BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 69(6), 2021, Article number: e139390
DOI: 10.24425/bpasts.2021.139390

Experimental verification of H_∞ control with examples of the movement of a wheeled robot

Zenon HENDZEL and Paweł PENAR*

Department of Applied Mechanics and Robotics, Faculty of Mechanical Engineering and Aeronautics, Rzeszów University of Technology, ul. Powstańców Warszawy 12, 35-959 Rzeszów, Poland

Abstract. The paper presents the results of experimental verification on using a zero-sum differential game and H_∞ control in the problems of tracking and stabilizing motion of a wheeled mobile robot (WMR). It is a new approach to the synthesis of input-output systems based on the theory of dissipative systems in the sense of the possibility of their practical application. This paper expands upon the problem of optimal control of a nonlinear, nonholonomic wheeled mobile robot by including the reduced impact of changing operating condtions and possible disturbances of the robot's complex motion. The proposed approach is based on the H_∞ control theory and the control is generated by the neural approximation solution to the Hamilton-Jacobi-Isaacs equation. Our verification experiments confirm that the H_∞ condition is met for reduced impact of disturbances in the task of tracking and stabilizing the robot motion in the form of changing operating conditions and other disturbances, which made it possible to achieve high accuracy of motion.

Key words: differential game; H_∞ control; wheeled mobile robot.

1. INTRODUCTION

The planning of the motion of WMR is one of the problems of mobile robotics. The navigation of wheeled robots is done through various methods, divided into local and global ones, and their list referred to in papers [1, 2].

From a mechanical standpoint, algorithms are the higher layer in planning motion [3–5], generating the desired kinematic parameters of mobile robots' drive wheels, executed as trajectory tracking control. The selected algorithms implementing the higher control layer are: the potential field method [1], generating the optimal trajectory with neural networks (NN) [6] and kinematics control [5, 7]. Kinematics control is a simple method of executing the higher layer of point-to-point control, based on the WMR's kinematics. It is a stabilisation task in which the selected characteristic point of the WMR has to reach a specific set point in space. The robot frame's configuration can be arbitrary or specific in this task.

The generated kinematic parameters of WMR's wheels serve as the basis for determining WMR's tracking errors and generalized tracking errors. On their basis, trajectory tracking control is generated, which can be implemented as optimal. This issue is widely discussed in papers [8, 9], for example. Optimal control in non-linear dynamical system is reduced to the approximation of the solution to the Hamilton–Jacobi–Bellman (HJB) equation [8, 10]. This is a problem of approximative dynamic programming (ADP) [4, 8, 10–13]. Optimal control problem can be extended by taking disturbances into account. In this approach, the theory of differential games is applicable. In this problem, optimal controls are determined in the presence of the worst-case disturbances, so it is possible to set the H_∞ type control and ensure that the system is robust to the said disturbances, which can be understood as changing operating conditions of the controlled object. This approach results from the input-output control theory [14–17].

Determining the solution of a zero-sum differential game corresponds with determining the saddle point in the Hamilton-Jacobi-Isaacs (HJI) equation. Similarly to the HJB equation, the solution to the HJI equation is only possible in a linear case. An approximation solution is applied in a nonlinear case. Here, the ADP structure (composed of a critic approximating the value function and two actors) generates suboptimal solutions approximating optimal control and the worst-case disturbances, respectively. This approach is widely discussed in literature: [4, 18–31]

Paper [24] contains an actor-critic adaptive solution to the HJI equation, executed in real time. This case assumes the knowledge of the dynamics system and the law of adaptation of critic's NN weights is based on a modification of the Levenberg-Marquardt method. The requirement regarding the knowledge of the mathematical model for the dynamic system is also assumed in paper [4], in which the iterative recursive least square (RLS) algorithm was used to approximate the solution of a zero-sum differential game.

The condition regarding the knowledge of the model may not always be satisfied, and therefore papers [23, 26–28] present adaptation algorithms based on data from the system output. Paper [31] shows a synthesis of H_∞ control with the tracking task for a system of unknown dynamics. Papers [4, 21] present the use of a zero-sum differential game in tracking control of a

*e-mail: ppenar@prz.edu.pl

Manuscript submitted 2021-04-07, revised 2021-08-31, initially accepted for publication 2021-09-02, published in December 2021

wheeled mobile robot. Another problem concerning the WMR, which involves differential zero-sum games, is the stabilisation task, as described in paper [22]. In paper [30], ADP methods were used to solve a zero-sum differential game for the control of a modular manipulating robot.

Recognising the small number of applicable papers, this paper pays special attention to this aspect through experimental verification of the differential game theory and H_∞ control. It concerns the WMR's behavioral control in a go-to-goal task under changing operating conditions. The results of experimental tests confirm the ability to apply differential game theory and H_∞ control for practical solutions, which extends the possibilities of using advanced control methods in practical applications.

The paper is divided into several sections. Section 2 formulates the problem of the theory of zero-sum differential games, resulting from the dissipativity theory. To solve it, adaptive dynamic programming was used, which approximates the solution of the HJI equation using the neural network in the critic's structure. The critic approximates the value functions and enables the determination of control and disturbances signals in real time. Using the iterative recursive least squares procedure, the NN weights of the critic can be determined in real time (without pre-learning). Section 3 includes the WMR's kinematics equations and formulates the problem of go-to-goal control in terms of kinematics. The WMR's dynamics equations are presented and the WMR tracking task is formulated in Section 4. Section 5 contains the results of experimental verification and provides the application aspect to the zero-sum differential game theory. Additionally, it demonstrates that the H_∞ condition is met, so the resulting control is robust to changing operating conditions. The last section is a summary of the findings.

2. PROBLEM FORMULATION

This section contains information on the dissipativity theory, which is connected with input-state-output signals. Function $s_v(\mathbf{d}(t),\mathbf{z}(t))$ is called supply rate, while accumulated energy function $V(\mathbf{x})$ is a state function [14, 15, 32–34]. In the analysis and synthesis of the dissipativity properties, we focus on signals of finite energy. These signals belong to L_2 spaces.

Definition 1. [17]: The function $f(t)\colon [0,\infty)\to\mathbb{R}^n$ belongs to $L_p[0,\infty)$ if $\int_0^\infty \|f(t)\|_p^p)^{1/p} < \infty$ where $\|f(t)\|_p$ is p-norm vector.

In particular: $f(t)\in L_2[0,\infty]$ if $\int_0^\infty \|f(t)\|^2)^{1/2} < \infty$.

Consider the non-linear dynamical system given by

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})\mathbf{u} + k(\mathbf{x})\mathbf{d}, \tag{1}$$

where $\mathbf{x}\in\mathbb{R}^n$ is the state vector, $\mathbf{u}\in\mathbb{R}^m$ is the control vector and $\mathbf{d}\in\mathbb{R}^q$ is the vector of parametric and structural disturbances of a non-linear dynamical system. Functions $f(.)\in\mathbb{R}^n$, $g(.)\in\mathbb{R}^{n\times m}$, $k(.)\in\mathbb{R}^{n\times q}$ with appropriate dimensions are locally Lipschitzian in the compact set $\Omega=\{\mathbf{x}\in\mathbb{R}^n\colon \|\mathbf{x}\|\le m_x<\infty\}$, where m_x is constant, $f(0)=0$, so $\mathbf{x}=0$ is the point of equilibrium.

There are two outputs associated with the dynamic system (1). The first is

$$\mathbf{y} = h_1(\mathbf{x}), \tag{2}$$

where $\mathbf{y}\in\mathbb{R}^p$ is the measured output, while the other is

$$\mathbf{z} = h_2(\mathbf{x},\mathbf{u}), \tag{3}$$

where $\mathbf{z}\in\mathbb{R}^r$ is the controlled changing output (tracking errors, control cost).

The non-linear dynamical system of (1) is dissipative when the dissipative inequality is satisfied as

$$0 \le V(\mathbf{x}(\infty)) - V(\mathbf{x}(0)) \le \int_0^\infty s_v(\mathbf{d},\mathbf{z})\,\mathrm{d}t, \tag{4}$$

where $V>0$ is the accumulated energy and $\int_0^\infty s_v(\mathbf{d},\mathbf{z})\,\mathrm{d}t$ is the supplied energy [17, 34, 35].

The choice of a specific form of the supplied rate s_v function enables the synthesis of real control systems, while considering signals of finite energy, described in the L_2 space [34]. In this assumption:

$$s_v(\mathbf{d},\mathbf{z}) = \gamma^2\|\mathbf{d}\|^2 - \|\mathbf{z}\|^2, \tag{5}$$

$\|.\|$ is the vector norm, $\gamma>0$ is the design parameter. Substituting equation (5) into the inequality (4) and assuming $V(\mathbf{x}(0))=0$ has the following result:

$$\int_0^\infty \|\mathbf{z}\|^2\,\mathrm{d}t \le \int_0^\infty \gamma^2\|\mathbf{d}\|^2\,\mathrm{d}t. \tag{6}$$

Dynamic system (1) is dissipative with the L_2 gain lesser or equal to γ if inequality (6) is satisfied. Satisfying inequality (6) means that the dynamic system of (1) is dissipative and has the L_2 gain lower than or equal to γ for $\mathbf{d}\in L_2[0,\infty)$, which means that the dynamic system is stable [17, 34]. Satisfying inequality (6) ensures the stability of a dynamic system (1). Assuming the worst-case disturbances and minimising the L_2 gain for the dynamic system of (1) with a feedback loop leads to obtaining optimal H_∞ control.

2.1. H_∞ control

As indicated in Section 2 and in papers [14, 34], H_∞ control is a control of a non-linear dynamical system of the form (1), which minimizes the L_2 gain. The synthesis of H_∞ control leads to solution to solving a zero-sum differential game. This is equivalent to solving the HJI equation, which is a specific type of the HJB equation.

To define the H_∞ control, a dynamic system of the form (1) was introduced in paper [14] with full access to state variables. Optimal H_∞ control means determining control $\mathbf{u}^*$, for which

the $L_2[0,\infty]$ gain (between the disturbances $\mathbf{d}$ and the controlled variable output $\mathbf{z}$) i.e

$$\gamma^2 = \frac{\int_0^\infty \|\mathbf{z}\|^2 \,\mathrm{d}t}{\int_0^\infty \|\mathbf{d}\|^2 \,\mathrm{d}t} \tag{7}$$

is minimised. Due to its difficulty, the problem of optimal H_∞ control is replaced with the search for suboptimal H_∞ control. Such an approach means determining the control that cause the gain of the closed system to be less than or equal to γ for the assumed disturbances. This is done by successively reducing the assumed gain γ, for which the $L_2[0,\infty]$ gain γ^* is less than or equal to γ [14]. Therefore, the H_∞ control problem is reduced to a set of the lowest value of γ, for which the inequality is satisfied

$$\gamma^* \leq \gamma. \tag{8}$$

This approach leads to solution a two-person zero-sum differential game that solves the minmax problem.

There is a non-linear dynamical system of (1) with the full access output $\mathbf{y}$ and the norm squared controlled variable output $\mathbf{z}$ in a form

$$\|\mathbf{z}\|^2 = \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u}, \tag{9}$$

where $\mathbf{Q},\ \mathbf{R} > 0$ are the design matrices of appropriate dimensionality. Dynamic system (1) is associated with value functions

$$V = \int_0^\infty \left[\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} - \gamma^2\|\mathbf{d}\|^2\right] \mathrm{d}t, \tag{10}$$

where control $\mathbf{u}$ and disturbances $\mathbf{d}$ are understood as minimizing and maximizing players, respectively. Based on the condition

$$\frac{\partial H}{\partial \mathbf{u}} = 0, \qquad \frac{\partial H}{\partial \mathbf{d}} = 0, \tag{11}$$

where H is the Hamiltonian

$$H(\mathbf{x},V,\mathbf{u},\mathbf{d}) = \frac{dV}{d\mathbf{x}}\left[f(\mathbf{x}) + g(\mathbf{x})\mathbf{u} + k(\mathbf{x})\mathbf{d}\right] - \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} - \gamma^2\|\mathbf{d}\|^2 \tag{12}$$

the following was determined:

$$\begin{aligned} \mathbf{u}^* &= -\frac{1}{2}\mathbf{R}^{-1}g(\mathbf{x})^T\frac{dV^*}{d\mathbf{x}}, \\ \mathbf{d}^* &= \frac{1}{2\gamma^2}k(\mathbf{x})^T\frac{dV^*}{d\mathbf{x}} \end{aligned} \tag{13}$$

where V^* is the optimal function of the value of (10), $\mathbf{u}^*$ is optimal control, and disturbances $\mathbf{d}^*$ is the worst-case disturbances. Signals $\mathbf{u}^*$, $\mathbf{d}^*$ satisfy the Nash inequality of the form [16]

$$V(\mathbf{x},\mathbf{u}^*,\mathbf{d}) \leq V^*(\mathbf{x},\mathbf{u}^*,\mathbf{d}^*) \leq V(\mathbf{x},\mathbf{u},\mathbf{d}^*) \tag{14}$$

and therefore they set the saddle point.

Inserting (13) into (12) resulted in the following:

$$\mathbf{x}^T\mathbf{Q}\mathbf{x} + \nabla V^{*T} f(\mathbf{x}) - \frac{1}{4}\nabla V^{*T} g(\mathbf{x})\mathbf{R}^{-1}g^T(\mathbf{x})\nabla V^* + \frac{1}{4\gamma^2}\nabla V^{*T}k(\mathbf{x})k^T(\mathbf{x})\nabla V^* = 0, \tag{15}$$

where $\nabla V^*(\mathbf{x}) = \dfrac{\mathrm{d}V^*}{\mathrm{d}\mathbf{x}}$. Equation (15) is a HJI equation. The ADP algorithm will be used to solve it.

2.2. Zero-sum differential game

In the case of non-linear dynamical system, ADP methods are used for solving a zero-sum differential game, and their critic structure generates the aproximation of the value function, while the actors generate the approximation of the player's signals.

It was assumed that the weights $\mathbf{W}$ in the critic neural network cause the value function to be as follows [24, 36]

$$V(\mathbf{x}) = \mathbf{W}^T\boldsymbol{\Psi}(\mathbf{x}) + \varepsilon(\mathbf{x}), \tag{16}$$

where $\boldsymbol{\Psi}\colon \mathbb{R}^n \to \mathbf{R}^N$ is the linearly independent vector of basic functions, N is the number of neurons and $\varepsilon(\mathbf{x})$ is the approximation error. Paper [24] shows that if the number of neurons $N \to \infty$, the approximation error is then $\varepsilon \to 0$ and $\nabla\varepsilon \to 0$.

Ideal weights $\mathbf{W}$ in the critic's neural network remain unknown. Therefore, the output of the critic's structure is the approximation of the value function, i.e. [24]

$$\hat{V}(\mathbf{x}) = \widehat{\mathbf{W}}^T\boldsymbol{\Psi}(\mathbf{x}), \tag{17}$$

where $\widehat{\mathbf{W}}$ is an approximation of ideal weights in the critic's NN, a subject to the adaptation procedure. Inserting (17) to the HJI equation of (12) resulted in the following:

$$H(\mathbf{x},\widehat{\mathbf{W}},\mathbf{u},\mathbf{d}) = \widehat{\mathbf{W}}^T\nabla\boldsymbol{\Psi}^T\left[f(\mathbf{x}) + g(\mathbf{x})\mathbf{u} + k(\mathbf{x})\mathbf{d}\right] + \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} - \gamma^2\mathbf{d}^T\mathbf{d} = \varepsilon_1. \tag{18}$$

where ε_1 is the estimation error. Inserting (17) to equation (13) resulted in the following:

$$\hat{\mathbf{u}} = -\frac{1}{2}\mathbf{R}^{-1}k(\mathbf{x})^T\nabla\boldsymbol{\Psi}(\mathbf{x})^T\widehat{\mathbf{W}}, \tag{19}$$

$$\hat{\mathbf{d}} = \frac{1}{2\gamma^2}g(\mathbf{x})^T\nabla\boldsymbol{\Psi}(\mathbf{x})^T\widehat{\mathbf{W}}. \tag{20}$$

Equations (19) and (20) are approximations of player's signals in a zero-sum differential game.

By choosing the iterative recursive least squares (RLS) method, the process of weight adaptation is in accordance with the law of adaptation for [10]

$$\dot{\widehat{\mathbf{W}}} = \xi\mathbf{P}\frac{\hat{\boldsymbol{\sigma}}_1}{1 + \nu\hat{\boldsymbol{\sigma}}_1^T\mathbf{P}\hat{\boldsymbol{\sigma}}_1}H(\mathbf{x},\widehat{\mathbf{W}},\hat{\mathbf{u}},\hat{\mathbf{d}}). \tag{21}$$

Coefficients $\xi, \nu > 0$ are design parameters, while $\mathbf{P}$ is a symmetric matrix based on equation

$$dot\mathbf{P} = \xi_P \mathbf{P} \frac{\hat{\boldsymbol{\sigma}}_1 \hat{\boldsymbol{\sigma}}_1^T}{1 + \nu \hat{\boldsymbol{\sigma}}_1^T \mathbf{P} \hat{\boldsymbol{\sigma}}_1} \mathbf{P}, \mathbf{P}(0) = \alpha_C \mathbf{I}, \quad \alpha_C \gg 0, \tag{22}$$

where $\xi_P > 0$ is a design parameter,

$$\hat{\boldsymbol{\sigma}} = \nabla \boldsymbol{\Psi}(\mathbf{x}) \left(f(\mathbf{x}) + g(\mathbf{x}) \hat{\mathbf{u}} + k(\mathbf{x}) \hat{\mathbf{d}} \right).$$

3. WMR'S KINEMATICS

The description of the kinematics for the selected WMR point with drive wheels 1 and 2 and supporting wheel 3 (Fig. 1) requires kinematics equations, which are described in detail in papers [5, 37, 38]. Wheels 1 and 2, with a radius of $r_1 = r_2 = r$, rotate around their axis, which does not change its position in relation to frame 4. The moving wheels are driven by separate drives consisting of electric motors and gears. Self-rotation angles for drive wheels are α_1 and α_2, respectively.

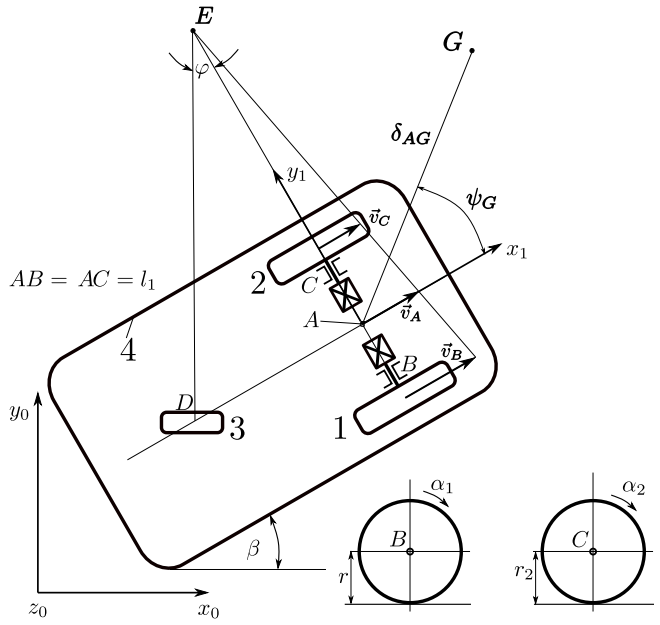

Fig. 1. Diagram of the Pioneer 2-DX two-wheel mobile robot in go-to-goal control

The angular velocity of drive wheels 1 and 2 is determined through the following equation [37]:

$$\dot{\alpha}_1 = \frac{v_A}{r} + \dot{\beta}\frac{l_1}{r}, \quad \dot{\alpha}_2 = \frac{v_A}{r} - \dot{\beta}\frac{l_1}{r}. \tag{23}$$

The mobile robot's basic problem is trajectory tracking control and stabilization control. The go-to-goal control task involves determining the control in which it is assumed that the selected WMR's point is to reach another point in the working space at any configuration of the robot's frame. In the fixed coordinate system x_0y_0, the WMR's positions are described by the coordinates defined as $\mathbf{q}_A = [x_A, y_A, \beta]^T$, where (x_A, y_A) are the coordinates of point A, while β is the rotation angle for the robot in fixed frame x_0y_0. The WMR's kinematics is described by the equation [38]

$$\begin{bmatrix} \dot{x}_A \\ \dot{y}_A \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} v_A^* \cos\beta & 0 \\ v_A^* \sin\beta & 0 \\ 0 & \omega^* \end{bmatrix} \begin{bmatrix} u_v \\ u_\beta \end{bmatrix}, \tag{24}$$

where v_A^*, ω^* is the maximum velocity of point A and the frame respectively, while $\mathbf{u}_B = [u_V, u_\beta]^T$ is the input (control) vector. A set configuration of WMR's point A was assumed

$$\mathbf{q}_G \in [x_G, y_G, 0]^T \tag{25}$$

and the actual and final configuration error

$$\tilde{\mathbf{q}} = \mathbf{q}_G - \mathbf{q}_A = \left[\tilde{x},\ \tilde{y},\ \tilde{\beta} \right]^T. \tag{26}$$

Error $\tilde{\mathbf{q}}$ can be defined in polar coordinates, i.e.:

$$\mathbf{q}_b = \begin{bmatrix} \delta_{AG} & \psi_G \end{bmatrix}^T, \tag{27}$$

where $\delta_{AG} = \sqrt{\tilde{x}^2 + \tilde{y}^2}$, $\psi_G = \arctan(\tilde{y}/\tilde{x}) - \beta$, $\psi_G \in [-\pi, \pi]$. If $\psi_G = 0$, then characteristic point A coincides with point G. Differentiating the error $\mathbf{q}_b$ vector resulted in the following

$$\begin{cases} \dot{\delta}_{AG} = -u_v \cos\psi_{AG}, \\ \dot{\psi}_G = \delta_{AG}^{-1} u_v \sin\psi_G - u_\beta. \end{cases} \tag{28}$$

To determine the control law $\mathbf{u}_B = [u_V, u_\beta]^T$ generating the WMR motion parameters, we have applied the Lapunov stability theory. Assuming the positively defined function [7]

$$V_L = 0.5\delta_{AG}^2 + 0.5\psi_G^2 \tag{29}$$

and differentiating equation (28), (29) resulted in

$$\dot{V}_L = -\delta_{AG} u_v \cos\psi_G + \psi_G (\delta_{AG}^{-1} u_v \sin\psi_G - u_\beta), \tag{30}$$

Assuming the control vector of

$$u_V = k_1 \delta_{AG} \cos\psi_G \tag{31}$$

and

$$u_\beta = k_1 \sin\psi_G \cos\psi_G + k_2 \psi_G, \tag{32}$$

where $k_1,\ k_2 > 0$, the derivative of Lapunov function (30) is as follows

$$\dot{V}_L = -k_1 \delta_{AG}^2 \cos^2\psi_G - k_2 \psi_G^2. \tag{33}$$

This means that $\dot{V}_L < 0$ for every $\delta_{AG},\ \psi_G \neq 0$ and the function of (29) is the Lapunov function, and the equilibrium state $\delta_{AG} = 0,\ \psi_G = 0$ is globally asymptotically stable. Based on the control synthesis, the distance between WMR's point A and the set point G is exponentially reduced to zero, while meeting condition $\psi_G = 0$ implies that point A of the robot is at point G.

The system of equations (24), (26), (27), (31), (32) enables determining the change of the linear velocity v_A of point A and the angular velocity of the robot frame $\dot{\beta}$. Knowing these values, we can determine other essential system motion parameters.

4. WMR'S DYNAMICS

The use of Lagrange's equations of the second kind to describe nonholonomic systems does not allow us to obtain a dynamics description suitable for the control problem, which is related to Lagrange's multipliers. Maggi's equations [39, 40] are an extension of Lagrangian formalism, in which the application of a projection operator (orthogonal complement matrix) eliminates the terms containing the multipliers. By applying Maggi's mathematical formalism, the dynamic equations of motion of the WMR can be expressed as [37, 38]:

$$\mathbf{M}\ddot{\boldsymbol{\alpha}} + \mathbf{C}(\dot{\boldsymbol{\alpha}})\dot{\boldsymbol{\alpha}} + \mathbf{F}(\dot{\boldsymbol{\alpha}}) + \Delta\mathbf{F}(\dot{\boldsymbol{\alpha}}) = \mathbf{u}, \tag{34}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]^T$ is the self-angle vector for the WMR's drive wheels, $\mathbf{u} = [M_1, M_2]^T$ is the vector of torque driving these wheels, and $\Delta\mathbf{F}(\dot{\boldsymbol{\alpha}})$ is the parametric and structural disturbances vector. Matrices $\mathbf{M}, \mathbf{C}$ and vector $\mathbf{F}$ are as follows

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} a_1 + a_2 + a_3 & a_1 - a_2 \\ a_1 - a_2 & a_1 + a_2 + a_3 \end{bmatrix}, \\ \mathbf{C}(\dot{\boldsymbol{\alpha}}) &= \begin{bmatrix} 0 & 2a_4(\dot{\alpha}_2 - \dot{\alpha}_1) \\ -2a_4(\dot{\alpha}_2 - \dot{\alpha}_1) & 0 \end{bmatrix}, \\ \mathbf{F}(\dot{\boldsymbol{\alpha}}) &= \begin{bmatrix} F_1(\dot{\alpha}_1) \\ F_2(\dot{\alpha}_2) \end{bmatrix}, \end{aligned} \tag{35}$$

where $\mathbf{F}(\dot{\boldsymbol{\alpha}})$ is the vector of friction of motion, $\mathbf{a} = [a_1, a_2, \ldots, a_6]^T$ are grouped parameters resulting from geometry, masses, robot mass distribution and friction of motion. Their values result from the parametric identification of the WMR model. Based on experimental tests, the friction of motion was approximated by

$$F_i(\dot{\alpha}_i) \approx a_{4+i} \frac{2 - 2e^{-c_F \dot{\alpha}_i}}{1 + e^{-c_F \dot{\alpha}_i}}, \quad i = 1, 2. \tag{36}$$

Parameters a_5, a_6 are interpreted as rolling resistance of wheels multiplied by the forces N_i (loads applied to corresponding wheels), while the remainder of the equation can be interpreted as an approximation of the signum function. The H_∞ control synthesis will concern the weakening of the impact of changing the friction of motion on the control accuracy by using ADP methods in verification tests in the go-to-goal control problem, the selected characteristic point of the WMR moves along the trajectory generated by the higher control layer in real time. The trajectory tracking error is defined as

$$\mathbf{e} = \boldsymbol{\alpha}_d - \boldsymbol{\alpha} \tag{37}$$

and the generalised error as

$$\mathbf{s} = \dot{\mathbf{e}} + \boldsymbol{\Lambda}\mathbf{e}, \tag{38}$$

where $\mathbf{s} = [s_1, s_2]^T$, $\boldsymbol{\alpha}_d$ is the desired self-turn angle of wheels, generated in the higher control layer in real time, while $\boldsymbol{\Lambda} > 0$ is the design parameter. The WMR's dynamics can be written in the generalised error space, i.e.

$$\mathbf{M}\dot{\mathbf{s}} = -\mathbf{C}(\dot{\boldsymbol{\alpha}})\mathbf{s} + \mathbf{f}_{WMR}(.) + \Delta\mathbf{F}(\dot{\boldsymbol{\alpha}}) - \mathbf{u}, \tag{39}$$

where $\mathbf{f}_{WMR}(.) = \mathbf{M}[\ddot{\boldsymbol{\alpha}}_d + \boldsymbol{\Lambda}\dot{\mathbf{e}}] + \mathbf{C}(\dot{\boldsymbol{\alpha}})[\dot{\boldsymbol{\alpha}}_d + \boldsymbol{\Lambda}\mathbf{e}] + \mathbf{F}(\dot{\boldsymbol{\alpha}})$. Relation (39) can be written in the form, as equation (1), i.e

$$\dot{\mathbf{s}} = f(\mathbf{s}) + g(\mathbf{s})\mathbf{u} + k(\mathbf{s})\Delta\mathbf{F}(\dot{\boldsymbol{\alpha}}), \tag{40}$$

where

$$\begin{aligned} f(\mathbf{s}) &= \mathbf{M}^{-1}\left[-\mathbf{C}(\dot{\boldsymbol{\alpha}})\mathbf{s} + \mathbf{f}_{WMR}(.)\right], \\ g(\mathbf{s}) &= -\mathbf{M}^{-1}, \qquad k(\mathbf{s}) = \mathbf{M}^{-1}. \end{aligned} \tag{41}$$

5. VERIFICATION TESTS ON OPTIMAL BEHAVIORAL CONTROL

A rapid prototyping test station using the dSpace card was used for verification tests [41]. The card generates a control (with torque unit) that is scaled by constant $1/19.7$ (19.7 – gear constant). In the case of using a WMR, the feedback is based on measurements from incremental encoders (500 PPR). Measuring the increase in the angle of rotation is the basis for determining the speed of the WMR driving wheels based on the equation

$$\dot{\alpha}_i = \pm \frac{2\pi}{500 \times 19,7 \times h} INC, \tag{42}$$

where h is discretization step and INC is the increase in the angle of rotation. Converted measurement is filtered by a filter whose transfer function is

$$F_\alpha(p) = \frac{1}{p^2 + 0.03p + 0.0001}, \tag{43}$$

where p is Laplace operator. The angle of rotation is determined by the numerical integration (Fig. 2).

The Pionier 2-DX Wheeled Mobile Robot [42] and a test station consisting of a dSpace signal processor, a PC computer with Matlab/Simulink software and a WMR with a power system (Fig. 3) were used to verify the simulation of optimal adaptive control [4, 22].

5.1. Go-to-goal task

The value function resulting from the dissipativity theory (point 2) is as follows

$$V(\mathbf{s}) = \int_0^{t_k} \left[\mathbf{s}^T\mathbf{Q}\mathbf{s} + \mathbf{u}^T\mathbf{R}\mathbf{u} - \gamma\Delta\mathbf{F}^T\Delta\mathbf{F}\right]. \tag{44}$$

To apply adaptive dynamic programming algorithms, the aproximation of the value function of

$$\hat{V} = \widehat{\mathbf{W}}^T\boldsymbol{\Psi}(\mathbf{s}) \tag{45}$$

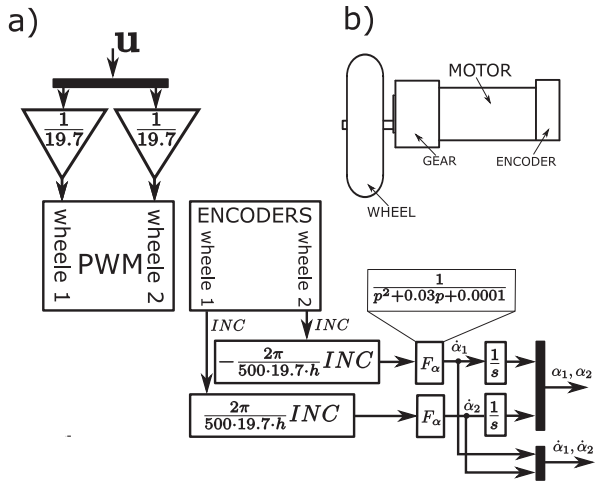

Fig. 2. A WMR control and kinematics parameter measurement scheme (a) WMR's drive module (b)

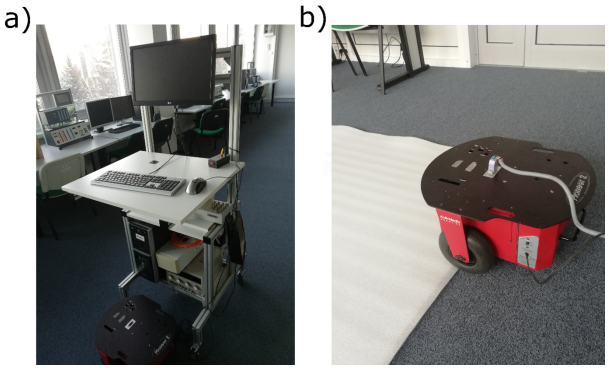

Fig. 3. Quick design test station (a) and the WMR entering a different surface (b)

was adopted, where the basic function vector is $\boldsymbol{\Psi}(\mathbf{s}) = [s_1^2, s_1 s_2, s_2^2]^T$ and $\widehat{\mathbf{W}} \in \mathbb{R}_{3\times 1}$.

The approximation of the minimizing and maximising player signals according to the equations (19) and (20) for the WMR is as follows

$$\begin{aligned} \hat{\mathbf{u}} &= \frac{1}{2}\mathbf{R}^{-1} g(\mathbf{s})^T \Delta\boldsymbol{\Psi}(\mathbf{s})\widehat{\mathbf{W}}, \\ \Delta\hat{\mathbf{F}} &= \frac{1}{2\gamma^2} k(\mathbf{s})^T \Delta\boldsymbol{\Psi}(\mathbf{s})\widehat{\mathbf{W}}. \end{aligned} \tag{46}$$

Using the iterative RLS method of (21), the law of weight adaptation in the case of behavioral control is as follows

$$\dot{\widehat{\mathbf{W}}} = \xi \mathbf{P} \frac{\hat{\boldsymbol{\sigma}}_1}{1 + \nu \hat{\boldsymbol{\sigma}}_1^T \mathbf{P} \hat{\boldsymbol{\sigma}}_1} H(\mathbf{s}, \widehat{\mathbf{W}}, \hat{\mathbf{u}}, \Delta\hat{\mathbf{F}}) \tag{47}$$

whereas

$$\hat{\boldsymbol{\sigma}} = \Delta\boldsymbol{\Psi}(\mathbf{s}) \left[f(\mathbf{s}) + g(\mathbf{s})\hat{\mathbf{u}} + k(\mathbf{s})\Delta\hat{\mathbf{F}} \right] \tag{48}$$

and

$$\begin{aligned} H(\mathbf{s}, \widehat{\mathbf{W}}, \hat{\mathbf{u}}, \Delta\hat{\mathbf{F}}) = \widehat{\mathbf{W}}^T \Delta\boldsymbol{\Psi}(\mathbf{s}) \left[f(\mathbf{s}) + g(\mathbf{s})\hat{\mathbf{u}} + k(\mathbf{s})\Delta\hat{\mathbf{F}} \right] \\ + \mathbf{s}^T \mathbf{Q}\mathbf{s} + \hat{\mathbf{u}}^T \mathbf{R}\hat{\mathbf{u}} - \gamma \Delta\hat{\mathbf{F}}^T \Delta\hat{\mathbf{F}}, \end{aligned} \tag{49}$$

The verification of this task was carried out according to the diagram shown in Fig. 4

Figure 4 shows a go-to-goal control diagram. In the proposed approach, the determination of the set parameters for the motion of the WMR's drive wheels is based on the current position and orientation of the WMR in relation to the selected endpoint G. The determined generalised following error **s** is the basis for the suboptimal tracking control, resulting from an approximation solution to the zero-sum differential game.

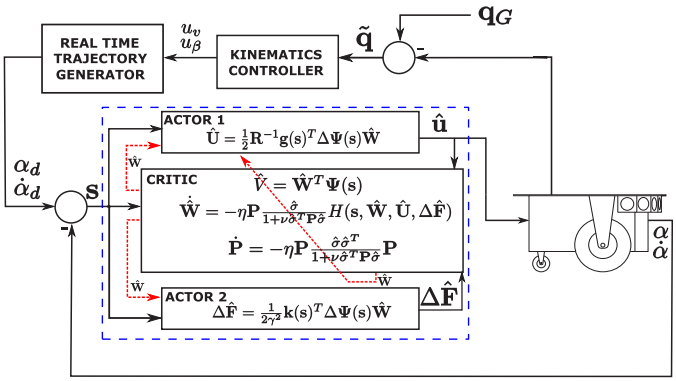

Fig. 4. Diagram of the goal-to-goal task implementation

Using the diagram in Fig. 4, verification tests of optimal behavioral control under changing operating conditions were performed. The higher control layer performed the task of WMR's movement from point $(0,0)$ to points $G_1(2,5), G_2(2,-3), G_3(4,0)$. To take into account the characteristic phases of the WMR's motion, i.e. acceleration and deceleration, velocity v_A^* is modified by the velocity profile of

$$v_A = v_A^* \left(\frac{1}{1+e^{-c(t-b_1)}} - \frac{1}{1+e^{-c(t-b_2)}} \right), \tag{50}$$

where c is the design parameter, b_1, b_2 are the starting and braking phases, respectively. The braking phase is determined by associating it with the WMR reaching the target point. If at time t_G the WMR's position is in the proximity ε_G of the target point G, then $b_2 = t_G + 1$. The verification assumes a number of constants, the values of which are summarized in Table 1, and the law of adaptation for matrix $\mathbf{P}$ was modified in the following way:

$$\dot{\mathbf{P}} = \begin{cases} \xi_P \mathbf{P} \dfrac{\hat{\boldsymbol{\sigma}}_1 \hat{\boldsymbol{\sigma}}_1^T}{1 + \nu \hat{\boldsymbol{\sigma}}_1^T \mathbf{P} \hat{\boldsymbol{\sigma}}_1} \mathbf{P}, & \|\mathbf{P}\| > P_0, \\ \mathbf{0}_{3\times 3}, & \|\mathbf{P}\| \le P_0. \end{cases} \tag{51}$$

This modification helps reduce the impact of disturbances during the weight $\widehat{\mathbf{W}}$ adaptation process.

Using parameters presented in Table 1 and the WMR's power system gain, the verification tests were performed for the go-to-goal task. During the test, adaptation for weights $\widehat{\mathbf{W}}$ was performed. The curves for the target point $G_1(2,5)$ are shown in Fig. 5.

Figure 5a,b shows the self-turn angle error and the angular velocity error for wheels in the verification. The highest value

Table 1
Constants used in verification

Constant	Sym.	Value
WMR dynamics parameters	$[\mathbf{a}_{1-6}, c_F]$	$\begin{bmatrix}0.0691\\0.0511\\0.0291\\0.0014\\4.3001\\4.3001\\0.226\end{bmatrix}$
Matrix proportional to the generalised error $\mathbf{s}$	$\mathbf{\Lambda}$	$5\mathbf{I}_{2\times2}$
WMR starting point	(x_A, y_A)	$(0,0)$
WMR end points	(x_G, y_G)	$\begin{bmatrix}(2,5)\\(2,-3)\\(4,0)\end{bmatrix}$
Maximum velocities	$v_A^*,\ \dot{\beta}^*$	$0.4, 0.7$
Proximity of point G	ε_G	0.03
Kinematics control gains	k_1, k_2	$0.4, 0.25$
Speed profile slope	c	10
Start phase time	b_1	5
Design matrix in value function	$\mathbf{Q},\ \mathbf{R}$	$\mathbf{I}_{2\times2},\ 0.5\mathbf{I}_{2\times2}$
H_∞ control gain	γ	11
Initial values of weights $\widehat{\mathbf{W}}(0)$, $\hat{\mathbf{P}}(0)$	$\begin{bmatrix}\widehat{\mathbf{W}}(0)\\\hat{\mathbf{P}}(0)\end{bmatrix}$	$\begin{bmatrix}0.05\mathbf{I}_{1\times3}\\100\mathbf{I}_{3\times3}\end{bmatrix}$
Adaptation gain	$\xi,\ \xi_P$	1.2, 80
Coefficient ν	ν	0.1
Constant matrix	P_0	5×10^{-4}
Sample rate (discrete step)	h	0.002
Verification time	t_k	50
Weight adaptation time	–	0–39

of the rotation angle tracking error occurs at the beginning of the motion, in the acceleration phase, and immediately after it. This is related to the sudden increase in the control value and inertia of the WMR. Then, tracking error $\mathbf{e}$ reaches the values of 0.0747 [rad] and 0.1015 [rad] (static error) for wheel 1 and 2, respectively, which is related to the end of the WMR's movement when it reaches proximity ε_G of point G.

Figure 5c shows the curve of control $\hat{\mathbf{u}}$, which has the dimension of torque. Its value is the highest in the start phase and immediately after it. This is related to the high value of the generalised error, as indicated by the nature of the curve, analogous to the tracking error curve. The response to the generated control aproximation curve is the disturbances aproximation $\Delta\hat{\mathbf{F}}$, as shown in Fig. 5d. Its curve is analogous to the aproximation of the minimizing player, but its amplitude is 250 times lower. This is due to the presence of the term $1/\gamma^2$ in equation (46). Because of the non-zero static error, the control value

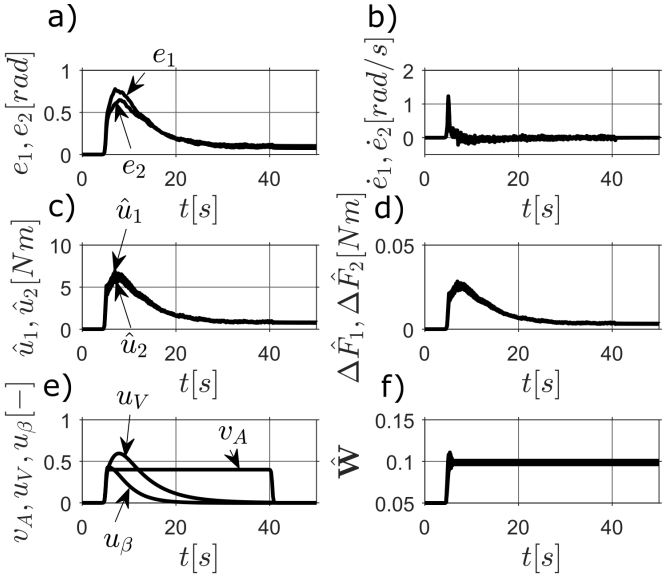

Fig. 5. Verification results from point G_1: curves of tracking errors (a, b), control (c) and disturbances (d) signals, curves of kinematics control and velocity profile (e) and curves of weight adaptation for the critic's adaptive structure (f)

after the end of the WMR's movement for wheels 1 and 2 is 0.74 and 0.81 [Nm], respectively. These values do not cause the WMR's movement. Figure 5e shows the curves of kinematics control signals generated by the higher control layer. In addition, Fig. 5e shows the velocity profile v_A, indicating that the WMR was in the proximity of point G at time $t = 40$ [s]. Figure 5f shows the curve of the adaptation process of the critic's weights $\widehat{\mathbf{W}}$. Their values change in the acceleration phase and remain constant at later stages of the WMR's movement.

Figure 6a shows the curve of value $\hat{V}$ function aproximation in relation to the zero-sum differential game. Its value in the case of the go-to-goal $G_1(2,5)$ task is the highest in the initial movement phase, which results from the fact that generalised error $\mathbf{s}$ and control $\hat{\mathbf{u}}$ reach high values at that time. During the later stages of movement, its value decreases. Due to the occurrence of a static error, its value is different than zero after the completed movement. Figure 6b shows the trajectory of the characteristic point along which the WMR's point A moved to perform the task of driving from point $(0,0)$ to point $G_1 - G_3$. The position the WMR reached for the go-to-goal $G_1 - G_3$ task

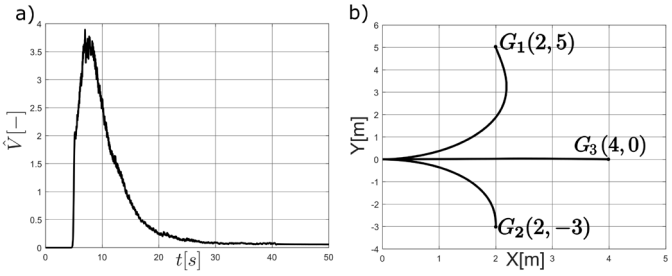

Fig. 6. Aproximation of the value function in the go-to-goal $G_1(2,5)$ task (a), trajectories of the characteristic point A of the WMR, obtained during the verification on a real object for target points $G_1 - G_3$ (b)

is (2.0033, 4.9759), (2.0008, –2.9751), (3.9760, 0.0023), respectively.

The weights obtained in the adaptation process are used for H_∞ control, which is the second element of the verification test. The WMR drove onto an additional foam surface (Fig. 3), which caused a change in the friction of motion. The time in which the change of friction of motion occurs is evaluated based on curves of tracking errors and control. This way, the verification was supplemented with changing operating conditions in order to determine the H_∞ condition by evaluating the disturbances level.

Figure 7a shows the curve of the tracking error $\mathbf{e}$ obtained in the verification (for $G_1(2,5)$) with changing conditions. The change in operating conditions, related to the change of friction of motion, caused disturbances in the curve of errors, as it can be seen in the comparison of curves 7a, b and 5a, b. The tracking error determined for wheel 1 and wheel 2 in 40 [s] of movement is different. The same situation occurs in the case of the static error, which is 0.0628 [rad] and 0.1783 [rad] for wheel 1 and 2 respectively. Figure 7b presents the curve of the tracking error for the angular velocity of the WMR's wheels.

At $t = 9$ [s], there are disturbances related to the entry of the robot onto another surface and an increase in the friction of motion. Figure 7c shows the curve of control $\hat{\mathbf{u}}$, reflecting the change of the surface on which the WMR was moving. Entry on a different surface increases the amplitude of the control signal to aproximately 8.4 [Nm]. The remainder of the control process is analogous to movement without changes in the friction of motion, but its amplitude is higher. This difference is approximately 0.4 [Nm] and will be the basis for determining the disturbances aproximation in the condition related to the H_∞ control. After the end of the deceleration phase, the control $\hat{\mathbf{u}}$ is different from zero due to a static error. The impact of the surface change is also visible in the curve of disturbances approximation $\Delta\hat{\mathbf{F}}$ (Fig. 7d).

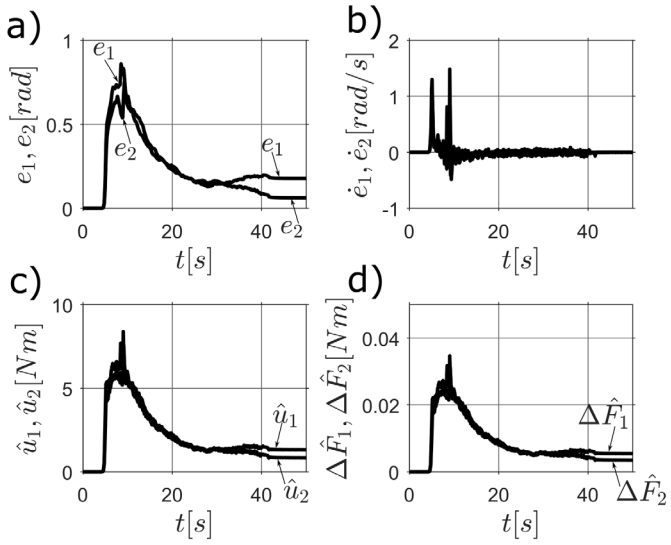

Fig. 7. Tracking errors (a, b), aproximation of optimal control (c) and disturbances (d) for the behavioral control problem under changing operating conditions

Based on the differences between the control signals, the disturbances aproximation $\Delta\bar{\mathbf{F}}$ was determined according to the equation

$$\Delta\bar{F}_i(\dot{\alpha}_i) = \begin{cases} 0 & t \le t_F \\ \hat{u}_{zi} - \hat{u}_i & t > t_F \end{cases}, \quad i = 1,2, \tag{52}$$

where $\hat{u}_{zi}$ is the control generated during the test using an additional foam surface, $\hat{u}_i$ is control for driving on the reference surface, and t_F is the time when the disturbances occurred; the time was read from the disturbance course. The $\Delta\bar{\mathbf{F}}$ curve is shown in Fig. 8a. In addition, knowing the time of the disturbances' occurence, it was assumed that their value for $t < 8.38$ [s] was equal to zero.

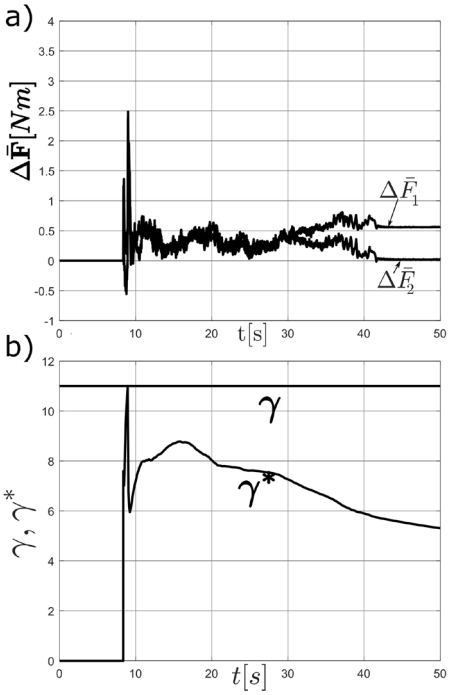

Fig. 8. Aproximation disturbances (a) and γ^* gain curve (b)

Based on disturbances aproximation, the H_∞ condition was determined, resulting from a relation (7), i.e:

$$\frac{\int_0^t \left[\mathbf{s}^T\mathbf{Q}\mathbf{s} + \hat{\mathbf{u}}^T\mathbf{R}\hat{\mathbf{u}}\right] \mathrm{d}t}{\int_0^t \left[\Delta\bar{\mathbf{F}}^T\Delta\bar{\mathbf{F}}\right] \mathrm{d}t} \le \gamma^2. \tag{53}$$

The obtained curve of γ^* gain for $G_1(2,5)$ is shown in Fig. 8b. It demonstrates that $\gamma^* \le \gamma$, which confirms the theoretical considerations regarding the input-output stability. The same condition is satisfied for target points G_2 and G_3 (Fig. 9). In these cases, the γ^* gain value is lower than aproximately 6 and 8, respectively, therefore one could attempt to search for a suboptimal control by reducing the value of γ.

Table 2
Values of control quality indicators in verification of the go-to-goal task

	wheel	$G_1(2,5)$		$G_2(2,-3)$		$G_3(4,0)$	
		RMSE		RMSE		RMSE	
		e_i	$\dot{e}_i$	e_i	$\dot{e}_i$	e_i	$\dot{e}_i$
Movement without disturbances	$i=1$	0.2889	0.1037	0.1261	0.0805	0.2064	0.2329
	$i=2$	0.2608	0.0857	0.2271	0.0957	0.2445	0.2487
Movement under changing operating conditions	$i=1$	0.3214	0.1191	0.1253	0.0938	0.1456	0.1517
	$i=2$	0.2766	0.1121	0.2616	0.1141	0.1546	0.1480

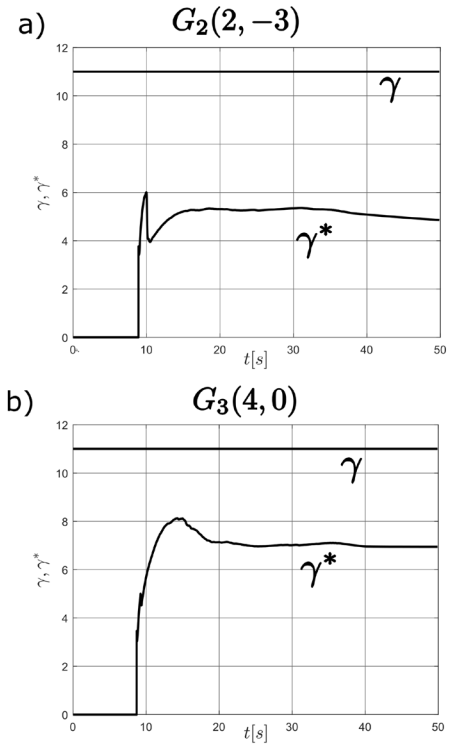

Fig. 9. Curve of γ^* gain for G_2 (a) and G_3 (b)

Quality indicators were used to determine the quality of control in the stabilization tasks involving target points $G_1 - G_3$. Their values are shown in Table 2 where

$$RMSE(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{54}$$

is the mean square error of variable x.

5.2. Trajectory tracking control task

In order to verify the H_∞ control in another control task, tests of the optimal control in tracking control of the robot were also performed. This task was verified according to the diagram shown in Fig. 10

The difference between the verified tasks is that in the first case the desired trajectory is generated in real time, but in the tracking control task the desired trajectory is generated offline.

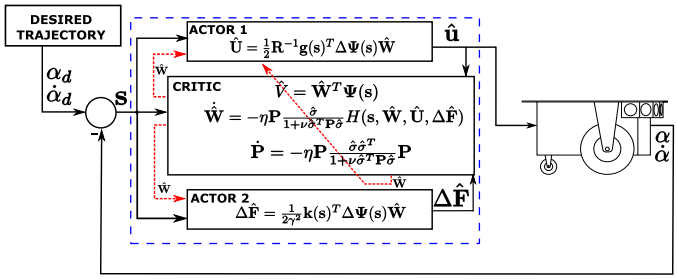

Fig. 10. Diagram of the tracking task implementation

In this task, the selected point if robot A is to move in a straight line. By carrying out the same course of the experiment, the obtained solutions are presented in Fig. 11.

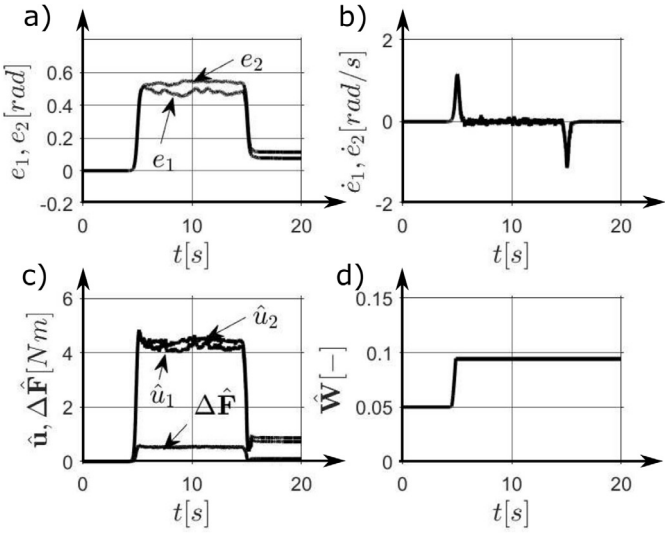

Fig. 11. Tracking errors (a, b), aproximation of optimal control (c) and disturbances (d) for the trajectory tracking control and curves of weight adaptation for the critic's adaptive structure (d)

Assuming selected constants from Table 1 and the values of the coefficients from the equation (50), i.e. $b_1 = 5, b_2 = 15[s], c = 10$, a verification test of the straight line trajectory tracking control was performed. Figure 11a shows the rotation angle tracking error obtained during the verification. The obtained curves are different in the case of wheels 1 and 2, which is caused by measuring noise and unmodelled dynamics phe-

nomena, including unevenness of the ground, and difference in dynamics in the actuators. The average rotation angle tracking error in the steady motion phase is aproximately 0.5 [rad]. Figure 11b shows the curves of the angular velocity trajectory tracking error. Its amplitude is the highest in the starting and braking phases (approximately 1 [rad/s]). The desired path of point A of the WMR is realized by suboptimal control, which is a solution to the differential zero-sum game, the curves of which is shown in Fig. 11c. In the starting phase, the control (a signal has a torque unit) is increased to a level above 4 [Nm], which is then maintained for the entire duration of the motion. After the end of the braking phase, the control is not equal to zero, which is related to a non-zero static error. The control level achieved, during this time does not cause the WMR's wheels to move. Control $\hat{\mathbf{u}}$ is also a minimizing player in a zero-sum differential game. The response to the minimizing player's signal is the maximizing player's signal. Its approximation generated by the actor-critic algorithm, is shown in Fig. 11c. In order to unify the scales, the $\Delta\hat{\mathbf{F}}$ score was multiplied by 30, and its value results from the WMR's dynamics. The signals $\hat{\mathbf{u}}$ and $\Delta\hat{\mathbf{F}}$ were determined on the basis of the estimation of weights of the critic's neural network $\widehat{\mathbf{W}}$, which approximates the adopted value functions. Their curves are presented in Fig. 11d. It shows that the increase in NN's weights is related to the starting phase, which is related to the increase in the generalized error $\mathbf{s}$ and the decrease in the values of the matrix elements $\mathbf{P}$. The value of weights stabilize at approximattely 0.1, which results from the dynamics of the WMR.

In order to determine the H_∞ control condition during the trajectory tracking control task, disturbances in the form of a change of friction of motion were introduced. Using the critic's weights obtained in the adaptation process (Fig. 11d), the trajectory tracking verification test was performed under changing operating conditions. Trajectory tracking errors are shown in Figs. 12a, b. Friction change is a characteristic point of these curves, occuring aproximattely at 9 [s]. The sudden increase

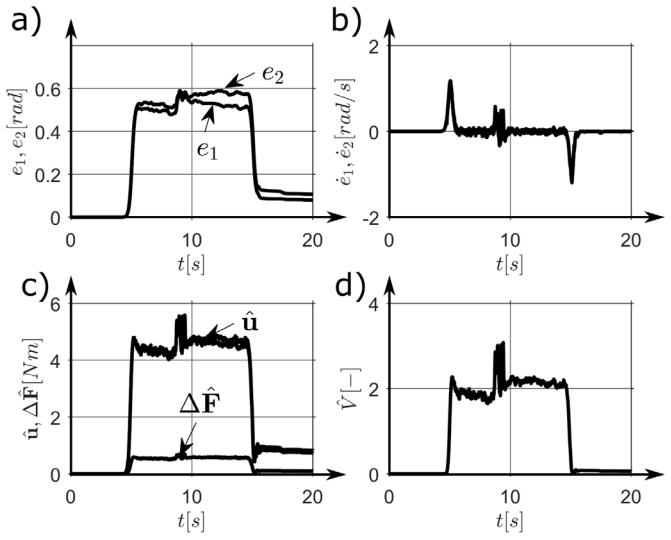

Fig. 12. Tracking errors (a,b), aproximation of optimal control (c) and disturbances (d) for the trajectory tracking control problem under changing operating conditions

in the tracking error's value results from the WMR crossing the contact edges of the two surfaces on which it was moving. Entering the foam surface causes an increase in the control amplitude, which is shown in Fig. 12c. In the phase of steady motion on the foam surface, the amplitude of the control signal is higher (approximattely 0.5 [Nm]), which is related to the change in friction. After the completion of the braking phase, the control signal is different than zero, which results from the presence of a static error apparent in the curves in Fig. 12a. Additionally, Fig. 12c shows the curves of the signal of the maximizing player $\Delta\hat{\mathbf{F}}$, the value of which was multiplied by 30 to unify the scale.

To determine the H_∞ control condition, it is necessary to know the generalized error, control evaluation, and the worst-case disturbances. During verification, the generalized error $\mathbf{s}$ and the estimation of the control signal $\hat{\mathbf{u}}$ are available for measurement. The immeasurable disturbances signal, which is the change in friction of motion, can be estimated on the basis of the control difference during verification with (Fig. 13b) and without disturbances (Fig. 13a), based on equation (52). From the equation (53), based on the disturbances estimation (Fig. 13c), the H_∞ condition was determined.

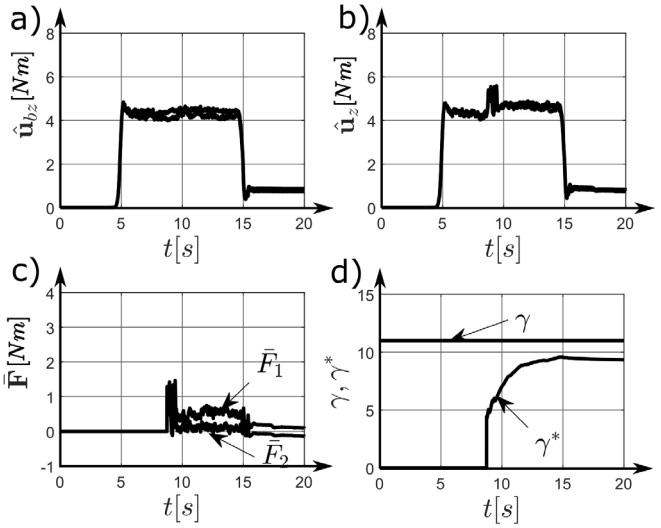

Fig. 13. Course control in tracking control problem with (a) and without (b) with changing operating conditions; course estimate disturbances (c) and γ, γ^* gain

Figure 13d shows the value of the gain γ^*, determined on the basis of the WMR's disturbances estimation, which proves that the H_∞ condition is met. This fact proves that the input-output system is stable in the presence of selected worst-case disturbances. Comparing the results of the experimental tests given in Tables 2 and 3, respectively, it can be stated that the errors obtained in the go-to-goal task are less severe than the errors obtained in the trajectory tracking control task. A more precise error analysis of the verified tasks was omitted, because special attention was paid to the possibility of using the theory of differential games as one of the methods of approximating the solution of the HJI equation and its practical implementation.

Table 3
Values of control quality indicators in the trajectory tracking verification task

	wheel	RMSE	
		e_i	$\dot{e}_i$
Movement without disturbances	$i=1$	0.3412	0.1885
	$i=2$	0.3818	0.1871
Movement under changing operating conditions	$i=1$	0.3959	0.2016
	$i=2$	0.3674	0.1942

6. SUMMARY

The paper presents the results of experimental verification of using a zero-sum differential game and H_∞ control in the problems of stabilizing motion and tracking control of a wheeled mobile robot. The concept of dissipativity was employed in the adopted solution. The ADP methods were used to obtain an approximation solution to the HJI equation. Taking into account changing operating conditions, understood as changing friction of motion, the H_∞ control condition was verified. Figures 8b, 9, and 13d demonstrate that the obtained H_∞ suboptimal control reduce the impact of the estimated disturbances on the control quality, characterised by quality indicators, shown in Tables 2 and 3. On the basis of the conducted experimental tests, and with taking into account the results obtained in the two cases of the robot's movement, the robustness of the adopted solution to changing working conditions of the robot was demonstrated. The obtained solutions suggest that it is possible to implement advanced optimal control methods in practical solutions.

REFERENCES

[1] B. Kovács, G. Szayer, F. Tajti, M. Burdelis, and P. Korondi, "A novel potential field method for path planning of mobile robots by adapting animal motion attributes," *Rob. Auton. Syst.*, vol. 82, pp. 24–34, 2016, doi: 10.1016/j.robot.2016.04.007.

[2] A. Pandey, "Mobile Robot Navigation and Obstacle Avoidance Techniques: A Review," *Int. Robotics Autom. J.*, vol. 2, no. 3, pp. 96–105, 2017, doi: 10.15406/iratj.2017.02.00023.

[3] R.C. Arkin, *Behavior-based robotics*. The MIT Press, 1998.

[4] M. Szuster and Z. Hendzel, *Intelligent Optimal Adaptive Control for Mechatronic Systems*. Springer, 2018.

[5] M.J. Giergiel, Z. Hendzel, and W. Żylski, *Modeling and control of mobile wheeled robots*. PWN, 2013, [in Polish].

[6] P. Bozek, Y.L. Karavaev, A.A. Ardentov, and K.S. Yefremov, "Neural network control of a wheeled mobile robot based on optimal trajectories," *Int. J. Adv. Rob. Syst.*, vol. 17, no. 2, pp. 1–10, 2020, doi: 10.1177/1729881420916077.

[7] P. Gierlak and Z. Hendzel, *Control of wheeled and manipulation robots*. Publishing House Rzeszow Univ. of Technology, 2011, [in Polish].

[8] B. Kiumarsi, K.G. Vamvoudakis, H. Modares, and F.L. Lewis, "Optimal and Autonomous Control Using Reinforcement Learning: A Survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2042–2062, 2018.

[9] F.L. Lewis, D. Vrabie, and V.L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.

[10] K.G. Vamvoudakis and F.L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[11] F.-Y.Wang, H. Zhang, and D. Liu, "Adaptive Dynamic Programming: An Introduction," *IEEE Comput. Intell. Mag.*, vol. 4, no. May, pp. 39–47, 2009.

[12] A.G. Barto, W. Powell, J. Si, and D.C. Wunsch, *Handbook of learning and approximate dynamic programming*. Wiley-IEEE Press, 2004.

[13] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li, *Adaptive Dynamic Programming with Applications in Optimal Control*. Springer, Advances in Industrial Control, 2017.

[14] A.J. van der Schaft, *L2-Gain and Passivity Techniques in Nonlinear Control*. Springer International Publishing, 2017.

[15] B. Brogliato, R. Lozano, B. Maschke, and O. Egeland, *Dissipative Systems Analysis and Control*. Springer-Verlag London, 2007.

[16] A.W. Starr and Y.C. Ho, "Nonzero-sum differential games," *J. Optim. Theory Appl.*, vol. 3, no. 3, pp. 184–206, 1969.

[17] M. Abu-Khalaf, J. Huang, and F.L. Lewis, *Nonlinear H2 Hinf Constrained Feedbacka Control*. Springer-Verlag London, 2006.

[18] D. Liu, H. Li, and D. Wang, "Neural-network-based zero-sum game for discrete-time nonlinear systems via iterative adaptive dynamic programming algorithm," *Neurocomputing*, vol. 110, pp. 92–100, 2013.

[19] C. Qin, H. Zhang, Y. Wang, and Y. Luo, "Neural network-based online Hinf control for discrete-time affine nonlinear system using adaptive dynamic programming," *Neurocomputing*, vol. 198, pp. 91–99, 2016.

[20] D. Liu, H. Li, and D. Wang, "Hinf control of unknown discrete-time nonlinear systems with control constraints using adaptive dynamic programming," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–6.

[21] Z. Hendzel and P. Penar, "Zero-Sum Differential Game in Wheeled Mobile Robot Control," Int. *Conf. Mechatron.*, vol. 934, pp. 151–161, 2017.

[22] Z. Hendzel, "Optimality in Control for Wheeled Robot," *Adv Intell. Syst. Comput.: Autom*. 2018, vol. 743, pp. 431–440, 2018.

[23] Y. Fu and T. Chai, "Online solution of two-player zero-sum games for continuous-time nonlinear systems with completely unknown dynamics," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 12, pp. 2577–2587, 2015.

[24] K.G. Vamvoudakis and F.L. Lewis, "Online solution of nonlinear two-player zero-sum games using synchronous policy iteration," *Int. J. Robust. Nonlinear Control*, vol. 22, pp. 1460–1483, 2012.

[25] S. Yasini, A. Karimpour, M.-B. Naghibi Sistani, and H. Modares, "Online concurrent reinforcement learning algorithm to solve two-player zero-sum games for partially unknown nonlinear continuous-time systems," *Int. J. Adapt Control Signal Process.*, vol. 29, no. 4, pp. 473–493, 2015.

[26] B. Luo, H.-N. Wu, and T. Huang, "Off-policy reinforcement learning for Hinf control design," *IEEE Trans. Cybern.*, vol. 45, no. 1, pp. 65–76, 2014.

[27] H.-N. Wu and B. Luo, "Neural Network Based Online Simultaneous Policy Update Algorithm for Solving the HJI Equation in Nonlinear Hinf Control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 12, pp. 1884–1895, 2012.

[28] Y. Zhu, D. Zhao, and X. Li, "Iterative adaptive dynamic programming for solving unknown nonlinear zero-sum game based on online data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 714–725, 2016.

[29] J. Zhao, M. Gan, and C. Zhang, "Event-triggered Hinf optimal control for continuous-time nonlinear systems using neurodynamic programming," *Neurocomputing*, vol. 360, pp. 14–24, 2019.

[30] B. Dong, T. An, F. Zhou, S. Wang, Y. Jiang, K. Liu, F. Liu, H. Lu, and Y. Li, "Decentralized Robust Optimal Control for Modular Robot Manipulators Based on Zero-Sum Game with ADP," in *International Symposium on Neural Networks*. Springer, 2019, pp. 3–14.

[31] H. Modares, F.L. Lewis, and Z.-P. Jiang, "Hinf Tracking Control of Completely Unknown Continuous-Time Systems via Off-Policy Reinforcement Learning," *IEEE Trans. Neural Netw. Learn. Syst*., vol. 26, no. 10, pp. 2550–2562, 2015.

[32] J.C. Willems, "Dissipative Dynamical Systems. Part I: General Theory," *Arch. Ration. Mech. Anal*., vol. 45, pp. 321–351, 1972.

[33] D.J. Hill and P.J. Moylan, "Dissipative Dynamical Systems: Basic Input-Output and State Properties," J. Franklin Inst., vol. 305, no. 5, pp. 327–357, 1980.

[34] A.J. van der Schaft, "L2-gain Analysis of Nonlinear Systems and Nonlinear State Feedback Hinf Control," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 770–784, 1992.

[35] S. Boyd, L.E. Ghaoui, E. Feron, and V. Balakrishnam, *Linear Matrix Inequalities in System and Control Theory*. SIAM studies in applied mathematics: 15, 1994.

[36] S. Yasini, M.B.N. Sistani, and A. Karimpour, "Approximate dynamic programming for two-player zero-sum game related to Hinf control of unknown nonlinear continuous-time systems," *Int. J. Control Autom. Syst*., vol. 13, no. 1, pp. 99–109, 2014.

[37] W. Zylski, *Kinematics and dynamics of mobile wheeled robots*. Publishing House Rzeszow Univ. of Technology, 1996, [in Polish].

[38] J. Giergiel and W. Żylski, "Description of motion of a mobile robot by Maggie's equations," *J. Theor. Appl. Mech*., vol. 43, no. 3, pp. 511–521, 2005.

[39] J. Garca De Jaln, A. Callejo, and A.F. Hidalgo, "Efficient solution of Maggi's equations," *J. Comput. Nonlinear Dyn*., vol. 7, no. 2, 2012, doi: 10.1115/1.4005238.

[40] A. Kurdila, J.G. Papastavridis, and M.P. Kamat, "Role of Maggi's equations in computational methods for constrained multibody systems," J. Guidance Control Dyn., vol. 13, no. 1, pp. 113–120, 1990, doi: 10.2514/3.20524.

[41] DS1103, *Hardware Installation and Configuration*. dSpace, 2009.

[42] ActiveMedia, *Pioneer 2DX Operation Manual Peterborough*, 1999.