where a virtual reality helmet was attached to the test subject, from which data on gaze direction and head tilt angles were read. The rotation angles of the eyes around the vertical axis, as well as the angular velocities obtained by numerical differentiation, were used as the state vector. The head angular velocities, which were also obtained by numerically differentiating angles taken from the sensors on the virtual reality helmet, served as the external influence u.

III. Building an Identifier with Guaranteed Sensitivity to External Influences

A. Relative Sensitivity

Let's consider a similar system with uncertain dynamics (1) with similar assumptions and limitations. The corresponding model of the non-parametric identifier is:

$$\frac{d}{dt}x = Ax + W_1^*\sigma_1(x) + W_2^*\sigma_2(x)u + \tilde{f}_e(x(t), t) + \eta(t), \qquad x(0) = x_0 \in \mathbb{R}^n. \quad (4)$$

Here: $x \in \mathbb{R}^n$ – identifier state; $A \in \mathbb{R}^{n\times n}$ – a Hurwitz matrix, responsible for the linear portion of the identifier's dynamics; $W_1^* \in \mathbb{R}^{n\times p}$ and $W_2^* \in \mathbb{R}^{n\times p}$ – weight matrices; $\sigma_1 : \mathbb{R}^n \to \mathbb{R}^p$, $\sigma_2 : \mathbb{R}^n \to \mathbb{R}^{p\times m}$ – activation function matrices; the parameter p is chosen freely; $\tilde{f}_e = \tilde{f}_e(x(t), t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ – represents the approximation error produced by a finite number of activation functions in the identifier. It can be assumed that the modeling error belongs to the following set Ω:

$$\Omega = \left\{\tilde{f}_e \mid \|\tilde{f}_e\|^2 \leq \tilde{f}_0 > 0\right\}. \quad (5)$$

The activation functions $\sigma_1(\hat{x})$ and $\sigma_2(\hat{x})$ should be sigmoidals satisfying Cybenko's Universal Approximation Theorem described in [17].

Considering the approximate dynamic model (4), the corresponding identifier structure would be similar to described in

(3) (here and below we use the notations $\hat{x} = \hat{x}(t)$, $W_1 = W_1(t)$, $W_2 = W_2(t)$, assuming that the state and weights implicitly depend on time).

We would like to be sure that the term $W_2\sigma_2(\hat{x})$ will react more to the external influence u, and also take into account that the inner layer $W_1\sigma_1(\hat{x})$ participates in how the state of the identifier $\hat{x}$ follows the real trajectory of the system (1) x.

Definition 1: Given a continuously differentiable function $y = h(z, \theta)$, which constantly depends on the parameter θ. Then the **relative sensitivity** of y with respect to θ can be calculated:

$$S_{y,\theta} = \frac{\partial y}{\partial \theta}$$

Let's calculate the relative sensitivity of the norm of identification error time derivative in relation to the activation function outputs to estimate the contribution of each component:

$$S_{\dot{\Delta},H_i} = \frac{\partial \|\dot{\Delta}\|}{\partial H_i} = \frac{1}{2\|\dot{\Delta}\|}\left(\frac{\partial \dot{\Delta}}{\partial H_i}\right) = \frac{1}{2\|\dot{\Delta}\|}\left(\frac{\partial (\dot{x} - \dot{\hat{x}})}{\partial H_i}\right) = \frac{1}{2\|\dot{\Delta}\|}\left(-\frac{\partial \dot{\hat{x}}}{\partial H_i}\right). \quad (6)$$

Here:

- $\Delta = \Delta(t) = x - \hat{x}$;
- $H_1 = H_1(\sigma_1(\hat{x}))$; $H_2 = H_2(\sigma_2(\hat{x})u)$;
- $H_i(v) = \begin{bmatrix} v & v & \dots & v \end{bmatrix}^\top \in \mathbb{R}^{n \times n \cdot p}$; $i = \{1, 2\}$.

For convenience of derivative calculation in (6), let's rewrite the initial form of the identifier.

$$\dot{\hat{x}} = \frac{d}{dt}\hat{x} = A\hat{x} + H_1(\sigma_1(\hat{x}))W_{1,v} + H_2(\sigma_2(\hat{x})u)W_{2,v}, \quad (7)$$

where $W_{i,v} = \text{vec}\{W_i\}$. In (7), indices at H_1 and H_2 have a purely cosmetic nature to distinguish components further.

We also add a regularization term $L\Delta$ for better identification error compensation during training. Such L should be selected as Hurwitz matrix. Notice that even with this correction term, the identifier will keep its serial configuration.

Thus, we have:

$$\frac{d}{dt}\hat{x} = A\hat{x} + H_1(\sigma_1(\hat{x}))W_{1,v} + H_2(\sigma_2(\hat{x})u)W_{2,v} + L\Delta. \quad (8)$$

In terms of sensitivity indices (6), the desired result can be formulated as follows:

$$\left\|S_{\dot{\Delta},H_2}\right\| \gg \left\|S_{\dot{\Delta},H_1}\right\|, \quad (9)$$

or equivalently

$$\frac{\left\|S_{\dot{\Delta},H_1}\right\|}{\left\|S_{\dot{\Delta},H_2}\right\|} \leq \beta. \quad (10)$$

The relations given by (9) and (10) offer valuable insight into the prevailing notion that external factors have a significant impact on the weighting coefficients, profoundly influencing the entire identification process. Essentially, the sensitivity of the identification error is expected to be considerably higher concerning the second component of the weighting coefficients due to its susceptibility to external influences. By substituting the values from (6) into (10), we get:

$$\frac{\left\|S_{\dot{\Delta},H_1}\right\|}{\left\|S_{\dot{\Delta},H_2}\right\|} = \frac{\left\|\frac{1}{2\|\dot{\Delta}\|}\left(-\frac{\partial \dot{\hat{x}}}{\partial H_1}\right)\right\|}{\left\|\frac{1}{2\|\dot{\Delta}\|}\left(-\frac{\partial \dot{\hat{x}}}{\partial H_2}\right)\right\|} = \frac{\frac{1}{2\|\dot{\Delta}\|}\left\|\frac{\partial \dot{\hat{x}}}{\partial H_1}\right\|}{\frac{1}{2\|\dot{\Delta}\|}\left\|\frac{\partial \dot{\hat{x}}}{\partial H_2}\right\|}$$
$$= \frac{\left\|\frac{\partial \dot{\hat{x}}}{\partial H_1}\right\|}{\left\|\frac{\partial \dot{\hat{x}}}{\partial H_2}\right\|} = \frac{\|W_{1,v}\|}{\|W_{2,v}\|} \leq \beta \quad (11)$$

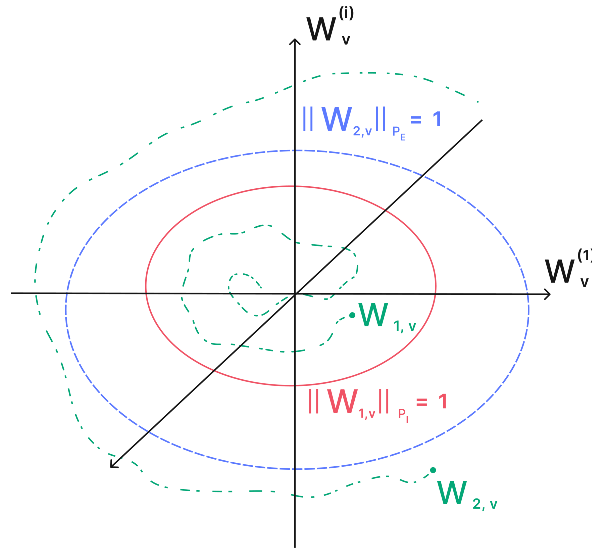

Fig. 1. Illustration of the elliptical constraints concept

To achieve the required sensitivity, we can constrain the norms of the weights to specific values. Specifically, $W_{2,v}$ should have a lower boundary, ensuring that small variations in H_2 result in significant variations of $\hat{x}$. In a similar vein, $W_{1,v}$ should have an upper boundary.

The idea of sensitivity indices and elliptical constraints can be described as follows. Consider a vector space of dimension np, where each coordinate corresponds to an element of vectorized weight matrices. Then, the task of achieving the ratio (10) reduces to the limitation of weight norms with the help of two ellipsoids (Fig. 1), set by the matrices of quadratic forms $P_E, P_I \in \mathbb{R}^{np \times np}$ for the external and internal ellipsoids respectively.

B. Learning laws with predefined sensitivity

One way to represent the proposed constraints in new learning laws, in order to enforce the desired sensitivity ratio, will be to introduce a certain multiplier, which will account for constraints in the evolution of weights dynamics.

We have the identifier of the following form:

$$\frac{d}{dt}\hat{x} = A\hat{x} + H_1(\sigma_1(\hat{x}))W_{1,v} + H_2(\sigma(\hat{x})u)W_{2,v} + L\Delta. \quad (12)$$

Let us introduce learning laws in the following way:

$$\begin{aligned} \frac{d}{dt}W_{1,v}(t) &= \frac{1}{2\gamma_1(t)}\tilde{W}_{1,v}\frac{d}{dt}\gamma_1(t) + \frac{1}{k_1\gamma_1(t)}H_1(\sigma_1(\hat{x}))^\top P\Delta, \\ \frac{d}{dt}W_{2,v}(t) &= \frac{1}{2\gamma_2(t)}\tilde{W}_{2,v}\frac{d}{dt}\gamma_2(t) + \frac{1}{k_2\gamma_2(t)}H_2(\sigma_2(\hat{x})u)^\top P\Delta. \end{aligned} \quad (13)$$

$$W_{1,v}(0) = W_{1,v}^0, \quad W_{2,v}(0) = W_{2,v}^0,$$

where $k_1 \in \mathbb{R}^+$ and $k_2 \in \mathbb{R}^+$ are two positive numbers regulating the evolution of the weights; $\tilde{W}_{j,v} = W_{j,v}^* - W_{j,v}$, $j = 1, 2$; the functions $\gamma_1(t)$ and $\gamma_2(t)$ represent distances to corresponding constraints (and therefore $\gamma_1, \gamma_2 > 0$):

$$\gamma_1(t) = d_1(1 - \|W_{1,v}\|_{P_I}^2), \gamma_2(t) = d_2(\|W_{2,v}\|_{P_E}^2 - 1). \quad (14)$$

Here d_1 and d_2 should be continuous and differentiable with respect to time; $P \in \mathbb{R}^{n\times n}$ is a quadratic form of an attracting ellipsoid, being the solution to the Riccati matrix inequality [18]:

$$P(A + L) + (A + L)^\top P + PRP + Q < 0. \quad (15)$$

The matrix $Q \in \mathbb{R}^{n\times n}$ is positive definite, and $R \in \mathbb{R}^{n\times n}$ can be positive semi-definite. The properties that justify the positive definite and bounded of the class of Riccati equations in (15) have been discussed before in [1], which is a necessary condition for the convergence of the learning laws of the identifier. $P_I \in \mathbb{R}^{np\times np}$ and $P_E \in \mathbb{R}^{np\times np}$ are quadratic forms of ellipsoids: inner and outer respectively.

The learning laws can be derived from stability analysis (in the Lyapunov sense) of the identification error Δ. The application of Lyapunov stability analysis confirms the existence of an upper bound for the identification error [1] [19]. The proposed Lyapunov candidate function uses a quadratic form, depending on the identification error, time, and weights that describe the identifier's dynamics. These weights should be calibrated in such a way that the identification error is bounded eventually. The existence of such a boundary is described by the following theorem.

Theorem 1: Consider a class of nonlinear systems with uncertain dynamics (1), subjected to external disturbances from the set (2). Assume that there exists an identifier model (4) with bounded modeling error, determined by the set (5). Suppose all the above assumptions are true. Then, if there exists a positive definite matrix $P = P^\top, P > 0$, which is a solution for the matrix inequality (15), the application of the learning laws (13) leads to practical stability of the identification error Δ within a sphere centered at the origin:

$$\limsup_{T\to\infty}\left\{\sup_{\eta\in\Sigma,;\tilde{f}_e\in\Omega}\|\Delta(T)\|_P^2\right\} \le \lambda_{\max}(\Lambda)(f_0 + \eta_0), \quad (16)$$

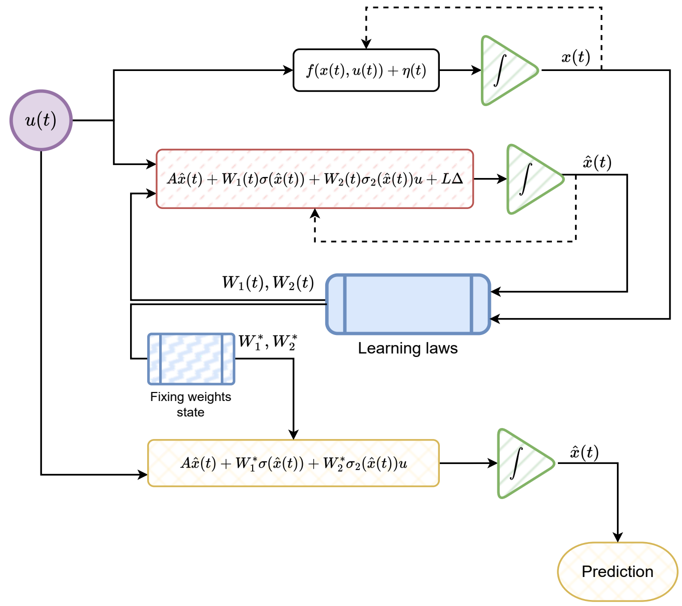

Fig. 2. Identification (upper part) and prediction (lower part) workflows.

where Λ is positive definite matrix.

The identifier algorithm with new learning laws with guaranteed sensitivity to external disturbances is illustrated in Fig. 2. As a rule, it can be divided into two parts: the training cycle and the prediction process. The training cycle consists of several subsections: the extraction of identifiable signals (black rounded rectangle), learning laws (13) (blue solid rounded rectangle), and calculation of an estimated approximation of the target signal using (8) (red dashed rounded rectangle). The prediction process uses fixed weight coefficients derived from learning laws in the training cycle (blue zigzagged rounded rectangle) passed into the estimation equation (yellow cross-hatched rounded rectangle), the integration of which gives the resulting predicted signal. External influences are denoted by solid purple circles.

IV. Experimental Results

Numerical experiments were conducted on data obtained from tests on a platform with four degrees of freedom (Fig. 3).

During the trials, 18 cycles of rotations around the vertical axis were performed with an amplitude of up to 25° (Fig. 4). The data was collected using an HTC Vive Pro Eye virtual reality headset. The headset's position and orientation were obtained through the SteamVR tracking system. The SRanipal software collected data provided by the built-in eye-tracking system, outputting gaze origin and direction vectors for each eye at a maximum frequency of 120 Hz. The resulting eye movements and head dynamics were recorded and later processed for modeling using the proposed identifier.

The experiment was conducted as follows. First, the subject would put on the headset and adjust the straps so that the headset was firmly fixed on the head throughout the experiment. Then the eye tracker was calibrated in accordance

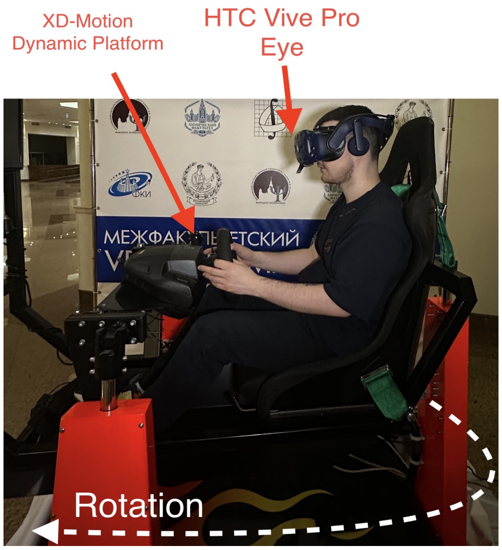

Fig. 3. Experimental platform

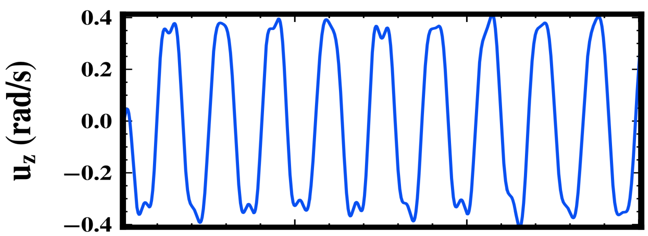

Fig. 4. Example of post-processed head angular velocities

with the SRanipal documentation and recommendations. After the calibration procedure was completed, any adjustments to the headset by the subject resulted in the experiment being repeated in accordance with the SRanipal recommendations. The subject was then seated directly on the dynamic platform.

The angular velocities of the subject's head, obtained by numerically differentiating the angles recorded by the tracker, were used as external disturbances. The deviation angle of the pupil from the frontal horizontal axis of rotation was used as the state of the system to be identified. The presented block diagram (Fig. 2) was implemented using the Python programming language. Two simulations were conducted with two types of data: integration over the entire period with a step of 0.001 and integration only over a portion of the data with the same step, using the remainder for prediction. The identification results over the entire time span using post-processed data are presented in section IV-A. The prediction results on the same data can be found in section IV-B. The post-processing consisted of passing the data through a direct Fast Fourier Transform (FFT), removing frequencies above the 0.99 threshold, and returning to the original space through the inverse FFT. Such processing was applied to both external disturbances (Fig. 4) and pupil deviation angles to neutralize the effect of high-frequency saccades [20] and improve the operation of the identifier. As activation functions, sigmoidal functions were used.

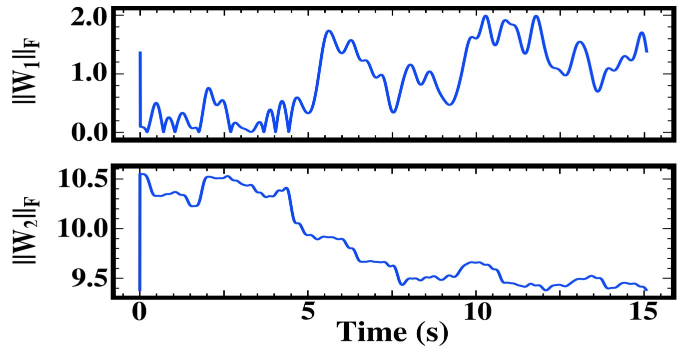

Fig. 5. The dynamics of weight's Frobenius norm

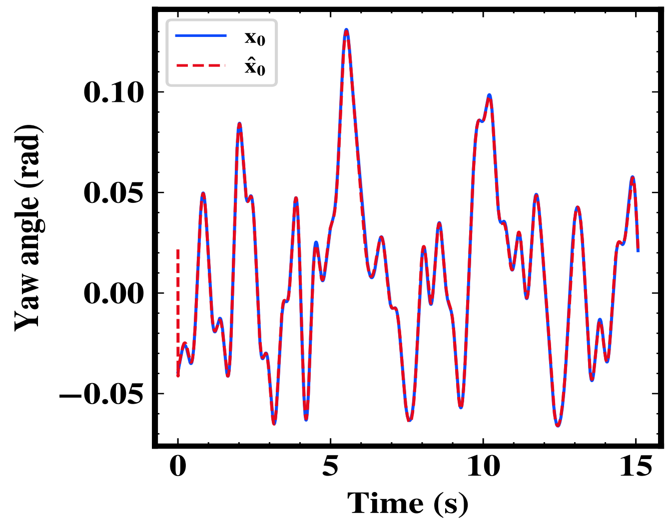

Fig. 6. Identification result of the left eye's yaw using post-processed data and whole dataset

A. Numerical results

The Fig. 6 shows the identification results when trained over the entire time interval. The graph shows the motion of a yaw angle of the left eye.

In Fig. 5, the dynamics of the Frobenius norm of the weight coefficients can be observed. It can be noticed that the norm of W_2 is always greater than the norm of W_1, indicating higher sensitivity to external influences and, consequently, a larger contribution of the second weight matrix.

B. Forecasting results

The prediction was made by fixing the weight coefficients W_1 and W_2 and removing regularization term $L\Delta$ from the dynamics (12). Fig. 7 shows the identification result when all weight coefficients was fixed at the same time $t = 9$. Given that the identifier is unaware of past system states, it can only generate predictions for a short time span in the future. That said, the precision of these predictions is somewhat enhanced by the elimination of high frequencies.

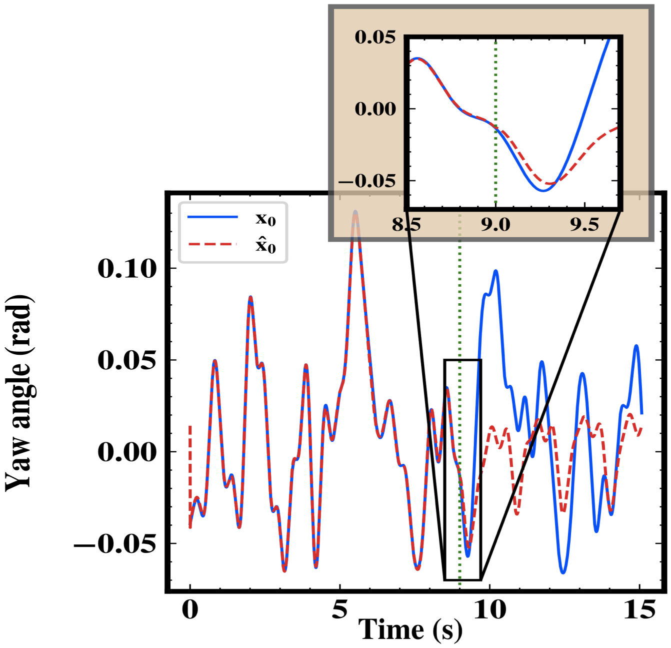

Fig. 7. Identification result of the left eye yaw (in radians) using part of the post-processed dataset in [0, 9] time span

V. General Discussion

This paper introduces sensitivity in the form of weight boundaries. They are defined inside the learning laws in order to manage the importance of different terms of the DNN-based model. Encapsulating the two weight matrices within ellipsoidal bounds ensures that the selected matrix doesn't become negligible. In this study, constraints were added to the control-associated term. However, the approach can be adapted to other setups and term priorities. Similarly, systems with three or more learned terms may use different combinations of restrictions depending on whether lower or upper boundary is more important.

Ideally, system sensitivity should factor in the product $W_1\sigma_1(\hat{x})$ and $W_2\sigma_2(\hat{x})$. Otherwise, the outputs of activation functions may overwhelm the bounded weights in calculations, reintroducing insensitivity into the system. This can be especially problematic if their output varies by orders of magnitude in different parts of the process; for instance, if activation functions output values close to 0 in the initial stages and values near 1 in the later stages. While this isn't critical in our periodic control setup, in some cases, it could necessitate different sensitivity considerations.

VI. Conclusion

In this work, new learning laws with predefined sensitivity to external disturbances for a nonparametric identifier based on a dynamic neural network were proposed. The identifier was used to track the vestibulo-ocular reflex during eye movement. However, the model considered does not claim to reproduce the reflex with physiological accuracy due to insufficient nonlinearity generated by the selected sigmoidal activation functions. However, the introduction of predefined sensitivity into the learning laws of the DNN shows that the model can track the general trend of systems with uncertain dynamics.

References

[1] A. Poznyak, E. Sanchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control: Identification, State Estimation and Trajectory Tracking*. World-Scientific, 09 2001.

[2] L. Radanovic, *Sensitivity Methods in Control Theory: Proceedings of an International Symposium Held at Dubrovnik, August 31-September 5, 1964*. Elsevier, 1964.

[3] R. Rohrer and M. Sobral, "Sensitivity considerations in optimal system design," *IEEE Transactions on Automatic Control*, vol. 10, no. 1, pp. 43–48, 1965.

[4] C. P. Neuman and A. K. Sood, "Sensitivity functions for linear discrete systems with applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-8, no. 6, pp. 764–770, 1972.

[5] N. Gupta and R. Mehra, "Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 774–783, 1974.

[6] A. Simpkins, "System identification: Theory for the user, 2nd edition (ljung, l.; 1999) [on the shelf]," *IEEE Robotics and Automation Magazine*, vol. 19, no. 2, pp. 95–96, 2012.

[7] D. M. Hamby, "A review of techniques for parameter sensitivity analysis of environmental models," *Environmental Monitoring and Assessment*, pp. 135–154, 1994.

[8] C. B. Storlie, L. P. Swiler, J. C. Helton, and C. J. Sallaberry, "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models," *Reliability Engineering and System Safety*, vol. 94, no. 11, pp. 1735–1763, 2009.

[9] X. Li, Y. Li, Z. Liu, and J. Li, "Sensitivity analysis for optimal control problems described by nonlinear fractional evolution inclusions," *Fractional Calculus and Applied Analysis*, vol. 21, pp. 1439–1470, Dec 2018.

[10] Y.-R. Jiang, Q. Zhang, A. Chen, and Z. Wei, "Sensitivity analysis of optimal control problems governed by nonlinear hilfer fractional evolution inclusions," *Applied Mathematics and Optimization*, vol. 84, 12 2021.

[11] Z. Yan, "Sensitivity analysis for a fractional stochastic differential equation with S^p-weighted pseudo almost periodic coefficients and infinite delay," *Fractional Calculus and Applied Analysis*, vol. 25, pp. 2356–2399, Dec 2022.

[12] I. Chairez, O. Andrianova, T. Poznyak, and A. Poznyak, "Adaptive modeling of nonnegative environmental systems based on projectional differential neural networks observer," *Neural Networks*, vol. 151, pp. 156–167, 2022.

[13] A. Guarneros-Sandoval, M. Ballesteros, I. Salgado, and I. Chairez, "Stable learning laws design for long short-term memory identifier for uncertain discrete systems via control lyapunov functions," *Neurocomputing*, vol. 491, pp. 144–159, 2022.

[14] A. Poznyak, A. Polyakov, and V. Azhmyakov, *Attractive ellipsoids in robust control*, p. 348. Birkhauuser, 2014.

[15] A. Kurzhanski and I. Valyi, *Ellipsoidal calculus for estimation and control*. Systems and control : foundations and applications., Boston: Birkhauuser, 1997.

[16] P. García and K. Ampountolas, "Robust disturbance rejection by the attractive ellipsoid method – part i: Continuous-time systems∗∗a companion paper (garcía and ampountolas, 2018) deals with the robust stabilization of discrete-time systems by the aem.," *IFAC-PapersOnLine*, vol. 51, no. 32, pp. 34–39, 2018. 17th IFAC Workshop on Control Applications of Optimization CAO 2018.

[17] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, vol. 2, p. 303–314, Dec 1989.

[18] T. Poznyak, I. Chairez Oria, and A. Poznyak, *How to estimate the quality of applied DNNS*, p. 68–71. Elsevier, 2019.

[19] R. Q. Fuentes-Aguilar and I. Chairez, "Adaptive tracking control of state constraint systems based on differential neural networks: A barrier lyapunov function approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 12, p. 5390–5401, 2020.

[20] M. D. Binder, N. Hirokawa, and U. Windhorst, eds., *Saccade, Saccadic Eye Movement*, pp. 3557–3557. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.