The plasma shape control system in the tokamak with the artificial neural network as a plasma equilibrium reconstruction algorithm

Artem A. Prokhorov*, Yuri V. Mitritshkin**, Pavel S. Korenev*
Mikhail I. Patrov***

*V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia
**Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
(e-mail: prohorov.artem@physics.msu.ru; yvm@mail.ru; pkorenev92@mail.ru)
***A.F. Ioffe Institute of Russian Academy of Sciences, St. Petersburg, Russia
(e-mail: michael.patrov@mail.ioffe.ru)

Abstract: The problem of accurate and reliable plasma shape control is significant, both for modern operating tokamaks, for example for the Globus-M/M2 spherical tokamak, and for future thermonuclear tokamak-reactors using magnetic plasma confinement. The article presents the new results of design and modeling the plasma shape control system for the Globus-M/M2 tokamak with the pre-trained artificial neural network as a plasma equilibrium reconstruction algorithm, which is included in the feedback of the control system. To collect the necessary data for training the artificial neural network and to model the plasma control system the developed magnetic plasma evolutionary code was used.

Keywords: Artificial neural networks, Plasma equilibrium reconstruction, Plasma shape control, Robust control, Quantitative Feedback Theory, Tokamak Plasma Magnetic Evolution Code

1. INTRODUCTION

At present, the perspective Globus-M/M2 spherical tokamak with the aspect ratio of 1.5 (Iofee Institute, St. Petersburg, RF) (Minaev et al., 2017) can be considered as an example of a tokamak from the alternative way to achieve controlled thermonuclear fusion compared to traditional tokamaks with a large aspect ratio (Chuyanov and Gryazneviceh, 2017). In tokamaks, unlike the SISO plasma position and current control loops, in which the controlled signals can be directly measured, the plasma shape can only be estimated from external magnetic measurements by Rogowski coils, flux loops, and magnetic probes (Lao et al., 1985; Mitritshkin et al., 2019). At the same time, the extensive experience in conducting plasma discharges, as well as a vast database of plasma discharges, together with the need to implement a feedback plasma shape control system in Globus-M2 experiments, led to the proposal to employ artificial neural network (ANN) as an online algorithm for plasma equilibrium reconstruction in a feedback loop. The novelty of this article is a step forward from the application of the ANN approach for identifying non-circular plasma equilibrium to controlling the shape of the plasma in the spherical Globus-M2 tokamak in real-time.

An application of ANN models for various objectives related to nuclear fusion research begins in the early 1990s. Significant results were obtained in (Bishop et al., 1995; Windsor et al., 1997) on the COMPASS-D tokamak (GB), where real experiments only the plasma vertical elongation control system was applied, which could not be considered as a comprehensive plasma shape control system. In addition, in (Windsor et al., 1997) an attempt was made to employ an ANN as a nonlinear mapping from measured magnetic signals and soft X-ray signals to geometric parameters of the plasma shape, such as the plasma inner radius, vertical position, minor radius, elongation $k$, triangularity, etc., for the 501 simulated shots. Unfortunately, the accuracy of estimation on the real data was insufficient for the adaptation of the ANN to experiments in real-time.

In addition, a series of results dedicated to the ASDEX Upgrade tokamak is worth mentioning. In (Morabito, 1995) the ANN model was used to classify the type of plasma shape (upper and lower X-point, inner and outer limiter) and outperformed the real-time version of the Function Parameterization technique (Braams et al., 1986) implemented on the plasma control computer of the machine in terms of computation time. In (Coccorese et al., 1994) the technique of using the multidimensional interpolation capability of the multilayer ANN to establish a nonlinear mapping between a set of magnetic flux measurements and shaping parameters (major and minor radius, elongation, triangularity, internal inductance, poloidal beta, $R$ and $Z$ co-ordinates of the X-point) of a non-circular plasma was demonstrated. The shaping parameters were generated by means of a specially adapted version of an MHD equilibrium code on ASDEX Upgrade. However, this work was not brought to the stage of applying the ANN in the control system in experiments.

A similar method (as for the ASDEX Upgrade tokamak) of ANNs for the fast extraction of equilibrium parameters from measurements was applied to the DIII-D tokamak (Lister and Schnurrenberger, 1991) using the EFIT code (Lao et al., 1985) to generate a database, but still, the ANN approach was not applied to plasma shape control purposes.

In (Greco et al., 2007) the ANN model was exploited to classify magnetic variables useful to determine the plasma
shape and position based on the ITER geometry configuration with reduced computational complexity. Similar results for the identification of non-circular plasma equilibrium with the help of the ANN approach were obtained in (Wang et al., 2016) for the EAST tokamak.

All of the above papers are united by the fact that the trained ANN was not used in the feedback loop for real control of the plasma shape, but was only used for offline estimation of plasma equilibrium parameters. Given the hardware implementation of ANNs is straightforward, as well as the property of the ANN to extract useful features from a large data set is well-known. The aim of this work is to demonstrate the possibility of using a feedforward ANN with one hidden layer as an algorithm for plasma equilibrium reconstruction in the plasma shape control loop. To justify the credibility of this approach, several facts should be noted. Firstly, in our era of big data, when the amount of accumulated information is growing rapidly, it is necessary to be able to extract and process useful information from the data. For instance, in JET (GB) more than 45 gigabytes of data can be produced in a well-diagnosed discharge (Murari and Vega, 2014). Secondly, a feed-forward ANN with a single hidden layer has the well-known ability to approximate continuous functions on compact subsets in Euclidian space $\mathbb{R}^n$ (the universal approximation theorem for sigmoid functions is in (Cybenko, 1989)). Recent results have also been obtained for the Rectified Linear Unit (ReLU) activation function (Lu et al., 2017), the use of which significantly accelerates the training of the ANN. Thirdly, the development of modern optimization methods for the backpropagation algorithm, for example, ADAM namely adaptive moment estimation (Kingma and Ba, 2014), as well as the development of methods to avoid overfitting of the ANN (Goodfellow et al., 2016), have significantly improved the performance of ANNs for various objectives. Finally, yet importantly, online versions of the well-known plasma equilibrium reconstruction codes (for example, (Luo et al., 2009)) require powerful computing servers.

Section 2 describes the generation of the dataset for training the ANN based on the developed magnetic plasma evolutionary code. It also uses two methodologies for the synthesis of robust controllers in the application to the tokamak plasma: Quantitative Feedback Theory (QFT) for the SISO control loops, and loop shaping design using McFarlane-Glover method for the MIMO plasma shape loop. Section 3 presents the details of the training of the ANN. Section 4 shows the results of the simulation of the plasma control system for the Globus-M tokamak taking into account the ANN as the plasma equilibrium reconstruction algorithm. Conclusions underline the reasons for the possibility of using the ANN in the plasma shape control closed-loop and the high performance of the presented new plasma shape control system as well as the necessity to develop more advanced plasma magnetic control systems in elongated tokamaks.

2. GENERATION OF THE DATASET

2.1 The Globus-M Tokamak

The Globus-M tokamak (Gusev et al., 2013) is the spherical tokamak (Fig. 1a) with a large accumulated database of plasma discharges, so the simulation results are presented for this version of the tokamak, despite the recent modernization of the plant under control (Minaev et al., 2019).

There are two SISO fast control closed-loops for plasma vertical and horizon position stabilization on the Globus-M/M2 tokamak. In these loops, original thyristor current inverters $A_{HFC, VFC}$ in self-oscillations modes with a frequency of about 3 kHz (Kuznetsov et al., 2019) are used as actuators together with analog PID-controllers $C_{ZF}$. Six SISO feedback systems for the control of currents in the CS and PF coils with multiface thyristor rectifiers $A_{PF-CS}$ as actuators and analog P-controllers $C_{PF-CS}$ form the inner MIMO cascade of the multivariable control system (Fig. 2). The model of the multi-phase thyristor rectifier is investigated in (Mitriyshkin et al., 2016).

Fig. 1. (a) Vertical cross-section of the Globus-M tokamak where PF, CC, HFC, and VFC are poloidal, correction, horizontal, and vertical field coils, respectively; (b) Comparison between the plasma equilibrium obtained with the aid of reconstruction from experimental data using the FCDI code and modelled with the TOPMEC (see Section 2.2).

Fig. 2. Block diagram of the multi-loop hierarchical plasma magnetic control system of the Globus-M/M2 tokamak.

2.2 Tokamak Plasma Magnetic Evolution Code
The plasma evolution codes such as TSC (Jardin et al., 1986) and DINA (Khayrutdinov and Lukash, 1993) allow for simulation of plasma tokamak discharges but they model many transport processes not essential for plasma shape and position control and thus often are too slow and cumbersome for plasma magnetic control system design and simulation.

In order to model plasma dynamics during the entire discharge without solving complex transport equations, the numerical Tokamak Plasma Magnetic Evolution Code (TOPMEC) has been developed in (Mitrishkin et al., 2018a; Mitrishkin et al., 2018b). On each time step of the numerical TOPMEC, the plasma linear model is calculated (Walker and Humphreys, 2006, Mitrishkin et al., 2019) and then used to advance tokamak currents and plasma position. Then new currents and plasma position are used to calculate new plasma equilibrium. The plasma current density is calculated from boundary conditions on plasma current density, total plasma current, and plasma parameters namely the poloidal beta \( \beta_p \) and the normalized plasma internal inductance \( \lambda \) specified by the user. The updated plasma equilibrium is then used to create plasma linear model on the next time step (Fig. 3). To verify the code, we have compared the equilibria obtained by the TOPMEC and equilibria reconstructed from experimental data by the FCDI (Flux and Current Distribution Identification) code (Mitrishkin et al., 2017b). As shown in Fig. 1b the resulting plasma shapes are practically identical.

\[
\frac{d}{dr} - \frac{1}{r} \frac{d}{d\theta} \Phi + \frac{d^2}{dz^2} \Phi = -8 \pi \mu_r r J = -8 \pi^2 \mu_r r^2 \frac{d}{d\Psi} p - 2 \pi^2 \frac{d}{d\Psi} F^2
\]

where \( J \) is the toroidal plasma current density distribution. The coefficients of polynomials are calculated from boundary conditions on plasma current density, total plasma current, and plasma parameters namely the poloidal beta \( \beta_p \) and the normalized plasma internal inductance \( \lambda \) specified by the user. The updated plasma equilibrium is then used to create plasma linear model on the next time step (Fig. 3). To verify the code, we have compared the equilibria obtained by the TOPMEC and equilibria reconstructed from experimental data by the FCDI (Flux and Current Distribution Identification) code (Mitrishkin et al., 2017b). As shown in Fig. 1b the resulting plasma shapes are practically identical.

2.3 Magnetic plasma shape control system with the TOPMEC

Blocks marked in black in Fig. 2 were tested on the base of experiments on the Globus-M tokamak, while blocks marked in red were implemented only in the simulation environment with usage of LT1 (Linear time-invariant) and LPV (Linear parameter-varying) plasma models (Mitrishkin et al., 2017a; Mitrishkin, et al., 2017b). In the case of using the TOPMEC as the controlled plant model, when each new SISO loop is added (starting from the CS, then the vertical position Z control loop, then horizontal \( R \), and then the PF-coils, one by one), the parameters of linear plasma models generated by the TOPMEC are being changed. For example, the unstable eigenvalue of the matrix \( A \) of the state space representation of the linear plasma model could change considerably. Therefore, for modeling purposes, the \( C_2, C_8, C_{CS}, K_{CS}, \) and \( C_P \) controllers were synthesized, given the modified plasma model due to the previous control loop and using the QFT methodology (Garcia-Sanz, 2017; http://codypower.com), which is suitable for the synthesis of robust SISO controllers for the set of linear models with uncertainties. In QFT frequency analysis, a set of curves (QFT bounds) are considered on the Nichols chart for a finite number of representative frequencies (Fig. 4). These curves describe the frequency constraints for a closed-loop system at each characteristic frequency taking into account various specifications (stability specification, reference tracking specification, etc.) for the whole set of linear plasma models at once. This is the main advantage of the QFT methodology. It is only necessary to determine the nominal plant \( P_0(s) \) and perform tuning of the controller \( G(s) \) for the nominal open-loop transfer function \( L_0(s) = P_0(s)G(s) \) without violation of the frequency constraints.

![Fig. 4. Loop shaping procedure for the controller of the plasma for (a) vertical position Z, (b) horizontal position R with QFT bounds for characteristic frequencies (rad/s), which are marked with figures.](image)

The outer cascade of plasma shape control (Fig. 2) is based on the modified isoflux control technique (Mitrishkin et al., 2017a) and incorporates the Improved Moving Filaments method (Mitrishkin et al., 2019) for plasma equilibrium reconstruction, which has sufficient accuracy and speed of response. The Improved Moving Filaments algorithm calculates \( r, z \)-components of the magnetic field \( B_r \) and \( B_z \) in the desirable X-point location, as well as the poloidal flux differences between the X-point and points of a desirable location of plasma boundary. The objective of the MIMO controller \( C_{shape} \), which is designed by the well-suited for this task robust loop-shaping approach (McFarlane and Glover,
1990; Skogestad and Postlethwaite, 2005; Mitrishkin et al., 2017a) is to reduce these values to zero.

2.4 Training, validation and test sets

To train the ANN, the training data set was compiled based on the TOPMEC of the plasma. For this, 75 plasma discharges of various durations were generated with different references for the vertical $Z_{ref}$, horizontal $R_{ref}$ position of the plasma, and plasma current $I_{ref}$ using the plasma shape control system shown in Fig. 2. For modelling the plasma shape control system, the parameters $\beta_p$ and $l_r$ were chosen equal to 0.405 and 1.95, respectively. The outer plasma shape control loop was turned on at some time moment; therefore, the data includes both the limiter phase and the divertor phase of plasma discharge and besides both with and without the plasma shape control loop. In total, the database contains 255024 time slices 10 percent of which was used for validation and 10 percent for testing. The task of the ANN is a nonlinear mapping of the input vector $\hat{x} = (I_p, I_{HFC}, I_{VFC}, I_{PF-CC}, \Psi_M, I_{yy}) \in \mathbb{R}^{341}$, where $I_p$ is the plasma current, $I_{HFC}, I_{VFC}, I_{PF-CC}$ are currents in the CS&PF coils, $\Psi_M$ is the poloidal fluxes measured by 21 flux loops, $I_{yy}$ is the vacuum vessel current to the output vector $\hat{y} = (B_x, B_y, \theta_{\Psi_1}, \theta_{\Psi_2}) \in \mathbb{R}^{41}$, where $B_x, B_y$ are components of the magnetic field in the desirable X-point location, $\theta_{\Psi_1} = \Psi_x - \Psi_1, \theta_{\Psi_2} = \Psi_x - \Psi_2, \Psi_1$ and $\Psi_2$ are shown in Fig. 1b. In order to increase the convergence rate of the ANN learning algorithm, both the input and output data were normalized: $\hat{x}' = (\hat{x} - M_x) / \sqrt{D_x}$, $\hat{y}' = (\hat{y} - M_y) / \sqrt{D_y}$, where $M_x, M_y, D_x, D_y$ are the sample mean and the sample variance of $\hat{x}$ and $\hat{y}$, respectively.

3. THE MULTILAYER ANN WITH BACK-PROPAGATION TRAINING METHOD AS THE ISOFLUX SIGNALS ESTIMATOR

To train the ANN, an architecture with one hidden layer with ReLU activation functions (Lu et al., 2017) and with a different number of neurons was chosen (Fig. 5a).

In order to reduce the chances of ANN overfitting, we also tested ANNs with a dropout layer with the fraction rate equal to 0.5 (Fig. 5b). For training the ANNs, the ADAM algorithm (Kingma and Ba, 2014) with 30 epochs was used, and the batch size was 4000. Tables 1 and 2 show the results of training the ANNs in terms of the Mean Square Error (MSE), which was also a loss function for this multidimensional regression task. As expected, the accuracy in terms of MSE increases with the number of neurons in the hidden layer for all parts of the dataset, which means that the overfitting of the ANNs does not occur during the 30 epochs of learning.

**Table 1. MSE for the ANNs with one hidden layer**

| The number of neurons | Train data | Validation data | Test data |
|-----------------------|------------|-----------------|-----------|
| 32                    | 0.0177     | 0.0342          | 0.0256    |
| 64                    | 0.0082     | 0.0175          | 0.0151    |
| 128                   | 0.0042     | 0.0074          | 0.0066    |
| 256                   | 0.0022     | 0.0052          | 0.0038    |
| 512                   | 0.0016     | 0.0042          | 0.0030    |
| 1024                  | 0.0013     | 0.0038          | 0.0026    |
| 2048                  | 0.0012     | 0.0034          | 0.0024    |

**Table 2. MSE for the ANNs with one hidden layer and the additional dropout layer**

| The number of neurons | Train data | Validation data | Test data |
|-----------------------|------------|-----------------|-----------|
| 32                    | 0.1722     | 0.0778          | 0.0615    |
| 64                    | 0.1018     | 0.0434          | 0.0384    |
| 128                   | 0.0585     | 0.0263          | 0.0221    |
| 256                   | 0.0319     | 0.0107          | 0.0105    |
| 512                   | 0.0172     | 0.0066          | 0.0057    |
| 1024                  | 0.0101     | 0.0042          | 0.0037    |
| 2048                  | 0.0063     | 0.0029          | 0.0031    |

4. SIMULATION RESULTS OF THE PLASMA SHAPE CONTROL SYSTEM

The plasma shape control system with the ANN as the isoflux signals estimator in the plasma shape control loop was simulated in the Matlab/Simulink environment. Figs. 6 and 7 show a comparison of the magnetic output ANN signals used to control the shape together with the robust MIMO controller (see Section 2.3) with the ground truth signals from the TOPMEC. The plasma shape control loop was turned on at 0.18 s of the plasma discharge. Here, the ANN with 1024 neurons from Table 1 was used. The coordinates of the desired location of X-point, point 1 and 2 (Fig. 1b) for calculation of $\Psi_1, \Psi_2$ were the coordinates (0.30, 1.24), (0.57, 0.81), (0.30, 0.44), respectively.

Fig. 8 depicts the evolution of the plasma separatrix during the plasma discharge with the ANN in the feedback loop and with reference values $Z_{ref} = 0.01$ m, $R_{ref} = 0.365$ m, $I_{ref} = 150$ kA. As can be seen from the figures, the ANN with high accuracy recovers the output signals, and the robust
margin of the MIMO shape controller is enough to control the plasma shape.

![Fig. 6. Plasma shape control: $B_r$, $B_z$ at X-point are obtained by the ANN and the TOPMEC.](image)

![Fig. 7. Plasma shape control: $\partial \Psi_1 = \Psi_x - \Psi_1$, $\partial \Psi_2 = \Psi_x - \Psi_2$, $\Psi_x$ is the flux at X-point, $\Psi_1$, $\Psi_2$ are fluxes at the separatrix (Fig. 1b) from the ANN and the TOPMEC.](image)

![Fig. 8. Evolution of the plasma separatrix during the plasma discharge from the TOPMEC code (red lines) and the Improved Moving Filament Code (blue lines) with 5 filaments at (a) 0.17 s, (b) 0.18 s, (c) 0.19 s, (d) 0.20 s. The coordinates of the desired location of the plasma separatrix are marked with black asterisks.](image)

5. CONCLUSIONS

In the paper, the possibility of using the ANN to control the plasma shape for tokamaks as the algorithm for plasma equilibrium reconstruction, which is included in the feedback together with the MIMO robust controller, is demonstrated. A further direction of work will be related to the application of this approach in real experiments for the Globus-M2 tokamak.

The results of the work done are in line with the trend to develop plasma magnetic control systems for effective control of plasma position, current, and shape in vertically elongated tokamaks with unstable plasma that has been shown in the survey of Mitrishkin et al. (2018c). The plasma magnetic control systems are being evolved in the class of digital hierarchical multivariable parameter-varying systems having various directions and principles, which try to improve MIMO system robust performance and stability margins by different approaches. This is because plasma in any elongated tokamak is an extremely complicated plant under control that requires to control collectives of nano-particles specifically ions and electrons of the plasma.

ACKNOWLEDGMENTS

The work was supported by Russian Science Foundation, grant No 17-19-01022.

REFERENCES

Ariola, M. and Pironi, A. (2016). Magnetic Control of Tokamak Plasmas. 2nd ed. Cham: Springer International Publishing, Switzerland. doi:10.1007/978-3-319-29890-0

Bishop, C., Haynes, P., Smith, M., Todd, T., and Trotman, D. (1995). Real-Time Control of a Tokamak Plasma Using Neural Networks. Neural Computation, 7(1), pp. 206-217. doi:10.1162/neco.1995.7.1.206

Braams, B., Jilge, W., and Lackner, K. (1986). Fast determination of plasma parameters through function parametrization. Nuclear Fusion, 26(6), pp. 699-708. doi:10.1088/0029-5515/26/6/001

Chuyanov, V. and Gryaznevich, M. (2017). Modular fusion power plant. Fusion Engineering and Design, 122, pp. 238-252. doi:10.1016/j.fusengdes.2017.07.017

Coccorese, E., Morabito, C., and Martone, R. (1994). Identification of noncircular plasma equilibria using a neural network approach. Nuclear Fusion, 34(10), pp. 1349-1363. doi:10.1088/0029-5515/34/10/I05

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems, 2(4), p. 303–314. doi:10.1007/BF02551274

Garcia-Sanz, M. (2017). Robust Control Engineering: Practical QFT Solutions. Boca Raton: CRC Press.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. MIT Press.

Greco, A., Mammone, N., Morabito, F., and Versaci, M. (2007). Artificial Neural Networks for Classifying Magnetic Measurements in Tokamak Reactors. International Journal of Physical and Mathematical Sciences, 1(7), pp. 313-317. doi:10.5281/zenodo.1327638
